

THE CAMBRIDGE WORLD HISTORY OF
LEXICOGRAPHY

A dictionary records a language and a cultural world. This global history of lexicography is the first survey of all the dictionaries which humans have made, from the ancient civilizations of Mesopotamia, Egypt, China, India, and the Greco-Roman world, to the contemporary speech communities of every inhabited continent. Their makers included poets and soldiers, saints and courtiers, a scribe in an ancient Egyptian 'house of life' and a Vietnamese queen. Their physical forms include Tamil palm-leaf manuscripts and the dictionary apps which are supporting endangered Australian languages. Through engaging and accessible studies, a diverse team of leading scholars provide fascinating insight into the dictionaries of hundreds of languages, into the imaginative worlds of those who used or observed them, and into a dazzling variety of the literate cultures of humankind.

JOHN CONSIDINE is Professor of English at the University of Alberta. He is the author of *Dictionaries in Early Modern Europe: Lexicography and the Making of Heritage* (Cambridge, 2008), *Academy Dictionaries 1600–1800* (Cambridge, 2014), and *Small Dictionaries and Curiosity: Lexicography and Fieldwork in Post-Medieval Europe* (2017); he has edited or co-edited six other books on lexicography. He has contributed to the *Oxford English Dictionary* for the past thirty years, formerly as library researcher and as assistant editor, and now as a consultant.

THE CAMBRIDGE
WORLD HISTORY OF
LEXICOGRAPHY

★

Edited by
JOHN CONSIDINE



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of
education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107178861

DOI: 10.1017/9781316827437

© Cambridge University Press 2019

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2019

Printed and bound in Great Britain by Clays Ltd, Elcograf S.p.A.

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

NAMES: Considine, John (John P.) editor.

TITLE: The Cambridge world history of lexicography / edited by John Considine,
University of Alberta.

DESCRIPTION: Cambridge, United Kingdom ; New York, NY : Cambridge
University Press, 2019. | Includes bibliographical references and index.

IDENTIFIERS: LCCN 2018048304 | ISBN 9781107178861

SUBJECTS: LCSH: Lexicography – History.

CLASSIFICATION: LCC P327 .C343 2019 | DDC 413.028–dc23

LC record available at <https://lcn.loc.gov/2018048304>

ISBN 978-1-107-17886-1 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of
URLs for external or third-party internet websites referred to in this publication
and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

THE CAMBRIDGE WORLD HISTORY OF
LEXICOGRAPHY

A dictionary records a language and a cultural world. This global history of lexicography is the first survey of all the dictionaries which humans have made, from the ancient civilizations of Mesopotamia, Egypt, China, India, and the Greco-Roman world, to the contemporary speech communities of every inhabited continent. Their makers included poets and soldiers, saints and courtiers, a scribe in an ancient Egyptian 'house of life' and a Vietnamese queen. Their physical forms include Tamil palm-leaf manuscripts and the dictionary apps which are supporting endangered Australian languages. Through engaging and accessible studies, a diverse team of leading scholars provide fascinating insight into the dictionaries of hundreds of languages, into the imaginative worlds of those who used or observed them, and into a dazzling variety of the literate cultures of humankind.

JOHN CONSIDINE is Professor of English at the University of Alberta. He is the author of *Dictionaries in Early Modern Europe: Lexicography and the Making of Heritage* (Cambridge, 2008), *Academy Dictionaries 1600–1800* (Cambridge, 2014), and *Small Dictionaries and Curiosity: Lexicography and Fieldwork in Post-Medieval Europe* (2017); he has edited or co-edited six other books on lexicography. He has contributed to the *Oxford English Dictionary* for the past thirty years, formerly as library researcher and as assistant editor, and now as a consultant.

THE CAMBRIDGE
WORLD HISTORY OF
LEXICOGRAPHY

★

Edited by
JOHN CONSIDINE



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of
education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107178861

DOI: 10.1017/9781316827437

© Cambridge University Press 2019

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2019

Printed and bound in Great Britain by Clays Ltd, Elcograf S.p.A.

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

NAMES: Considine, John (John P.) editor.

TITLE: The Cambridge world history of lexicography / edited by John Considine,
University of Alberta.

DESCRIPTION: Cambridge, United Kingdom ; New York, NY : Cambridge
University Press, 2019. | Includes bibliographical references and index.

IDENTIFIERS: LCCN 2018048304 | ISBN 9781107178861

SUBJECTS: LCSH: Lexicography – History.

CLASSIFICATION: LCC P327 .C343 2019 | DDC 413.028–dc23

LC record available at <https://lcn.loc.gov/2018048304>

ISBN 978-1-107-17886-1 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of
URLs for external or third-party internet websites referred to in this publication
and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

Contents

List of Contributors page ix
Editor's Acknowledgements xi

Introduction 1

PART I

THE ANCIENT WORLD 9

1 · Ancient Mesopotamia 11
NIEK VELDHUIS

2 · Ancient and Coptic Egypt 36
FRANK FEDER

3 · Ancient China 51
FRANÇOISE BOTTÉRO

4 · Ancient India 67
LATA MAHESH DEOKAR, WITH JEAN-LUC CHEVILLARD

5 · The Greco-Roman World 84
ROLANDO FERRI

PART II

THE PRE-MODERN WORLD 107

6 · China, c. 600–c. 1700 109
NATHAN VEDAL

Contents

7 · India and Tibet, c. 500–c. 1750	130
LATA MAHESH DEOKAR, WITH JEAN-LUC CHEVILLARD	
8 · Arabic to c. 1800	159
RAMZI BAALBAKI	
9 · Hebrew to c. 1650	182
AHARON MAMAN	
10 · The Chinese Periphery to c. 1800	202
MÅRTEN SÖDERBLOM SAARELA	
11 · The Turkic Languages and Persian to c. 1700	223
MAREK STACHOWSKI	
12 · Byzantine Greek	247
STEFANO VALENTE	
13 · Medieval Latin Christendom	267
JOHN CONSIDINE	
14 · Early Modern Western and Central Europe	290
JOHN CONSIDINE	
PART III	
THE MODERN WORLD: CONTINUING TRADITIONS 315	
15 · China from c. 1700	317
HENNING KLÖTER	
16 · Japanese, Korean, and Vietnamese from c. 1800	340
DAVID LURIE, HEOKSEUNG KWON, AND JOHN D. PHAN	
17 · Turkish and Persian from c. 1700	366
LUCIANO ROCCHI AND ARTHUR DUDNEY	
18 · South Asia from c. 1750	388
WALTER HAKALA AND LISA MITCHELL	

Contents

- 19 · Arabic from c. 1800 414
JAN HOOGLAND
- 20 · Modern Hebrew 431
TSVI SADAN
- 21 · The Slavic and Baltic Languages 445
RICK DERKSEN
- 22 · The Germanic Languages Other than English from c. 1700 460
ULRIKE HAB
- 23 · Standard Varieties of English from c. 1700 484
CHARLOTTE BREWER
- 24 · Regional Varieties of English 509
MICHAEL ADAMS
- 25 · The Romance Languages from c. 1700 530
PASCALE RENDERS
- PART IV
- THE MODERN WORLD: MISSIONARY AND
SUBSEQUENT TRADITIONS 553
- 26 · Missionary Traditions in South America 555
OTTO ZWARTJES
- 27 · Missionary Traditions in Mesoamerica 579
OTTO ZWARTJES
- 28 · Missionary and Subsequent Traditions in North America 597
WILLEM DE REUSE
- 29 · Missionary Traditions in East Asia 614
OTTO ZWARTJES
- 30 · European Traditions in India and Indonesia 634
TOON VAN HAL

Contents

31 · Missionary and Subsequent Traditions in Africa	658
GONÇALO FERNANDES	

32 · Missionary and Other Traditions in Australia	682
WILLIAM B. MCGREGOR	

<i>Appendix 1 The Language Varieties</i>	706
--	-----

<i>Appendix 2 The Lexicographers</i>	738
--------------------------------------	-----

<i>Primary Sources</i>	806
------------------------	-----

<i>Secondary Sources</i>	863
--------------------------	-----

<i>Index</i>	923
--------------	-----

Contributors

MICHAEL ADAMS (Indiana University)
RAMZI BAALBAKI (American University of Beirut)
FRANÇOISE BOTTÉRO (CNRS, Paris)
CHARLOTTE BREWER (Hertford College, Oxford)
JEAN-LUC CHEVILLARD (CNRS, Paris)
JOHN CONSIDINE (University of Alberta)
LATA MAHESH DEOKAR (Savitribai Phule Pune University/Philipps University, Marburg)
WILLEM DE REUSE (University of North Texas)
RICK DERKSEN (independent scholar, Leiden)
ARTHUR DUDNEY (University of Cambridge)
FRANK FEDER (Academy of Sciences and Humanities, Göttingen)
GONÇALO FERNANDES (University of Trás-os-Montes and Alto Douro)
ROLANDO FERRI (University of Pisa)
WALTER HAKALA (University at Buffalo, SUNY)
ULRIKE HAß (University of Duisburg-Essen)
JAN HOOGLAND (Radboud University, Nijmegen)
HENNING KLÖTER (Humboldt University of Berlin)
HEOKSEUNG KWON (Seoul National University)
DAVID LURIE (Columbia University)
AHARON MAMAN (Hebrew University of Jerusalem)
WILLIAM B. MCGREGOR (Aarhus University)
LISA MITCHELL (University of Pennsylvania)
JOHN D. PHAN (Columbia University)
PASCALE RENDERS (Université de Liège)
LUCIANO ROCCHI (University of Trieste)
TSVI SADAN (Bar-Ilan University)
MÅRTEN SÖDERBLOM SAARELA (Academia Sinica)

List of Contributors

MAREK STACHOWSKI (Jagiellonian University, Kraków)
STEFANO VALENTE (University of Hamburg)
TOON VAN HAL (KU Leuven)
NATHAN VEDAL (Washington University in St Louis)
NIEK VELDHUIS (University of California, Berkeley)
OTTO ZWARTJES (Université de Paris)

Editor's Acknowledgements

First and foremost, I want to thank the contributors to this volume. We worked together to bring a large project to completion in good time and in a very friendly spirit. The appendices of languages and lexicographers are our joint work. Every contributor fielded questions from me, of one sort or another. One example may stand for all: confronted with a Turcological question at four o'clock one afternoon, I gathered that expert opinion on the matter at hand was at variance. So I sent a long email to Marek Stachowski in Kraków, asking his advice. Although four in the afternoon in western Canada is midnight in Poland, Marek sent me a learned and detailed answer within thirty minutes. Editing this volume has been an impressive experience.

I have turned with other questions to correspondents across the world, and am most grateful to Janet Afary (University of California, Santa Barbara), Adam Balogh (Törwe, Germany), Johannes Bronkhorst (University of Lausanne), Éva Buchi (ATILF, Nancy), Tim Buckwalter (Qamus LLC), Eleanor Dickey (University of Reading), Dominic Goodall (EFEO, Pondicherry), Marc Greenberg (University of Kansas), Jürgen Hanneder (University of Marburg), Christoph Harbsmeier (University of Oslo), Edmund Herzig (Oxford University), †Ian Jackson (Berkeley, California), Monika Jaglarz (Jagiellonian Library, Kraków), Moshe Kahan (Ben-Gurion University), Eivind Kahrs (Cambridge University), Tomasz Kamusella (University of St Andrews), Han Lamers (University of Oslo), Todd Lawson (University of Toronto), Alexander Maxwell (Victoria University, Wellington), Robin Meyer (Oxford University), Eun Kyung Min (Seoul National University), Dilworth Parkinson (Brigham Young University), Francesca Schironi (University of Michigan), Wolfgang Schweickard (Saarland University), Mark Van Mol (KU Leuven), Massimo Verdicchio (University of Alberta), and James Zetzel (Columbia University). I owe particular thanks to the Inter-Library Loans staff at the University of Alberta.

I presented an outline of the project at the Henry Sweet Society Colloquium at Gargnano del Garda in 2015, and am grateful to all the members of the society who discussed the project with me at Gargnano and afterwards, and to Giovanni Iamartino, who convened the colloquium and invited me to speak there.

Linda Bree of Cambridge University Press first suggested that I might edit such a book as this, and shepherded the project from its first beginnings in 2014 until her retirement; since then it has been looked after by Bethany Thomas. I am very grateful to them, and to their colleagues Isobel Cowper-Coles, Liz Hanlon, Tim Mason, and Sarah Starkey. The two anonymous readers who commented on the prospectus of the volume for Cambridge University Press were most kind and encouraging, and I am grateful for their advice. Karen Anderson's sensitive copy-editing has made this a better book.

I learned from the conversation and example of my beloved father that it might be possible to think about human languages with a sufficiently wide-ranging curiosity to imagine the outline of a book such as this (which was one half of my activity as editor), and to read the contributions to it attentively (which was the other half).

I owe the most, as always, to my wife, Sylvia Brown.

Introduction

This is the first global history of lexicography. There are, I think, two reasons why no such book has been written before.

The first reason is that there have been so many lexicographical traditions in the world over the past five thousand years: hundreds if not thousands of languages have been documented in wordlists of some sort, and scores of them have been documented in wordlists so numerous, and often so large, that their individual traditions are almost ungraspable by a single historian. There have been global bibliographies of wordlists since the eighteenth century, and, since at least the time of William Marsden's *Catalogue of Dictionaries* (1796), some of these have presented the wordlists of each language in chronological order.¹ Information of historical value is naturally present in such bibliographies even when the order is not primarily chronological. In Wolfram Zaunmüller's *Bibliographisches Handbuch der Sprachwörterbücher*, the last part of the entry for each language is, where appropriate, an overview of early dictionaries in reverse chronological order, century by century, from the nineteenth as far back as the fifteenth. Likewise, Andrew Dalby's *Guide to World Language Dictionaries* gives some very useful outline information about dictionaries of the past two centuries, and occasionally ranges back further, for instance to du Cange on Byzantine Greek or to the seventeenth-century Kikongo dictionary published as *Le plus ancien dictionnaire Bantou* in 1928. But a bibliography is not a history and, indeed, contributors to the present volume were asked from the beginning to treat the bibliographical record selectively enough for their chapters to have a narrative rather than an enumerative quality.

Likewise, a great encyclopedia of the late twentieth century, the three-volume *Wörterbücher/Dictionaries/Dictionnaires* edited by Franz Josef Hausmann, Oskar Reichmann, Hans Ernst Wiegand, and Ladislav Zgusta,

¹ Marsden, *Catalogue of Dictionaries*, 89–152.

was rich in historical material (and is therefore cited extensively in the present work), as is the supplementary volume edited by Rufus Gouws and others. But an encyclopedia, like a chronological bibliography, is not a history, because it is not primarily in the business of narrative.

The present volume, by contrast, seeks to tell a story. A good way to read it would be as a story, starting at the beginning and going on to the end. And the story it tells is a truly global one: English is the language in which this book is written, and the lexicography of English plays an important part in the book, but the language to which the most chapters are devoted is Chinese (indeed, the commonest surname in the biographical appendix is *Lǐ* 李, with representatives from the third century BC to the present day), and, to give another example, after Indo-European, the language family of which most members are mentioned (albeit only in one chapter) is Pama-Nyungan, to which many of the languages of Australia belong. The story which this volume tells is also a long one: the lexicography of the twenty-first and twentieth centuries is widely surveyed, but sustained individual attention is paid to each of the five traditions which began more than two thousand years ago. No individual could write equally well about the emergence of the lexicography of Sumerian and about developments in the electronic dictionaries of modern Hebrew and Japanese: this book is the first collectively authored history of lexicography and, as such, it tells a story which has not been told with unity and authority before.

A second reason why there has been no global history of lexicography is that the concept of 'lexicography' is somewhat elusive. For the purposes of this volume, it means the making of lists of words and their equivalents or interpretations. This definition does not free us from all doubt. For one thing, not all makers of such lists have thought of themselves as engaged in an activity different from other ways of collecting and transmitting information: to put that another way, many people have done what we would call lexicography although they themselves had no word for 'lexicography'.² Another problem is that the distinction between the lexicographical and the encyclopedic is notoriously uncertain, and is, indeed, made differently in different learned traditions.³ A borderline instance is that of the distinction between dictionaries of synonyms, which bring words of related meaning together (they have been very important in some Indian traditions, and are treated here) and thesauruses of the sort which are primarily oriented

² For a case study, see Considine, 'History of the concept of lexicography'.

³ See, e.g., Haß, *Große Lexika*, 1–2.

towards mapping the relationships between concepts (Roget's *Thesaurus*, for instance, is not treated here). Yet another problem is that we may ask when a text – for instance, a poem about words, or a traveller's account of unfamiliar things and their names in a local language – becomes sufficiently list-like to be counted as lexicographical.⁴ Here again, different learned traditions operate differently: the typical European dictionary is made up of entries in which the lemma, or headword, is 'an obligatory building-block of the text', but, for instance, a classical Tamil dictionary in verse may include statements of equivalence between words in which no one word seems to have headword status.⁵ However, 'the making of lists of words and their interpretations' gives this volume a central subject. The word *dictionaries* might not have done so: many short lists of words are lexicographical, but can hardly be called dictionaries. So, this is a history of lexicography, not of dictionaries.

The volume falls into four parts: a first on the lexicographical traditions of the ancient world; a second on those which originated in the next thousand years, and on their continuations up to the seventeenth or eighteenth century; a third on those of the past two or three centuries; and a fourth on the traditions which originated in European missionary activity from the sixteenth century onwards. One alternative to reading the whole volume from beginning to end would therefore be to read synchronically, across a given part, to get a picture of the lexicographical activity which was taking place at a particular time. Since each part is divided into chapters on different languages, or on geographical and cultural areas, and these chapters have been designed to fall into sequences, a third way of reading would be to follow one thread in the whole story from period to period, looking, for instance, at the three chronologically sequenced chapters on Chinese lexicography as a series (supplemented, perhaps, with the chapter on missionary lexicography in East Asia). Some key dictionaries – the Sanskrit *Amarakośa*, the Greek *Lexicon* of Hesychius of Alexandria, the Spanish–Latin *Dictionarium* of Nebrija, the Chinese *Kāngxī zidiǎn* – appear, viewed from multiple perspectives, in multiple chapters, though the aim has of course been to avoid extended duplication of treatment.

Reading widely in the volume will reveal that different chapters are handled in different ways. This is a good thing for the reader, because the

⁴ The latter point is made in, e.g., Alston, *Bibliography*, XIV.vii.

⁵ Wolski, 'Das Lemma und die verschiedenen Lemmatypen', 363, 'Das Lemma ist ein obligatorischer Textbaustein des Wörterbuchartikels'; Chevillard, 'On a 1968 incarnation of the *Piṅkalam*', 19–21.

effect of thirty-two strongly isomorphic chapters would be monotonous. It is also a natural consequence of the great variety of material under discussion: an expository schema which would fit the printed dictionaries of contemporary France would not fit the lexicography of ancient Egypt nearly as well. And it reflects the perspectives of different contributors: one, for instance, may be particularly interested in the kinds of grammatical analysis done by lexicographers, and another in dictionaries as records of cultural contact. These perspectives could not all be equally weighted in a single chapter; the book as a whole gives an overview of different ways of doing the history of lexicography.

The first part begins in ancient Mesopotamia, where the oldest extant wordlists in the world were made, and then surveys the four other lexicographical traditions of the ancient world: those of Egypt – the Egyptian tradition, from the Old Kingdom to the death of Coptic, is the longest of all dictionary traditions – and of China, India, and the Greco-Roman world. (It is curious that one highly developed literate culture of the ancient world, that of the speakers of Hebrew, did not develop a lexicographical tradition, although learned Jews were doubtless well aware of Greek and Roman lexicography.)

The second part begins with chapters on two of the great traditions of lexicography which continued from the ancient world into this middle period: those of Chinese and of the languages of India. The next four chapters in this part remain based in Asia, surveying the Arabic and Hebrew traditions, those of the Chinese periphery, and those of the Turkic languages and of Persian, before the last three turn westwards, to the lexicography of the Byzantine world, that of medieval Latin Christendom, and that of early modern western Europe.

The third part of the book is the longest, and that is not surprising: the lexicographical production of the Eurasian intellectual world in the past three centuries or so has been enormous. Highly literate societies; print and, latterly, digital publishing technologies which have made the mass production of books easier than ever before; and a high degree of intercultural contact have all played their part. The first chapter in Part III presents the lexicographical tradition which, at the present day, has remained highly active for longest, that of the Chinese language as documented by its native speakers. It is followed by an account of Japanese, Korean, and Vietnamese lexicography from the beginning of the nineteenth century onwards, a period in which these were no longer languages of ‘the Chinese periphery’ but had their own complex and vigorous traditions. The next two chapters take our

gaze westwards and then southwards, to return to the lexicography of the Turkish and Persian languages, and then to survey that of South Asia. The next two address two Semitic traditions: the continuing one of Arabic and the revived one of Hebrew. Then the next five chapters turn to the languages belonging to four of the language families of Europe: one chapter on Slavic and Baltic together, three on Germanic, and one on Romance. The emphasis on Germanic is of course because English is a Germanic language, and one which is familiar to the readers of this book: so, after a chapter on the Germanic languages other than English, there is one on supra-regional varieties of English and one on regional varieties.

The fourth and final part of the main text of this book brings together seven chapters on missionary and subsequent lexicography across the world: in South America, Mesoamerica, and North America; in East Asia, India and Indonesia, and Africa; and in Australia. As recently as the late twentieth century, giving this much space to missionary lexicography in a global history might have seemed disproportionate, but the historical importance of missionary linguistics, including missionary lexicography, is now more widely appreciated than it used to be.⁶ Moreover, the chapters speak for themselves.

Two appendices follow: one giving a brief account of each of the languages of which the lexicography is discussed in this volume, and one giving a brief account of each of a generous selection of the lexicographers whose work is discussed. The latter is, as far as I know, the first biographical survey of lexicographers on a global scale.

Much has had to be omitted from this book. By no means, for instance, does it refer to all of the languages of which interesting dictionaries have been made. Many natural spoken languages have had to be omitted altogether, as in the cases of many of the Caucasian and Malayo-Polynesian languages, or to be represented only by references to part of their lexicographical tradition, as in the cases of Irish and Tibetan. In particular, recent dictionaries have had to be treated very selectively, with the result that, for instance, more is said in this book about the early lexicography of Malay than the modern lexicography of Malay and Indonesian.⁷ Dictionaries of signed languages have not been treated; nor of invented languages such as Esperanto and the languages of J. R. R. Tolkien's fictions. A chapter on learned lexicography in Europe after 1700 formed part of the initial plan of this book, but could not in the end be included.

⁶ For recent activity in the field, see Zwartjes, 'Historiography of missionary linguistics'.

⁷ For the latter, see Echols, 'Presidential address: dictionaries and dictionary-making', 16–24.

The bibliographical side of the study of lexicography has only been touched on lightly. So, for instance, little has been said of the writing media with which lexicographers have worked, though they have sometimes been mentioned for their effect on the lexicographical record, as in the contrasting cases of the highly durable clay on which Mesopotamian word-lists were inscribed, and the relatively short-lived palm-leaf manuscripts of the classical Tamil tradition. The relationship between lexicography and papermaking alone would be a topic for a fascinating chapter, starting with the advantages enjoyed by early Chinese lexicographers, to whom paper was available as a writing material from the first century AD onwards, and exploring the spread of paper across the literate world, for instance to Mesoamerica, where Antonio de Ciudad Real could fill two sacks with the discarded working papers of his Maya dictionary at the end of the sixteenth century.⁸

The relationship of lexicography and other kinds of language study has likewise been touched on very lightly. The making of dictionaries and of grammars often went hand in hand: the missionary lexicographers in South America who pointed out that the *Vocabulario* of a given language must be used in conjunction with the *Arte* which accompanied it, often in the same volume, exemplify a relationship of widespread importance. The making of dictionaries and of translations could likewise be presented as a single story: an account of lexicography in Anglo-Saxon England which includes the interlinear glossing of texts in Latin might do more justice to the subject than one which is confined to the listing of words.

To give a final example, the constraints of space have rarely permitted a lively picture of the experience of making dictionaries. The Urdu lexicographer Sayyid Aḥmad Dihlavī recalled that, in his youth,

a passion appeared in my heart for the science of language and the compilation of dictionaries. Although at first glance . . . it was considered pointless and unworthy of serious regard, nevertheless that irresponsible beloved spirit of passion made peace with me as one human does with another.⁹

Not that all the human dramas of lexicography lead to a good peace: 'More stunned than suicidal', wrote Robert Laughlin of the Smithsonian Museum, recalling a computational meltdown as he worked on the *Great Tzotzil*

⁸ For the early use of paper as a writing material in China, see Tsien, *Paper and Printing*, 86; for Ciudad Real and the dictionary of which 'los borradores llenaua dos costales', see Smith-Stark, 'Lexicography in New Spain', 28.

⁹ Translated by Hakala, *Negotiating Languages*, 127.

Dictionary, 'I trundled my rubbish back to Washington. My dictionary became known around the museum as *The Great Tzotzil Disaster*.'¹⁰ It was successfully published in the end, and Laughlin ended the acknowledgements of another dictionary on a merrier note, imagining a party to which everyone connected with the dictionary should be invited, from Calepino onwards: 'What fun it will be to toast to the long life of the *Diccionario grande* . . . You who read these lines are now associated, too, whoever and whatever you are. Thank you. Welcome to the party. Let it be great!'¹¹ As those words remind us, the experience of reading dictionaries has also had, for the most part, to be excluded from the present volume.

But these exclusions were inevitable if the *Cambridge World History of Lexicography* was to appear as a single volume, capable of being handled without discomfort, and capable of being read from end to end by a not abnormally chalcenterous reader.¹² Despite the omissions, the story of the history of lexicography is told more fully in the following pages than it has ever been told before in a single volume.

¹⁰ Laughlin, *Great Tzotzil Dictionary of San Lorenzo Zinacantán*, 13.

¹¹ Laughlin, in *Great Tzotzil Dictionary of Santo Domingo Zinacantán*, I.xiii.

¹² For *chalcenterous* and lexicography, see Gilliver, *Making of the Oxford English Dictionary*, 432–3.

PART I



THE ANCIENT WORLD

Ancient Mesopotamia

NIEK VELDHUIS

The history of Mesopotamian lexicography in cuneiform extends from the very beginning of writing (around 3200 BC) to the demise of cuneiform in the first centuries of our own era.¹ Lexical texts accompany cuneiform writing in all periods and all areas where this writing system was used. They have been found not only in Babylonia (southern Iraq) and Assyria (northern Iraq), but also in present-day Iran, Syria, Turkey, Israel, and Egypt.

One may think of lexical lists as something between a dictionary and an encyclopedia – they organize, transmit, and preserve knowledge. Lexical lists have drawn the attention of cuneiform scholars from the earliest days of Assyriology, because they explain how to read and understand the ancient writing system. As such, they were invaluable in the early days of decipherment, and they still fulfil that essential function in modern philological research. Lexical lists are the earliest scholarly genre in ancient Mesopotamia, and thus they may claim to be the earliest scholarly genre in the history of humanity. Lexicography, therefore, also plays an important role in discussions of the history of Mesopotamian scholarship and education.

In the secondary literature, Mesopotamian lexical lists have often been categorized under the concept ‘*Listenwissenschaft*’ (‘list-making as scholarship’), a derogatory term that was introduced by Wolfram von Soden in 1936.² In von Soden’s view, Mesopotamian scholarship did not really deserve that name – it was poor in content and hopelessly stagnant. In the past few decades, scholars have recognized the limitations of von Soden’s approach, which defined scholarship in a presentist way and was clearly inspired by the

¹ An early account of the history of the cuneiform lexical tradition can be found in Meissner, *Babylonien und Assyrien*, vol. II. More recent histories include Civil, ‘Lexicography’; Civil, ‘Ancient Mesopotamian lexicography’; Cavigneaux, ‘Lexikalische Listen’; and Veldhuis, *History of the Cuneiform Lexical Tradition*, of which the second is most suitable for non-specialists.

² Von Soden, ‘*Leistung und Grenze sumerischer und babylonischer Wissenschaft*’.

prevailing anti-Semitic intellectual atmosphere in Germany in the 1930s.³ Since this discussion has been sufficiently covered elsewhere, I will not go into the details of von Soden's paper or the influence it had on Assyriology and surrounding disciplines.

The main part of this chapter will be a concise chronological discussion of cuneiform lexicography over the more than three millennia of its history. Before doing so, it will be useful to introduce three distinctions that are important for discussing the thousands upon thousands of cuneiform tablets that form the primary material of this chapter: sign lists versus wordlists; school exercises versus editions; and one-dimensional versus two-dimensional lexical texts.

Vertical Organization: Sign Lists and Wordlists

On the most general level, cuneiform lexical lists may be divided into sign lists and wordlists. Sign lists are designed to document and teach the inventory of signs and their proper uses. The most important and most frequently attested type of sign list is the one that is organized by the form of the sign – with simple signs that consist of just a few strokes coming first. Sign lists may provide glosses to indicate the proper reading of the sign in Sumerian; the use of those same signs in Akkadian is rarely taken into account. Since most signs are polyvalent (they may represent multiple words, morphemes, or syllables in Sumerian), a sign may be repeated several times with different glosses. Each line is introduced by the item sign, a single vertical. Sign lists are rarely attested in the third millennium, but become common in the second and first millennia. Although sign lists may not be thought of as lexicographic in the common sense of the word, they form part and parcel of the Mesopotamian lexical tradition and cannot be ignored here.

Wordlists are known from the very beginning of cuneiform writing. Wordlists are primarily organized by semantics and therefore invite a thematic organization, such as lists of professions, birds, stones, and geographical names. Over the millennia the contents of the lists changed considerably, but the idea of grouping together words of a similar semantic content remained remarkably stable and was used until the very end of cuneiform culture. The thematic wordlist is by far the most frequent type of wordlist, but by no means the only one. So-called acrographic wordlists are

³ Hilgert, 'Von "Listenwissenschaft"'; Veldhuis, *History of the Cuneiform Lexical Tradition*, 21–2.

organized by first sign, not unlike modern alphabetical ordering. Some bilingual (Sumerian–Akkadian) wordlists are organized by the Akkadian terms, grouping together words that more or less use the same (Akkadian) root consonants. So-called group vocabularies list brief sequences of semantically related words (often in groups of three) in Sumerian with Akkadian translation. The overall organization of such wordlists is usually entirely opaque.

The nature of the cuneiform writing system allows for lexical compositions that fall somewhere in between wordlists and sign lists: wordlists that are organized by first sign, and sign lists in which each sign represents a Sumerian word, translated into Akkadian.

School Exercises Versus Editions

Many of the lexical compositions that are discussed in this chapter were used as exercises in scribal training. The tablets that are the product of such training are known as exercise tablets or school tablets, and they are often recognizable by their format. Neo-Babylonian exercises (c. 500 BC), for instance, tend to include six or eight very short (four- to eight-line) extracts from different literary and lexical compositions, separated by horizontal rulings. Old Babylonian exercise tablets (c. 1800) feature a model text on the obverse, written by a teacher or advanced pupil, with blank space to the right where the pupil could copy the extract. On the reverse of such tablets, an earlier exercise (one that the pupil already knew by heart) was copied without the benefit of a model. Another well-known Old Babylonian exercise type is the round tablet or lentil. Often such tablets feature a two- to four-line extract in a teacher's hand, repeated by a pupil.

Old Babylonian and Neo-Babylonian exercise tablets have in common that they frequently combine material from different sources and that they contain extracts from larger compositions. What exercise texts look like differs from one period to another (and from one stage of education to another), but often their format and layout are indicative of their use in scribal education. They are important because they provide clues for understanding how lexical (and other) texts figured in the practice of scribal training, and thus they allow us a glimpse beyond the bare lists of words or signs into the world in which these lists functioned. Most exercise texts were not meant to be preserved; their destiny was the clay bin in which exercise tablets were recycled.⁴

⁴ See Tanret, *Per aspera ad astra*, 143–51.

Editions or reference texts, on the other hand, were produced to be kept. These are often beautifully written multi-column tablets that contain one composition in its entirety. Reference texts may include a colophon at the very end which may provide information on the composition (its name, for instance), the copyist, the reason for copying, and the process of copying itself (for instance: ‘completed and collated’). Colophons that include names allow us to connect ancient knowledge to actual people. One type of colophon says, for instance,

Tablet of Iqīša, son of Ištar-šuma-ēreš
hand of Ištar-šuma-ēreš, his son.

Here the line between exercise tablet and reference tablet becomes blurry. In this case, Ištar-šuma-ēreš (who is named after his grandfather) wrote a tablet for his father, a well-known priest in early Hellenistic Uruk.⁵ Sons and pupils frequently copied learned texts, including lexical texts, for the tablet collections of their teachers or fathers (the two may often have coincided). In doing so, these junior scribes familiarized or refamiliarized themselves with the composition that they were copying, and thus they may be labelled exercises. At the same time, they were also producing a new edition. Although the difference between exercise tablet and edition or reference tablet (both terms are somewhat problematic) is important, we have to recognize that we do not always know the difference and that in some cases the distinction is futile.

One-Dimensional Versus Two-Dimensional

The earliest lexical lists from the late Uruk period (c. 3200 BC) were simple lists of words, without any explanation. This format remains the norm until the Old Babylonian period (early second millennium), with one notable exception: the bilingual lists from Ebla (c. 2350 BC). Among the c. 15,000 cuneiform tablets found at Ebla, which are mostly administrative in nature, there is a group of bilingual lexical texts, translating Sumerian or pseudo-Sumerian terms into the east Semitic dialect of the region. Unlike later lexical traditions, the translations are added directly after the main entry, not in a separate column.

In the first millennium, nearly all lexical texts are arranged in two or more columns. One of these columns represents the lemma itself, and the other columns provide different types of explanation. For wordlists, the left column

⁵ Farber, ‘Neues aus Uruk’; Veldhuis, *History of the Cuneiform Lexical Tradition*, 419–22.

contains the word or expression to be explained, usually in Sumerian, while the right column contains the Akkadian translation. There are occasional variants of this pattern. The 'Emesal wordlist' explains dialectal forms of Sumerian (the Emesal variety) first in regular Sumerian and then in Akkadian, resulting in a three-column format. Other three-column wordlists such as *Murgud* have a Sumerian–Akkadian–Akkadian order, where the third column gives a more common synonym for a rare or archaizing word in the middle column. Akkadian–Akkadian synonym lists are well known in the first millennium, but Akkadian–Sumerian is entirely unknown, even in lists such as *Nabnitu*, in which the Akkadian column is the organizing element.

Late second-millennium and first-millennium sign lists tend to display much more complex formats. The sign to be explained is usually found in the second column, which is often written in archaizing cuneiform, clearly distinguished from the rest of the text. To the left, in the first column, we find glosses that indicate the pronunciation of the sign in Sumerian. The third column includes one or more Akkadian translations of the sign, read as a Sumerian word. Some exemplars add a separate column with sign names (descriptive names of signs and sign combinations), in an unusual mix of Sumerian and Akkadian morphology.⁶ Multiple-column formats are also found, occasionally, in Middle Babylonian (c. 1400–1100) lexical texts from the so-called western periphery. In this period, the cuneiform writing system spread all the way to present-day Anatolia and Egypt (for the latter, see Chapter 2), and, as we shall see, we occasionally find sign lists (and a few wordlists) with additional translation columns in local languages such as Hurrian, Ugaritic, or Hittite. A few sign lists from the Hellenistic period add a separate commentary column in which the Sumerian and the Akkadian translation are subjected to complex hermeneutical analyses, or in which rare signs are illustrated with usage examples from traditional texts.⁷

The great majority of Old Babylonian lexical texts are school exercises in a simple one-column format, listing Sumerian words without further explanation. There is good evidence, however, that such texts were memorized in a bilingual format, where every Sumerian word was accompanied by an Akkadian translation. By this time, Sumerian had died out as a spoken language (exactly when Sumerian died out is still a matter of controversy).⁸

⁶ See Gong, *Namen der Keilschriftzeichen*.

⁷ Frahm, *Babylonian and Assyrian Text Commentaries*, 243–8; Veldhuis, *History of the Cuneiform Lexical Tradition*, 400–4.

⁸ See Sallaberger, 'Ende des Sumerischen'; Michalowski, 'Life and death of the Sumerian language'; Woods, 'Bilingualism, scribal learning, and the death of Sumerian'.

Scribal pupils, who copied and memorized these lists as part of the elementary curriculum, would not be able to make much sense out of them without translation. The point of the scribal exercise was to gain experience in reading and writing Sumerian, and so there was no need to write the Akkadian out. Occasional Akkadian glosses in otherwise monolingual Sumerian wordlists confirm the underlying bilingual character of these lists. Repeated (Sumerian) words, which are not infrequent in wordlists, make sense when we assume that they were accompanied by different Akkadian translations. Similarly, Old Babylonian sign lists very frequently omit the glosses. Again, those glosses were learned by heart but were not part of the writing exercise. We may conclude that in the early second millennium lexical lists existed in two formats: a written format, the format we have today, which often lacked any explanatory material, and a memorized format, in which each word and each sign was accompanied by translation (wordlists) or glosses (sign lists).

Invention (3200–2000 BC)

Archaic writing was developed in the late Uruk period, around 3200 BC. The great majority of archaic texts deal with the accounting of grain, fish, cattle, sheep and goats, agricultural personnel, metal products, and related topics.⁹ The Uruk period was one of rapid urbanization, with its associated increase in social complexity. The invention of a writing system is to be seen in the context of the development of standardized mass production and organized labour, described by Guillermo Algaze as the ‘domestication of human labor’.¹⁰ Writing provided the managers of the time with the means to keep track of raw materials, production, and payments in ways that are unimaginable in a purely oral society.

Approximately 10 per cent of the archaic corpus is not administrative in nature, and these tablets, some 1,500 in total, are usually considered to be lexical. This lexical corpus includes a number of more or less standardized compositions that are known in multiple exemplars, as well as a group of ad hoc exercises. The table indicates how many exemplars of each of the thirteen standardized compositions are known today, ranging from more than 200 (*Lu A*) to 5 (*Plants*). Each of these lists may be reconstructed to enumerate around 120 entries.

⁹ See Englund, ‘Texts from the late Uruk period’.

¹⁰ Algaze, ‘Initial social complexity’, 211–13.

Title	Uruk	Jemdet-Nasr	Unprovenanced
<i>Lu A</i> (professions)	185		16
<i>Vessels and Garments</i>	91	1	3 or 4
<i>Word List C</i> (Tribute)	56		
<i>Metal</i>	55		1
<i>Wood</i>	30		2
<i>Cattle</i>	24		1
<i>Officials</i>	23		
<i>Fish</i>	22		
<i>Cities</i>	17		
<i>Geography</i>	12	1	1
<i>Food</i> (numbers, bread, meat)	9		
<i>Birds</i>	6		
<i>Plants</i>	5		

The titles used in this table are conventional and only partly descriptive. Several compositions may be described as thematic wordlists, such as the list of professions, the list of wood (including trees and wooden objects), the list of metals (including metal objects), the list of birds, and the list of fish. Others are much more difficult to characterize with a single label. The list *Officials* includes many words for professional titles, but also has a section on points of the compass and division of time. The list *Vessels and Garments* includes a section on perfumes as well as numerous unclear entries. *Word List C* has plausibly been interpreted as an origin story, describing the founding of the city of Uruk.¹¹ We are still far away from a definite interpretation of this group of texts and their function in the development and transmission of writing in the late Uruk period. A few important points, however, have emerged over the last few decades. First, the corpus is not homogeneous and does not represent a single concept of what a lexical text is or how it should be organized. There does not seem to be a master plan behind it. Secondly, the lexicon covered in this corpus is closely tied to the general usage of writing in accounting. Archaic lexical texts are the first attempt to expand the usage of writing beyond accounting, and they thus lay the foundation for writing royal inscriptions, letters, and poetry in later periods. But initially they do not venture very far from their roots; by and large the themes covered in the archaic lexical lists are comparable to the vocabulary needed for the administration of labour and commodities of the time. As an

¹¹ See Civil, 'Remarks on AD-GI₄'.

example I quote the first few lines of the most frequently attested list, the list of professions (*Lu A*).¹²

1. NAMEŠDA	ruler
2. NAM ₂ KAB	vizier
3. NAM ₂ DI	adviser
4. NAM ₂ NAM ₂	counsellor
5. NAM ₂ URU _{at}	mayor

All of the translations here are very uncertain because these words appear only occasionally in the archaic administrative texts and are not known from later periods, presumably because social structure and official titles changed over the centuries.

The interpretation of the archaic lexical corpus has developed in two distinct schools of thought.¹³ One school saw the archaic lists as a direct reflection of the Sumerian understanding of the order of the world, providing these texts, essentially, with a cosmological significance. This line of thinking is primarily associated with von Soden's very influential 'Leistung und Grenze sumerischer und babylonischer Wissenschaft', but the argument may still be encountered in much more recent literature.¹⁴ A second school, largely in reaction to a theological overinterpretation of the available data, posited that the archaic lexical texts were simply school texts, used in the transmission of the new communication technology.¹⁵ Both interpretations have serious flaws. The selection of topics in the archaic lexical corpus (see the table), which has little on topography, nothing on gods or stars, but plenty on pottery and meat, seems to be a strong argument against an 'order of the world' approach. A counter-argument against the 'school texts' argument is that the lexical corpus contains much that is never used in contemporaneous administrative texts and that it was standardized at a very early stage, to be transmitted virtually unchanged for many centuries. The cultural

¹² The conventions for transliterating archaic cuneiform are somewhat different from those for later periods. All signs are rendered in capitals, to indicate that vocalizing these words is a rather hazardous undertaking. The sign names derive from the later usage of these same signs, distinguishing homophones with index numbers (such as NAM₂). Some signs have multiple variants which may either be allographs or distinct signs – in many cases we do not know. Such signs are distinguished with subscript letters, occasionally augmented with further numbers (as in URU_{at}). The conventions for transliterating archaic cuneiform and the history behind them are explained in Englund, 'Texts from the late Uruk period', 65–71.

¹³ Veldhuis, *History of the Cuneiform Lexical Tradition*, 53–9.

¹⁴ Westenholz, 'Thoughts on esoteric knowledge'; Glassner, *Writing in Sumer*.

¹⁵ For instance, Veldhuis, 'How did they learn cuneiform?'

significance of the texts in the corpus, therefore, is much larger than what a simple elementary literacy exercise would suggest. A more recent theory argues that the lexical texts may have functioned in scribal education not only in order to transmit technical knowledge (that is, writing), but also to create and circumscribe the group of people who legitimately employed the technique.¹⁶ The interpretation of the enigmatic *Word List C* as an origin story fits very well with the idea that these texts were the symbols of the new craft.

All through the third millennium, many new lexical compositions were introduced, introducing vocabulary that was more relevant to contemporary writing. These new lexical compositions are often known only in a single exemplar or had a relatively brief history of transmission. At the same time, the set of thirteen archaic lexical compositions kept being copied in the entire area where cuneiform was used until about 1800 BC, a period of one and a half millennia. It is doubtful that in the third (and early second) millennium these archaic compositions functioned in the acquisition of literacy, strictly speaking. Sumerian writing and vocabulary had, of course, evolved in many ways, and many of the entries in this corpus had become obsolete early in the third millennium. There is good evidence that the archaic wordlists were transmitted together with meta-text that clarified the proper reading and meaning of the entries. For instance the entries ‘sheep knife’, ‘cow knife’, ‘fish knife’, and ‘cucumber knife’ in the earliest version are rendered ‘knife for sheep slaughtering’, ‘knife for cow slaughtering’, ‘knife for fish splitting’, and ‘knife for cucumber splitting’ in a mid-third-millennium copy.¹⁷ There are good indications to suggest that the expanded entries were not improvised by whoever wrote this copy, but belonged to the knowledge that was transmitted with the written text.¹⁸ This mode of transmission suggests that these archaic lexical compositions defined the identity of elite scribes and the importance of the history of cuneiform writing for the cultural identity of the period.

The city of Ebla, close to the Syrian coast, has yielded a large number of cuneiform tablets dated to approximately 2350 BC. This group includes a large number of remarkable lexicographical texts. Ebla scribes imported many lexical texts known from Babylonia proper, but also developed their own lists, including a sign list with glosses (the earliest pronunciation glosses

¹⁶ Veldhuis, *History of the Cuneiform Lexical Tradition*, 57–9.

¹⁷ Gurney, ‘List of copper objects’.

¹⁸ Veldhuis, *History of the Cuneiform Lexical Tradition*, 94–5.

for Sumerian that we have) and a group of bilingual lists in which Sumerian is translated into a local east Semitic language. One may speculate that the Ebla scribes, far outside the traditional heartland of cuneiform writing, had more room to experiment with various types of lists than their brethren in Babylonia. Important as this development is, it seems to have had no impact on Babylonia proper, and bilingual lists were invented again, several centuries later, in Babylonia.

Innovation (2000–1000)

The Old Babylonian period (early second millennium) inherited the idea of lexical lists from earlier ages. The structure of the Old Babylonian lexical corpus, however, is radically different from what came before. The third-millennium lexical corpus is conservative, one-dimensional, and unstructured. The Old Babylonian corpus, by contrast, is variable and two-dimensional and has a curricular structure.

By the Old Babylonian period Sumerian had died out as a spoken language and had been replaced by Akkadian, a language that belongs to the Semitic family and is unrelated to Sumerian. Sumerian was still used for learned purposes and for temple rituals. Scribal education focused primarily on Sumerian.

Where third-millennium lists were characterized by (sometimes extreme) conservatism, Old Babylonian lexical texts freely add or omit items, adjust spellings, or change the order of entries. We may compare the section ^{dug}bur-zi (a kind of bowl) in three versions of the same list, from Nippur, Isin, and, perhaps, Sippar.¹⁹ The Nippur and Isin texts are approximately contemporary (second half of the eighteenth century); the Sippar text may be a little later.

Nippur	Isin	Sippar?
bowl	bowl	bowl
bowl for barley	large bowl	large bowl
large bowl	small bowl	small bowl
small bowl	sakar bowl	bowl of one small <i>sila</i> capacity
bowl of one small <i>sila</i> capacity	bowl for oil-sprinkled bread	bowl of half a <i>sila</i> capacity

¹⁹ *Digital Corpus of Cuneiform Lexical Texts*, Q000040, lines 293–9; P332826, IB 1622a + 1546 r iii 9'–15', courtesy of Claus Wilcke (see Sallaberger, *Babylonische Töpfer*, 44–5); P247858, CBS 1862 r ii 05'–16' (the provenance of this text is uncertain).

(cont.)

Nippur	Isin	Sippar?
bowl for blood	bowl for incense	bowl for incense
bowl for cereal	bowl for cereal	bowl for blood
bowl for incense		bowl for [...]
bowl for oil-sprinkled bread		bowl for spice miller
		multi-coloured bowl
		bowl for oil-sprinkled bread
		thin bowl

The passage demonstrates the type of variance that may be expected between Old Babylonian versions of the same lexical composition. Although the three versions differ significantly from each other, they clearly belong to the same lexical composition. The traditions from Isin and Nippur (two neighbouring cities) are close, but not identical. The ‘Sippar’ source has a longer list of *bur-zi* bowls; still, it may be understood as an elaboration of the Nippur/Isin text, not as an entirely independent treatment. The Nippur text as presented here is based on multiple exemplars, and, interestingly, the duplicates have variants among themselves: the lines ‘bowl for barley’ and ‘bowl for cereal’ appear in only one source each and are skipped by other exemplars.

The great majority of Old Babylonian lexical lists exist in this highly flexible mode. This variability not only exists within lexical compositions: there is also a proliferation of a wide variety of different types of lists. In addition to the thematic lists of nouns (to which the section quoted earlier belongs), there are various lists of human beings, a list of body parts, various acrographic lists (ordered by first sign), phrase books (expressions to be used in contracts), god lists, and sign lists and sign exercises of various levels of complexity. Some of these lists are attested all over Babylonia, whereas others are characteristic of certain areas. The Old Babylonian period is characterized by a burst of creativity in matters lexical.

Third-millennium lexical texts are, for the most part, one-dimensional lists of Sumerian words. The bilingual lists from Ebla are the exceptions that prove the rule – they found no following in the Mesopotamian heartland. By contrast, most Old Babylonian lists are designed in two dimensions, providing an explanatory column for the words and signs listed. An extract from an Old Babylonian sign list called *Ea* (after its first line) may illustrate this.²⁰

²⁰ *Digital Corpus of Cuneiform Lexical Texts*, P228700; UM 29–16-031, from Nippur.

1. ¶	ir	NIMGIR
2. ¶	ti-in	NIMGIR
3. ¶	[mi]-ir	NIMGIR
4. ¶	ib ₂	IB
5. ¶	da-la	IB
6. ¶	u ₄ -ra-aš	IB
7. ¶	un	UN
8. ¶	[k]a-[lam]	UN
9. ¶	ru-u ₃	RU
10. ¶	šu-u[b]	RU
11. ¶	i-la-[a]r	RU
12. ¶	ĝeš-pa	RU

Each line on this tablet is introduced by a single vertical, the item sign (here represented by ¶). The sign to be explained is provided in regular writing (here represented in capitals), and the gloss that provides the proper reading of the sign is in much smaller signs.

Most cuneiform signs have more than one reading. Lines 4, 5, and 6 explain the three main uses of the sign which is conventionally transcribed IB: as the syllable -ib- (primarily used in verbal morphology); in the word *dara* 'belt' (here, exceptionally, represented with /l/ instead of /r/); and in the name of the goddess of the earth, Uraš. The list does not explain the meaning and proper uses of each of these values; such knowledge may have belonged to the explanations by a teacher.

The extract is taken from one particular school text from Nippur. The list *Ea* is known from hundreds of Old Babylonian exemplars, mostly extracts, that show that the full list had almost 1,000 entries. Numerous exemplars of Old Babylonian *Ea* do not even include the glosses. The exercises of the Old Babylonian scribal school were primarily writing exercises, designed to drill the correct writing of Sumerian signs and words. The glosses, therefore, might as well be memorized rather than copied – copying them would not add to the student's skill in writing proper Sumerian.

Similarly, virtually all Old Babylonian thematic lists (such as the list of earthenware quoted above) are in Sumerian only. There is plenty of evidence that these lists were bilingual (Sumerian–Akkadian) in design – a few exemplars in fact preserve an Akkadian column or some Akkadian glosses. The existence of such a non-written column of Akkadian translations is evident, among other things, from the rather frequent appearance of duplicate entries, such as

sur₂-du₃^{mušen}
sur₂-du₃^{mušen}

falcon
falcon.²¹

The Sumerian word *sur₂-du₃^{mušen}* has two known translations in Akkadian (*surdû* and *kassûsu*), which is why all available Old Babylonian and later sources of the bird list repeat the entry. Similarly, the list of animals has four entries *eh* (parasite), each of them corresponding to a different Akkadian translation. These four Akkadian words (*uplu*, *nābu*, *kalmatu*, and *puršû'u*; sometimes also *sāsu*) are included in much later versions of the list of animals, but Old Babylonian versions simply list *eh* four times. The main reason to copy a thematic list is to learn how to write proper Sumerian. The Akkadian translations were presumably memorized but there was little reason to write them down.

The Old Babylonian lexical corpus is structured as an extensive set of textbooks to be used in (elementary) scribal education, which progressed along more or less the same lines across Babylonia. By contrast, third-millennium scribal education was probably organized (like other crafts) in a master/apprentice setting where the apprentice would learn the trade by watching and doing. Third-millennium lexical lists have little educational relevance; the corpus is largely structured around the archaic lists inherited from the late fourth millennium. The curricular structure of the Old Babylonian corpus may be visualized by distinguishing (somewhat artificially) between different stages of education, corresponding to increasing length of practice units:

Elementary education

Stage 1	Sign Exercises (Syllable Alphabet A or B; TuTaTi)
Stage 2	Name Lists
Stage 3	Thematic Lists
Stage 4	Advanced Lexical Lists
Stage 5	Proverbs and Model Contracts

Advanced education

Sumerian Literary Compositions

Stages 1 to 4 of elementary education mainly consists of lexical lists, supplemented in Stage 4 by mathematical and metrological tables. The lists

²¹ See Veldhuis, *Religion, Literature, and Scholarship*, 88, for this passage.

in Stage 1 focus on the correct execution and the syllabic values of an elementary set of signs. These lists contain mostly meaningless syllable combinations, simply to allow students to practise their skills on a basic set of signs. The list TuTaTi lists triples of signs with alternating vowels (u, a, i). The lists of names in Stage 2 then provide the first meaningful entries. There are various more or less standardized lists of names.²² Not all of these necessarily functioned in this initial stage of the curriculum, but names do play an important role among the initial exercises.

The first true lexical exercises appear in Stage 3 among the thematic lists. The list *Ura* is an encyclopedic list of nouns and noun phrases that became a staple of scribal education from the Old Babylonian period to the end of cuneiform literacy. In Old Babylonian Nippur, *Ura* had six chapters: (1) trees and wooden objects; (2) reed and reed objects, pottery and clay, hides and leather objects, metals and metal objects; (3) domestic animals, wild animals, meat cuts; (4) stones and stone objects, plants and vegetables, fish and birds, fibres and clothing; (5) names of fields, cities, bodies of water, stars and constellations, and kinds of rope; (6) food and drink. Each of the six chapters contains, on average, about 600 entries, for a total of approximately 3,600 lines. As discussed above, local versions of *Ura* differed considerably from the Nippur version, but the themes that were covered are largely the same in all known versions (it is possible, however, that some versions added temple names, a theme not included in the Nippur version).

The vocabulary in *Ura* includes not only independent words, such as ‘sheep’ and ‘goat’, but also noun phrases (noun plus qualifiers). The section ‘domestic animals’ in Chapter 3 begins as follows in the Nippur version:

fattened sheep
 good quality fattened sheep
 fattened sheep, shorn with a knife
 male sheep
 male sheep used as breeder
 grass-fed sheep

The section ‘sheep’ continues for a total of 106 entries, and is followed by another 16 lines on ‘ewe’. The ‘sheep’ section includes terminology for various breeds, colours, illnesses, sheep used for particular offering ceremonies, sheep eaten by a god (followed by sheep eaten by a lion, and sheep eaten

²² See Peterson, ‘Personal name lists’.

by a wolf), and so on. Because of the structure of the Sumerian language, where virtually all qualifiers follow the head noun, each of the 106 lines of the 'sheep' passage begins with the sign for sheep (transliterated *udu*), visually marking the section. Such visual clues are very common throughout *Ura* because of the use of graphic classifiers (called determinatives in Assyriology) put before or after a word (such classifiers are usually not pronounced). Thus, every entry in the list of trees and wooden objects (*Ura* 1) begins with the classifier 'wood', all bird names end with the classifier 'bird', all fish names end with the classifier 'fish', and so on.

The various advanced lexical lists in Stage 4 include the acrographic lists (named after their first lines: *Izi*, *Kagal*, and *Nigga*), a list of human beings (freely mixed with other types of entries on an associative basis), a list of body parts (*Ugumu*), multiplication lists, and metrological lists. Significantly, this stage also includes two important sign lists, named *Ea* and *Diri*. *Ea*, an extract from which was presented above, is a list of simple signs with glosses. *Diri* is a list of compounds, signs that consist of multiple simple signs. The sign DIRI (after which the list is named) consists of the sign sequence SI.A. The combination represents the word *dirig* which means 'to surpass'. The reading and meaning of such compounds cannot usually be deduced from the reading or meaning of the component signs.

Students who arrived at Stage 4 had already used many of the signs that they would encounter in *Ea* and *Diri*. These exercises, therefore, do not necessarily introduce new knowledge, but they systematize the knowledge that the students had already acquired. We may see the progression through the curriculum as approaching the Sumerian language and writing system from different angles, from concrete to abstract, in order to provide a deep understanding of the working and history of (Sumerian) cuneiform writing. Jay Crisostomo argues that the advanced lexical exercises (in particular the list *Izi*) foster a type of analogical hermeneutics that is central to cuneiform scholarship of the time.²³

At Stage 5 (Proverbs and Model Contracts), students used their knowledge, for the first time, in full sentences and coherent (brief) texts. The Model Contracts, of course, also provided them with knowledge they would need in scribal practice.²⁴ The so-called Proverbs are sayings, expressions, or brief fables in Sumerian, gathered in more or less standardized collections.²⁵ This stage provides the bridge to the advanced stage of education, where students copied a broad array of Sumerian literary texts.²⁶

²³ Crisostomo, *Translation as Scholarship*. ²⁴ *Old Babylonian Model Contracts* is an edition.

²⁵ See Alster, *Proverbs of Ancient Sumer*, and Alster, 'Some new Sumerian proverbs'.

²⁶ For the literary texts, see *Electronic Text Corpus of Sumerian Literature*.

Given the structure of the curriculum as outlined above, one might suggest that the primary goal of the lexical texts is to prepare students for studying the literary corpus, and to some extent that is certainly correct. Students, to whom Sumerian was a dead language, acquired knowledge of vocabulary, signs, and spellings that they needed to read, copy, and understand the legendary texts about kings of the past, the mythological stories about the beginning of time, the songs to kings and gods, and the more light-hearted debates between non-human actors (such as Hoe and Plough) that they would encounter in the final stage of their education.

This interpretation of the Old Babylonian lexical corpus runs into one significant problem. The vocabulary that was catalogued in this set of lexical texts went far beyond what one would need for understanding the literary corpus. An initial exploration indicates that approximately 30 per cent of the vocabulary in the lexical lists is used in the current corpus of Sumerian literature.²⁷ Recovery of more – or more complete – literary texts may well increase that number by one or two percentage points, but the fact remains that a large number of entries seem superfluous if we look at the lexical texts solely from the perspective of the literary corpus. A good example comes from the list of body parts *Ugumu*, which includes a relatively high number of common and very common words – for head, eye, mouth, tooth, arm, hand, finger, leg, foot, heart, kidney, and so on.²⁸ In its present reconstruction, this list has between 250 and 300 lines (towards the end the reconstruction becomes rather fragmentary). The word *ugumu*, the first entry of the list, means ‘my crown’; the text is unique among thematic lexical lists in that it consistently adds the first-person possessive to each entry. The list *Ugumu* is roughly organized from head to toe, with almost half of the entries concerning the face. Lines 55–78 deal with the eyes, beginning as follows:

igi-ĝu ₁₀	my eye
sig ₇ -igi-ĝu ₁₀	my eyebrow
ma-ad igi-ĝu ₁₀	my pupil
igi si-ĝu ₁₀	my . . .
giggi igi-ĝu ₁₀	the black of my eye
babbar igi-ĝu ₁₀	the white of my eye
ša ₄ igi-ĝu ₁₀	the inside of my eye

²⁷ This number is based on preliminary computations, comparing the vocabulary in digital editions of the Old Babylonian literary and lexical corpora. I will publish this research in the near future.

²⁸ See Couto-Ferreira, ‘Etnoanatomia y patonomía’.

na ₄ igi-ĝu ₁₀	my eyeball [the stone of my eye]
^d lamma igi-ĝu ₁₀	my pupil [the protective lamma deity of my eye]
murub ₄ igi-ĝu ₁₀	the middle of my eye
pa igi-ĝu ₁₀	my eyelid
gakkul igi-ĝu ₁₀	my eyeball
dim ₃ igi-ĝu ₁₀	the figurine in my eye
er ₂ igi-ĝu ₁₀	the tears in my eye
zar-bad ₃ igi-ĝu ₁₀	the redness in my eye

The literary corpus uses only a tiny portion of this vocabulary. Some of the words in this section may even be artificial creations. The entry *zar-bad₃ igi-ĝu₁₀* probably means ‘the redness of my eyes’, derived from Akkadian *šarāpu* ‘to be red’, which is used to describe eyes or eye conditions. The Sumerian form *zar-bad₃* is not found anywhere else – most likely it is invented to fill a gap in the existing Sumerian vocabulary. The list *Ugumu* is not an exception; all lexical lists include common words as well as rare or obsolete and artificial entries. This picture of the Old Babylonian lexical corpus is puzzling when we see this material mainly as preparation for studying a literary heritage in Sumerian. The lexical texts represent a value of their own, quite apart from their usefulness in studying literary texts.

Looking at the Old Babylonian scribal curriculum as a whole, it appears that the acquisition of literacy was a byproduct at the most. In studying the corpus of literary texts the students familiarized themselves with a Sumerian heritage; they looked back at a (largely imaginary) Golden Age, creating an imagined community of highly learned scribes.²⁹ Potent symbols of this imagined community were the Sumerian language and the cuneiform writing system. Sumerian, therefore, was worth studying for its own sake, and a deep knowledge of Sumerian vocabulary and Sumerian writing constituted social capital.

The Middle Babylonian Period

In the Middle Babylonian period (late second millennium BC), cuneiform writing spread over the entire ancient Near East from present-day Iran to Egypt and from Anatolia to Bahrain. In all these places lexical texts

²⁹ Cf. Anderson, *Imagined Communities*.

followed suit. The main lexical compositions that were created in the early second millennium continued to be used in scribal education, and they continued to be flexible to some extent. In many cases the bilingual format of the lists now became explicit. Lexical lists were written in two columns, with Sumerian on the left and Akkadian on the right. In places outside Babylonia and Assyria (Hattuša in Anatolia, Ugarit and Emar in Syria, or Amarna in Egypt) both Sumerian and Akkadian were foreign, but they represented the prestige and chronological depth of Babylonian scribal culture. Occasionally scribes added an additional column in a local language (Hittite, Ugaritic, or Hurrian). Such multilingual lists have attracted much attention, but they remain the exception: fewer than twenty exemplars are known today.

In Babylonia proper, several very extensive new lexical compilations were developed. The list *Nabnitu* may have had as many as fifty-four chapters (tablets), of which some twenty to twenty-five have been identified so far.³⁰ *Nabnitu* broadly follows the parts and activities of the human body from head to toe. Within each section the list collects words that use the same (or similar) root consonants. Thus Tablet 21 deals with bending and bowing (related to the middle part of the body), which is *gurum* (Sumerian) or *kanānu* (Akkadian). The first thirteen lines of the tablet collect various Sumerian expressions that may be translated by Akkadian *kanānu* and then continue with Sumerian words that equal Akkadian *kannu* ‘vessel stand’ and *qinnu* ‘bird nest’. Such etymological and pseudo-etymological associations are the main organizing principle of the individual chapters of *Nabnitu*; they allowed its compilers to include a very wide range of Sumerian–Akkadian equivalences.

Another important list that was introduced in this period and is equally innovative in its organization is *Erimhuš*.³¹ This list, which consisted of seven or eight tablets, collects series of three or four synonyms or near-synonyms, as follows:

erim-huš	<i>anantu</i>	battle
inbir	<i>ippīru</i>	struggle
zag nu-sa ₂ -a	<i>adammû</i>	unequalled = battle (Akkadian)

³⁰ Finkel (ed.), *Series SIG₇.ALAN = Nabnitu* is an edition; see also Edzard, ‘SIG₇.ALAN = Nabnitu-Liste’.

³¹ Cavigneaux et al. (eds.), *Series Erim-Huš = Anantu and An-Ta-Gál = Šaqû*, is an edition.

From the usage of *Erimhuš* in first-millennium commentary texts, it appears that all six expressions (Sumerian and Akkadian) were considered equivalent.³² In most cases the coherence of a three- or four-line section in *Erimhuš* is fairly easy to understand; the organization of the composition as a whole, however, is entirely unclear.

The synonym list *Malku* = *Šarru* ('king = king') is probably of Middle Babylonian or early first-millennium origin.³³ The list, which had five (Babylonia) or eight (Assyria) chapters, is organized in two columns, as are most other lexical lists of the period, but unlike other lists both columns are in Akkadian, providing synonyms for unusual or outdated terms.

The god list *An* = *Anum* organizes an inventory of almost 2,000 god names by assigning each god to a particular role in the family or entourage of one of the main gods.³⁴ In most cases the right column is not so much a translation of the left column, but rather indicates the relationship to the main god: 'his spouse', 'his son', 'his cook', and so on. God lists had existed since the mid third millennium, but this particular organization was new.

Relatively few lexical tablets from Babylonia proper date to the Middle Babylonian period.³⁵ Yet, *Nabnitū*, *Erimhuš*, *An* = *Anum*, and *Malku* = *Šarru* bear witness to the strong lexicographical creativity of this period. It is worth emphasizing that these new lexical compositions are not only new in content, but also use innovative ways to organize that content.

Consolidation (First Millennium BC)

The first millennium in Mesopotamia is the period of the large empires: the Neo-Assyrian, Neo-Babylonian, and Persian empires, followed by the Hellenistic period. Assyria, in northern Mesopotamia, did not have much of a literary history of its own. Although cuneiform writing had been used there for centuries, for their literary and scholarly heritage the Assyrians looked south, to their Babylonian neighbours. Babylonian scholarship and rituals were used and collected in Assyrian temples and palaces, often adapted only in the slightest of ways. Assyrian scribes and scholars did not try to hide the Babylonian origin of their handbooks – quite to the contrary, anything Babylonian seemed to be more trustworthy and valuable.

³² See Frahm, *Babylonian and Assyrian Text Commentaries*.

³³ Hruša, *Akkadische Synonymenliste malku = šarru*, 27–8.

³⁴ Litke (ed.), *Reconstruction of the Assyro-Babylonian God-Lists*, is an edition.

³⁵ See Bartelmus, *Fragmente einer grossen Sprache*.

The Assyrians, in adopting a foreign body of learning, introduced the notion of immutability. Scholarly texts became standardized (at least in theory) and were provided with a pedigree that went all the way back to the period before the Flood – or even to the gods themselves. In their attempt to preserve a Babylonian heritage, and preserve it exactly the way it was received, they thus changed a living tradition into a frozen body of texts.³⁶ The standardization of scholarly texts (often referred to in the literature as canonization) was a process that had its roots already in the Old Babylonian period, but received a decisive push in Assyria in the early first millennium.

Our evidence for Assyrian scholarship (lexical and otherwise) comes primarily, although not exclusively, from the realm of the king and his entourage. The collections of cuneiform tablets at the royal palace are related to the image that the king wanted to project of himself – the image of a king steeped in the traditional wisdom that enabled him to understand the will of the gods. Such collections primarily consist of omen compendia of various kinds, including astrology and extispicy, but also contain medical texts, literary compositions, and lexical compendia.

The most famous of these collections is the so-called Library of Assurbanipal, in fact a group of at least three libraries, located on the citadel of Nineveh, the residence of King Aššurbanipal (668–627 BC). The mid-nineteenth-century excavations at Nineveh yielded some 25,000 cuneiform texts, now in the British Museum, including approximately 1,000 lexical texts. The library (or libraries) of Assurbanipal at Nineveh, arguably the largest library of the ancient world, was assembled with the explicit intention to be comprehensive. Assyrian rulers were avid collectors of all kinds of things, such as exotic animals, trees, luxury furniture, and cuneiform tablets.³⁷ The centre of the empire sampled both the man-made luxuries and the natural features of the entire world for the pleasure of the king, exhibiting his absolute power. The collections of cuneiform tablets emphasized the chronological depth of the knowledge of the king and the tradition on which it was based. The professional mythology of the expert scribes of the time traced their text corpora back in time to the period of the *apkallus* who received their wisdom directly from the gods in the period before the Flood.³⁸ Lexical texts demonstrated their hoary antiquity by their use of the ancient Sumerian language.

³⁶ Veldhuis, 'Domesticizing Babylonian scribal culture'.

³⁷ Garrison, 'Antiquarianism, copying, collecting'; Thomason, *Luxury and Legitimation*.

³⁸ See Lenzi, *Secrecy and the Gods*.

The Nineveh lexical corpus included all known types of wordlists and sign lists in a standardized format. The thematic list *Ura*, which in the Old Babylonian period had six chapters, had expanded to twenty-four tablets. The topics treated in this series were still more or less the same, but the vocabulary had increased considerably (to include many questionably Sumerian words), and all exemplars had become bilingual Sumerian–Akkadian. Similarly, the sign list *Ea*, which had about 1,000 entries in the Old Babylonian period, was now divided over 8 tablets for a total of approximately 2,400 lines. Each entry has not only a gloss, but also an Akkadian rendering and in some exemplars a sign name. The expanded version of *Ea* (conventionally referred to as *Aa*) has no fewer than forty-two tablets in its first-millennium incarnation, but this series is only very partially preserved. The Middle Babylonian lexical series *An* = *Anum*, *Nabnitū*, *Erimhuš*, and *Malku* = *Šarru* are well represented in the Nineveh collection. The lexical corpus of the time consolidated all the lexical material available, of both Old Babylonian and Middle Babylonian origin. More than any other ancient Mesopotamian tablet collection, this library valued comprehensiveness. The vocabulary of these lexical compilations had fairly little relevance for contemporary scribal practice, but its antiquity, perceived or real, was what was valued.

Tablets from the Assurbanipal library often include colophons that provide us some idea of the meaning of these texts for their copyists and owners. An interesting example is K 2016a+, which contains the fourth tablet of the thematic series *Ura* (wooden furniture and boats).³⁹

Fourth tablet of the series *Ura* = *hubullum*.

For the consultation by Assurbanipal, the crown prince
of the succession house of Esarhaddon the king of the world, king of the
land Assur,

governor of Babylon, king of the land Sumer
and Akkad, Aplāya, the junior apprentice scribe
the son of Kēni the scribe of the crown prince wrote it
and provided it for the crown prince his lord as a prayer.

This passage refers to the education of both the crown prince Assurbanipal and Aplāya, the junior apprentice. Aplāya was advanced enough in his studies to be asked to copy Tablet 4 of *Ura* in its entirety and in a quality that was suitable for use by the crown prince. Aplāya's

³⁹ *Digital Corpus of Cuneiform Lexical Texts*, P289805.

connection to the crown prince ran through his father Kēni, who was scribe of the crown prince, and who may have instructed Assurbanipal in scribal and scholarly matters.

Lexical texts were not only collected and stored; they were also used in education, as they had been in previous centuries. Our evidence for the educational usage of lexical texts in Assyria is rather meagre, but large numbers of school texts are known from the Neo- and Late Babylonian periods. A typical Neo-Babylonian school tablet contains brief extracts from literary and lexical compositions.⁴⁰ An example is CBS 8801 from Nippur.⁴¹ The obverse of this tablet contains brief extracts from interlinear bilingual (Sumerian–Akkadian) incantations.⁴² The extracts are so brief that one does not get much sense of what is going on. Indeed, the first extract starts in the middle of one sentence and ends in the middle of another:

may they cut off his wings.
 The one who leans through the window
 may they slaughter him by the neck!
 The one who looks through the side window

Each of these incantations is known from other exemplars so that the modern reader can reconstruct the context of these lines. For the ancient pupil, the approach to these texts seems rather fragmentary.

The reverse has passages from the thematic series *Ura 3* (trees) and *Ura 4* (furniture). Here the text of the passage from *Ura 3* is laid out in columns (Sumerian left, Akkadian right).

mes poplar	poplar from Crete
mes poplar	zanzaniquum tree
ilru tree	ditto
illuru tree	ditto
zanzanikkum tree	ditto
willow	willow
dark willow	dark willow
fate tree	mandrake
tree of fate	ditto

⁴⁰ See Gesche, *Schulunterricht in Babylonien*.

⁴¹ *Digital Corpus of Cuneiform Lexical Texts*, P263622.

⁴² These are *Saġba I* 69–76 and *Saġba II* 61–66 (edited in Schramm, *Bann, Bann! Eine sumerisch-akkadische Beschwörungserie*), and extracts from a ritual against the Asak demon (edited in Schramm, *Compendium sumerisch-akkadischer Beschwörungen*) and from the incantation series *Udughul III* 124–5 (edited in Geller and Vacin (eds.), *Healing Magic and Evil Demons*).

fate tree	ditto
tree that carries fate	ditto
billa tree	ditto
billa tree plant	ditto

The extract from *Ura* 3 demonstrates how first-millennium *Ura* tends to collect all possible Sumerian equivalencies of a single Akkadian word. Such sequences have ‘ditto’ in the Akkadian column; the Sumerian column may include actual synonyms (or near-synonyms), but also variant writings of the same Sumerian word, or loans from Akkadian.

There are hundreds if not thousands of such school texts from first-millennium Babylonia. Most exercise texts of this type have a larger number of much shorter extracts. BM 36726, for instance, has a sequence of three literary extracts (including a passage from the creation epic *Enūma Eliš*) and five lexical extracts with five or six lines each from *Ura* 7B, 8, 9, 10, and 11.⁴³ Such sequences of extracts are always taken from consecutive chapters of a lexical series. Another example is BM 54203, which, after extracts from a bilingual incantation and a prayer to Marduk, includes extracts from *Ura* 10–17 as follows:

<i>Ura</i> 10 (earthenware)	421–7
<i>Ura</i> 13 (domestic animals)	76–81
<i>Ura</i> 14 (wild animals)	48–53
<i>Ura</i> 15 (meat cuts)	13–18
<i>Ura</i> 16 (stones)	9–13
<i>Ura</i> 17 (plants)	1–5. ⁴⁴

It appears that students wrote series of such extracts, working their way through several tablets of *Ura* at the same time, while gradually adding higher tablet numbers. The student who wrote BM 54203 had already worked his way through most of *Ura* 10, and presumably he had finished 11 and 12, which do not appear in this exercise. He had made some progress with 13 and 14, and had only recently started with *Ura* 15 and 16. This is the very first exercise in which he encountered the list of plants in *Ura* 17. In a few cases we can identify sequences of such school texts in which each new exercise picks up the extracts where the previous exercise ended.⁴⁵

⁴³ Gesche, *Schulunterricht in Babylonien*, 279. ⁴⁴ Gesche, *Schulunterricht in Babylonien*, 387.

⁴⁵ Veldhuis, *History of the Cuneiform Lexical Tradition*, 414.

Ura is not the only lexical list that is excerpted in these exercise tablets, but it is by far the most frequently attested one. Exercise texts and other types of data allow us to reconstruct the extent and order of the (ideal) Neo-Babylonian curriculum.⁴⁶ This curriculum may be summarized as follows:

<i>Syllabary A</i> (S ^a)	sign list	Elementary
<i>Syllabary B</i> (S ^b) 1–2	sign list	
<i>Weidner God List</i>	traditional god list	
<i>Ura</i> 1–2	business expressions	
<i>Ura</i> 3–24	thematic lists	Intermediate
<i>Lu₂</i> 1–2	list of professions	
<i>Malku</i> = <i>Šarru</i> 1–5	Akkadian synonyms	
<i>Erimhuš</i> 1–7	wordlist	
<i>Diri</i> 1–7	compound signs	Advanced
<i>Ea / Aa</i>	sign list	
<i>An</i> = <i>Anum</i>	god list	
<i>Nabnitū</i>	wordlist	

The standardization of the scholarly and literary tradition in the first millennium encouraged scholarly explanation of texts that were difficult to understand or were suspected to have a deeper, hidden meaning. Commentary texts provided explanations and interpretations of literary compositions, divinatory texts, and lexical texts. The lexical tradition of sign lists and wordlists provided the main material for the hermeneutic techniques developed by the scholars of the time, allowing them to use rare or obsolete sign readings and interpretations of Sumerian words to make sense of their source texts.

First-millennium cuneiform lexicography was a relatively conservative area of knowledge production that largely depended on the reproduction of earlier collections and stands in marked contrast to the lexicographical innovations of the second millennium. Yet, these traditional Babylonian lexical texts are part of the same scholarly milieu in which we find the latest in celestial sciences, mathematical astronomy, the zodiac, and horoscopy.⁴⁷ The restrictions of the scope of this chapter should not blind us to the multifaceted and multicultural aspects of a vibrant intellectual culture.

⁴⁶ Veldhuis, ‘Purity and access’. ⁴⁷ See Rochberg, *Heavenly Writing*.

Conclusions

From the beginnings of writing in the late fourth millennium to the latest usage of cuneiform in the Parthian period, cuneiform literacy was always accompanied by lexical texts. These lexical texts were in part school texts, in part cultural heritage. Cuneiform lexical texts have attracted much attention from Assyriologists because they provide an entry into the vocabulary of Sumerian, a linguistic isolate. Lexical texts still provide this essential function for cuneiform studies today. In addition, lexical texts may be studied for their contribution to intellectual history, and the history of education and scholarship. They reveal an intense scrutiny of the cuneiform writing system, its structure, and its history.

Editions of Primary Texts

Many cuneiform lexical texts have been published in the series *Materials for the Sumerian Lexicon*. This series was started by Benno Landsberger in 1937 under the title *Materialien zum sumerischen Lexikon* and at present has nineteen volumes with one more volume announced. MSL (as it is usually abbreviated) includes mostly composite text editions without translation. It is a true monument of scholarship that provides the foundation for all (modern) lexicographical work.

Since 2003 the Digital Corpus of Cuneiform Lexical Texts (DCCLT) has provided editions with translations, glossaries, and links to photographs of lexical texts of all periods. Initially, DCCLT focused on areas where MSL was weakest, namely the Old Babylonian period and the third millennium. Today, many first-millennium lexical texts have been incorporated into DCCLT, but MSL is still much more complete.

Ancient and Coptic Egypt

FRANK FEDER

Ancient Egypt and the Egyptians have been renowned for their culture and literacy since classical antiquity.¹ The storerooms of knowledge and literature were the libraries of the temples from the third millennium BC down to the fourth century AD. We distinguish three periods of ancient Egyptian history in the third and second millennia BC: the Old Kingdom (c. 2657–2120), the Middle Kingdom (c. 2119–1794), and the New Kingdom (c. 1550–1069).² A period of foreign rule followed, and the pharaonic culture and religion were gradually Hellenized under the rule of Alexander and (after 332 BC) his successors, the Ptolemaic dynasty, and in the Roman empire, and finally replaced by Christendom. We can easily speak of 4,000 years of written culture in the Nile Valley if we include the Christian or Coptic period, which ended when the Ancient Egyptian language, of which Coptic is the latest offspring, disappeared and was finally replaced by Arabic in the fourteenth century AD. This immense time period is the frame or scope of this chapter.

Although Egypt possesses the longest documented language history of mankind so far known – perhaps China will surpass it one day – and temples and tombs, but also numerous papyri, ostraca, and the like, still bear witness to it today, we must admit that only a very small part has been transmitted to us. There is, unfortunately, also a geographical imbalance between upper and lower Egypt, because of the more humid climate of the Nile Delta that made the conservation of texts, particularly on papyri, much more difficult. This is a huge disadvantage for our knowledge about lower Egypt as some of the most important cities and cultural centres, especially of the first millennium BC, were situated in the delta.

¹ The most comprehensive collection of all genres of Egyptian texts, lexically annotated and translated (mostly into German, but also into English and French) can be found in the *Thesaurus Linguae Aegyptiae*.

² Beckerath, *Chronologie*, 187–90.

It may be that the pyramids are the most impressive signs of Egyptian culture. However, no proper temple survives from the third millennium BC, and the written record, apart from the tomb inscriptions, is sparse. This record improves through time, and so we possess many more texts from the later periods, namely from the second half of the second millennium BC down to the era of Greco-Roman rule over Egypt.

What the priests, as guardians of writing and literacy, kept and transmitted for such a long time was predominantly the *materia sacra*: the mythological and cosmological genesis of the Egyptian pantheon, ritual texts for the daily service, names of holy objects and explanations of them, and all the data, often differing locally, of a very elaborated religious topography. Of course, *belles lettres* – stories and poetry – and the economic life of the temple community also had their place in the temple libraries. Texts which we would consider scientific literature, such as medical, mathematical, or astrological texts – unless, like some mathematical texts, they served purely economic purposes – were regarded by the Egyptians as part of the *materia sacra*. Thus, a medical treatment, although it may have had features in common with our medicine, was an act of religious ritual charged with magical power, and the person administering it was naturally a priest.

The texts of a temple library were preferably written on papyrus scrolls and stored in special rooms within the temple or in certain buildings in the neighbourhood of the temple complex.³ The Egyptians called such institutions, which may be compared to our libraries and archives, ‘house of books’ and ‘house of life’.⁴ Particularly, the ‘house of life’ seems to have had a special significance not only for the storage but also for the production of books. As it is improbable, for obvious reasons, that we will find exercises in lexicography among the tomb inscriptions or in the texts of the temple reliefs, we will have to rely on the material that was kept in the libraries of the temples; in archives of the administration; and, perhaps, in private collections deposited in a tomb for entertainment in perpetuity.

The Old Kingdom (c. 2657–2120 BC)

The first approaches made by the Egyptians to classify their lexicon that we know about today were thematic compilations. The same tendency can be

³ Almost the only known example of such a room in a temple is the ‘house of books’ in the temple of Edfu: see Burkard, ‘Bibliotheken im alten Ägypten’, 95–6. See also Ryholt, ‘Libraries in ancient Egypt’.

⁴ Burkard, ‘Bibliotheken im alten Ägypten’, 85–90.

observed in Mesopotamia (see Chapter 1).⁵ Although the hieroglyphic script was always a combination of phonetic and logographic signs, it used a special category of signs, which had in this function no phonetic value: the classifiers.⁶ These served to categorize lexemes according to semantic classes, in this way helping to identify groups of hieroglyphic signs as words. So, the natural approach for early Egyptians studying language was classification and encyclopedism: a description of the world they lived in, mirrored also in the hieroglyphic signs. And the natural form for such a classification was a grouping of the signs and words in tables according to semantic classes.

There are even examples of such tables from the remote era of the pyramids, the Old Kingdom. On a writing board discovered in a tomb of the famous Giza Necropolis near the three great pyramids we find, organized in sections, names of kings, gods, and places as well as hieroglyphic drawings of birds and fish.⁷ This kind of thematically arranged list became the most common lexicographic form in Egypt. However, despite the great number of such lists that must have existed, only a few examples seem to have survived, particularly from this early age.

The Middle Kingdom (c. 2119–1794 BC)

We have to move forward through time to find other noteworthy compositions of this kind. One relatively large papyrus dating to the late Middle Kingdom was found in a tomb at Thebes, beneath the mortuary complex built centuries later, in the nineteenth dynasty, by Ramses II (1279–1213 BC), the so-called Ramesseum. Alan H. Gardiner, who published this and other texts of this genre, termed them ‘Onomastica’.⁸ The *Ramesseum Onomasticon* contains 321 items in lists of liquids, plants, birds, fishes, cattle, quadrupeds, cereals, place-names (fortresses in Nubia and towns in Egypt), parts of the body of an ox, and so on.

The New Kingdom (1550–1069 BC)

The onomastic genre continued to be the preferred means for lexical categorization. Another famous example is the *Onomasticon of Amenemope*, which is actually preserved in nine sources of different origin. It was compiled by the

⁵ See also Boisson, Kirtschuk, and Béjoint, ‘Aux origines de la lexicographie’, 267–70; Osing (ed.), *Hieratische Papyri aus Tebtunis I*, 33 n. 86.

⁶ For the nature of hieroglyphic script, see Gardiner, *Egyptian Grammar*, 6–11.

⁷ Brovanski, ‘Two Old Kingdom writing boards’. ⁸ Gardiner, *Ancient Egyptian Onomastica*.

‘scribe of sacred books in the house of life’, Amenemope, in the late Ramesside period (twelfth century BC). It contains 610 items in several thematic classes: heaven, water, earth, professions, towns of Egypt, buildings, types of land, cereals, beverages, (again) parts of an ox and kinds of meat, and so on. Gardiner used the term ‘Onomasticon’ and not ‘glossary’ or the like, since he supposed that the authors of such compositions were not interested in lexical units to describe their language but rather in a kind of encyclopedic classification of the world, as Amenemope himself states in the preface to his compendium.⁹ But in the history of lexicography the borders between dictionary and encyclopedia have not always been that clear. Wordlists organized alphabetically are not the only kind of dictionary.¹⁰

With the New Kingdom, two new challenges to the Egyptian language emerged, whose features were already recognizable before but only now began to have an impact on language classification and lexical categorization. First, the texts for monumental inscriptions in the temples and tombs, the official language (of propaganda), and the religious literature had to be written in hieroglyphs and in a classical form of the Egyptian language, ideally the Middle Egyptian of the Middle Kingdom. Middle Egyptian was by that period already an obsolete language standard, but was held in high esteem as the ‘words of the gods’. By contrast, the daily correspondence of documentary texts and *belles lettres* were written in the contemporary idiom, which we call Late Egyptian, and in a cursive script, developed from the hieroglyphs already used in the Old Kingdom, which we know as ‘hieratic’. Consequently, a situation of diglossia came into being, roughly comparable to the use of Latin in Europe during and after the Middle Ages (see Chapters 13 and 14). This phenomenon became increasingly salient, and determined textual culture in Egypt until Christianization brought a profound change with a completely new writing system based on the Greek alphabet. I will come back to this point later.

Beyond any doubt, the ‘Onomastica’ must have been standard manuals to conserve and transmit traditional knowledge about everything in the Egyptians’ world, including their language. This, however, always served to maintain the cult and the rituals in the temples as well as the equipment and decoration of the tombs of the privileged, and it was therefore necessary to be able to read and explain the sacred books that had been copied from very old manuscripts in some cases.

⁹ See Gardiner, *Ancient Egyptian Onomastica*, I.1*–2*.

¹⁰ Boisson, Kirtschuk, and Béjoint, ‘Aux origines de la lexicographie’, 268–9.

A second phenomenon began to have an increasing influence on the categorization of language in Egypt: contact with foreign languages and writing systems. Probably the first foreign writing system the Egyptians came into contact with was the cuneiform script developed in Mesopotamia. This contact is recorded as early as the period after the Middle Kingdom, when the so-called Hyksos dominated lower and middle Egypt (c. 1645–1536 BC). With Egyptian rule over the Near East in the New Kingdom, whole libraries of diplomatic cuneiform correspondence between the pharaonic court and the Near Eastern vassals and, of course, the courts of the Mitanni and Hittite empires are recorded.¹¹ So, as the language of the diplomatic exchange was Akkadian, the Egyptian scribes had to learn to read and write in this language and writing system.¹²

Egypt Under Foreign Rule: The First Millennium BC

Official and religious texts had to be written in an Egyptian idiom which was obsolete and certainly only incompletely understood by the Egyptians (of the New Kingdom and later on) themselves. This called for special instruction in Classical Egyptian. But this education was received only by a privileged circle even among the scribes, those who were to write and copy the holy books. The necessity of teaching these scribes required elaborate lists of words and even phrases in order to keep the knowledge of Classical Egyptian alive.¹³ After the New Kingdom, in the first millennium BC – especially since Demotic, an even more cursive script, served for writing documentary texts (from the seventh century BC) – all religious texts were written in hieratic (or in hieroglyphs) and used Classical Egyptian. Therefore, the formal study of language was of course relevant.¹⁴

While the many foreign rulers of Egypt mostly adopted the Egyptian text culture or even created a renaissance of Classical Egyptian in art and literature – like the Kushite twenty-fifth dynasty (746–655 BC) and their successors, one of the last indigenous dynasties, the twenty-sixth dynasty (664–525 BC) – Persian rule over Egypt (as the twenty-seventh dynasty, 525–401 BC) brought another impact of a foreign language and writing system to the Nile Valley.

¹¹ See Moran, *Amarna Letters*; Rainey, *El-Amarna Correspondence*; Edel, *Ägyptisch–hethitische Korrespondenz*.

¹² Müller, *Akkadisch in Keilschrifttexten aus Ägypten*.

¹³ Osing (ed.), *Hieratische Papyri aus Tebtunis I*, 32–4.

¹⁴ Roquet, 'Savoir et pratique linguistiques'.

The use of Aramaic as the common vernacular for the administration of the vast empire also, of course, influenced Egypt, but without enduring impact.¹⁵

Egypt Under the Ptolemaic Dynasty (332–30 BC)

Alexander the Great's conquest of Egypt, on his campaign against the Persian empire, brought the longest and deepest language contact that Egyptian ever encountered before it was finally replaced by Arabic in the thirteenth and fourteenth centuries AD. Over the 300 years after Alexander, Greek and Egyptian interacted on several levels. Greek became the dominant language of administration and of the ruling class of the country. The royal family, its court, and the higher administration spoke and wrote exclusively Greek. For any native Egyptian who wanted to start a career in the middle or higher administration, this called for knowledge of Greek. On the other hand, Greek settlers and military personnel lived in the Egyptian *chora* (countryside), and a significant number of mixed bilingual families had already emerged in the third century BC.¹⁶

There was, of course, a need for interpreters. These came, however, mostly from an Egyptian or bilingual background. The Greek citizens in the *poleis* (Naukratis and Ptolemais Hermiou) and in Alexandria lived under Greek law and saw no necessity to learn Egyptian.¹⁷ The most erudite interpreters were certainly from the clergy of the Egyptian temples, which had since the oldest imaginable times been the home of literacy. The priests wrote and translated the famous bilingual sacerdotal decrees – I have only to mention here the notorious Rosetta Stone, which led J. F. Champollion to the re-decipherment of the hieroglyphs in 1822 – and an Egyptian priest from the Delta temple of Sebennytos, Manetho, wrote a history of Egypt in Greek for the Ptolemaic library at Alexandria.¹⁸

We have to suppose that translation and language-learning must have been based on glossaries and grammars. Unfortunately, the chances of transmission have not left many examples for us. At the least, one already badly preserved Greek–Egyptian glossary dated to the third century BC, which was lost in the turmoil at the end of the Second World War in

¹⁵ See Vittmann, *Ägypten und die Fremden*, 84–119; Thompson, 'Multilingual environment', 395–9.

¹⁶ Lewis, *Greeks in Ptolemaic Egypt*; Rutherford (ed.), *Greco-Egyptian Interactions*, 1–40; Thompson, 'Multilingual environment', 399–417.

¹⁷ Kramer (ed.), *Glossaria bilinguia altera*, 1–3.

¹⁸ Quirke and Andrews, *Rosetta Stone*; Dillery, 'Literary interaction between Greece and Egypt'.

Heidelberg, could be reconstructed after a photograph.¹⁹ Both the Greek and the Egyptian words are written exclusively with Greek letters and present an everyday vocabulary, giving Egyptian equivalents to Greek words with meanings like 'bed', 'axe', 'iron', 'sword', and 'wool', very likely an *aide-mémoire* for a Greek learner of Egyptian. Though this was not the first attempt to write the Egyptian language with Greek letters, it is an important early indication of the shift in the writing system that would take place later.²⁰

At the Dusk of Pharaonic Text Culture: Egypt Under the Roman Principate (30 BC–AD 285)

It is an astonishing fact that the Hellenization of the administration in Egypt – the exclusive use of Greek as the language of all official and legal documentation – was completed under Roman rule in the first two centuries AD. While under the Greek Ptolemaic dynasty, Egyptian written in Demotic script had never totally fallen out of use for documentary texts, legal documents in Demotic gradually disappeared in the first century AD. Roman law required that official documents had to be composed in Greek if they were to have legal validity.²¹

However, just before the traditional religion and culture of Egypt expired, no later than the fourth century AD, the encyclopedic lexicography of the 'Onomastica' was still vividly in use in the (restricted) area of the temples. This is demonstrated by an extraordinary find. In 1931, an Italian mission excavating the temple district of Tebtunis in the Fayyum region discovered a huge group of mostly fragmentary papyri scattered on the floor of a house next to the temple enclosure. Other fragments obviously belonging to the Tebtunis find appeared subsequently on the antiquities market, and a considerable portion was acquired by the Carlsberg Foundation in Copenhagen. Today most fragments are kept in Florence, Copenhagen, and Berlin, with some small groups in other collections in Europe and the United States.²² The papyri are written in Egyptian (Demotic, hieratic, and hieroglyphic) and Greek. They date from the late first century BC to the second century AD. The Greek texts document the necessity, for the Egyptian priests of the temple of 'Sobek Lord of Beten' (Tebtunis), to

¹⁹ Quecke, 'Griechisch-ägyptische Wortliste'; Kramer (ed.), *Glossaria bilinguia altera*, 3–5.

²⁰ Quack, 'How the Coptic script came about', 30–7.

²¹ Cf. Yiftach-Firanko, 'Law in Graeco-Roman Egypt'; Hoffmann, *Ägypten*, 69–70.

²² Osing (ed.), *Hieratische Papyri aus Tebtunis I*, 19–23; Ryholt, 'Contents and nature of the Tebtunis temple library'.

communicate with the Roman administration, for example about taxes, in Greek.

The Demotic texts are to a large extent literary, but there are also scientific works on astronomy, mathematics, and medicine. There are also herbals, legal texts, explanations of dreams, and so on, as well as vocabularies and grammatical paradigms. Demotic was also employed for the texts of the daily service in the temple, from oracular questions to some documentary texts. The few hieroglyphic and the hieratic texts were exclusively dedicated to the *materia sacra*: ritual texts and, again, thematically organized 'Onomastica'.

As in the New Kingdom 'Onomastica', we find here lists of animals, plants, and so on, but wordlists bearing on religious topography are especially noteworthy. These compilations, sometimes organized according to the *nomes* (districts) of upper and lower Egypt, include, for instance, names of places, gods, trees, animals, and sanctuaries and their staff: all the things traditionally regarded as sacred in a certain region and, moreover, also the objects essential for performing the rituals and processions in a given temple or in its vicinity. It is easy to see that this knowledge was of high importance to the priests, because it was indispensable for organizing religious feasts and holy days of the liturgical calendar for which they were responsible. We also find here, of course, lists of the feasts and their names, of the names of the hours of night and day, and of the months. The cycle of night and day, the lunar phases (fixed in the lunar calendar), and the constellations determined the beginning of feasts and religious events.

As far as lexicography is concerned, the hieratic texts from Tebtunis published by Jürgen Osing are extraordinarily interesting. Osing retained the term 'Onomasticon' because of the resemblance of some parts of the texts to the 'Onomastica' of the New Kingdom. But he felt inclined to call the texts as a whole 'Handbücher priesterlichen Wissens' ('manuals of priestly knowledge'), which describes their major purpose more accurately. One very large but also very fragmentarily preserved papyrus which he named 'Papyrus I' begins with a list of verbs arranged not only thematically but also in semantic categories.²³ Here, for the first time, we encounter such a concept in an Egyptian text. Verbs for 'to be upset', 'to be sick', 'to achieve', 'to accomplish', 'to write', and so on, are listed together with their synonyms in a semantic class. This is followed by a list of nouns likewise organized in semantic categories, for example 'writing and scribal practice', 'terms for localities', 'chronological terms', 'diseases', 'objects of wood and leather', and

²³ Cf. Osing (ed.), *Hieratische Papyri aus Tebtunis I*, 25–218; the list of verbs is at 67–95.

'animals', this last class being subdivided as mammals, crocodiles, serpents, worms, birds, fishes, and turtles. Interestingly, in the category of birds sometimes an exact description is given of the bird and where its preferred habitat is located. Here we are again in the encyclopedic area. In the third and fourth sections of the papyrus, the lists of the *materia sacra* and the religious calendar are treated as explained above. The hieratic texts are still written in Classical Egyptian, a language that had been out of everyday use for more than a thousand years. This language had to be learned and practised by the priests and, therefore, vocabularies and grammatical exercises were necessary.²⁴ The papyri provide them.²⁵

To facilitate the understanding of the Classical Egyptian in the hieratic texts the priests used a system of glossing. A hieratic word is transcribed, translated, or both into the Demotic script and language. But, in addition, sometimes a transcription is given using Greek letters with a set of Demotic signs for the sounds of Egyptian that Greek could not render, obviously to provide the appropriate pronunciation. We know such a writing system under the name of Coptic. Since the set of the Demotic signs complementing the Greek alphabet is not yet standardized in these papyri as it would be in the fully developed Coptic writing system of the fourth century – in which the translations from Greek (the Bible and Christian literature) into Egyptian were written – it has become customary among scholars to label this kind of writing system 'Old Coptic'.²⁶ Old Coptic had hitherto been known only from magical texts dating from the first century to the third century AD.²⁷

The glossing of hieratic manuscripts written in Classical Egyptian had already been in use well before the Greco-Roman era. Whole texts or parts of them could be glossed with Late Egyptian and later Demotic equivalents.²⁸ Especially interesting is the almost complete interlinear Late Egyptian translation of a ritual text of the Late period (eighth–fourth centuries BC) written in Classical Egyptian.²⁹ The tradition of marking the divisions of the text as well as corrections and *variae lectiones* (mostly with red ink) goes back to the

²⁴ For other and older examples of this Egyptian diglossia, see Kaplony-Heckel, 'Schüler und Schulwesen in der Ägyptischen Spätzeit'; Donadoni, 'Gli Egiziani e le lingue degli altri'; Roccati, 'Il bilinguismo interno dell'Egitto'; Vernus, 'L'égypto-copte'; Quack, 'Inhomogenität von Ägyptischer Sprache'.

²⁵ For grammatical exercises, see, for example, the treatment of the verb forms of classical Egyptian in Osing (ed.), *Hieratische Papyri aus Tebtunis I*, 61–4.

²⁶ See Kasser, 'Protodialectes Coptes'; Quack, 'How the Coptic script came about', 37–49.

²⁷ Quack, 'How the Coptic script came about', 55–74.

²⁸ Osing (ed.), *Hieratische Papyri aus Tebtunis I*, 42; Quack, 'How the Coptic script came about', 37.

²⁹ Vernus, 'Entre Néo-égyptien et Démotique'.

oldest known texts.³⁰ So we see here, in the Tebtunis archive, perhaps only the late climax of a long-practised philological tradition.

In the Greco-Roman period, and possibly earlier, the Egyptians produced wordlists organized by the initial sounds of the words, probably first inspired by Old South Arabian alphabetization (cf. the acrographic south Semitic wordlists mentioned in Chapter 1). One of the best-known examples is a hieroglyphic papyrus, also from Tebtunis.³¹ It displays a list of hieroglyphic signs evidently arranged according to the consonantal values of the signs, with conceptual and religious explanations added. It is intriguing that the list starts with the letter *h*, which occupies position 12 in the alphabet of modern Egyptologists.³²

It is very likely that the rich find of the Tebtunis Papyri once belonged to the ‘house of life’, the library of the temple. This find corrects and enriches our knowledge about the lexicographical achievements of the ancient Egyptians, and more besides. The bulk of the texts, and of the hieratic texts in particular, is still to be explored.³³ We learn that such a temple library consisted of all the data thought to be indispensable for maintaining the traditional religion. We are now better able to understand what the Greek authors who visited Egypt had in mind when they spoke of the immense wisdom and knowledge of the Egyptian priests.

Latin, Greek, and Coptic: Egypt as Part of the Late Antique Roman and Byzantine Empire (AD 285–642)

Although during the 300 years of the Principate, established by Augustus in 30 BC, Roman citizens had a privileged status, and Latin was now the language of the highest civil and military administration in Egypt, Greek remained the dominant language for all other official and legal communication.³⁴ Nevertheless, a certain level of translation from Latin into Greek, and vice versa, was necessary for legal and official purposes. Therefore, the Greek-using administration needed glossaries for Greek and

³⁰ Posener, ‘Sur l’emploi de l’encre rouge’.

³¹ Iversen (ed.), *Papyrus Carlsberg No. VII: Fragments of a Hieroglyphic Dictionary*. For recent discussions, see Quack, ‘Die spätägyptische Alphabetreihenfolge und das “südsemitische” Alphabet’; Quack, ‘How the Coptic script came about’, 29–30. For Demotic lists of this kind, see Osing (ed.), *Hieratische Papyri aus Tebtunis I*, 21 n. 24; and especially Volten, ‘An “Alphabetical” Dictionary and Grammar in Demotic’.

³² See Gardiner, *Egyptian Grammar*, 27.

³³ For a recent overview, see Quack, ‘Hieratischen und Hieroglyphischen Papyri aus Tebtynis’.

³⁴ See Adams, *Bilingualism and the Latin Language*.

Latin equivalences, and quite a few examples have survived.³⁵ This belongs, however, to the history of Greek and Latin recorded in Egypt, while I want to focus here on the Egyptian language and its lexicography.

Two administrative decisions changed the shape of the Roman empire and the status of its population with lasting effect. In 212, the emperor Caracalla, in his famous *Constitutio Antoniniana*, granted Roman citizenship to all free inhabitants of the empire. Now, the legal term 'Egyptian' stood for all citizens of the province regardless of their ethnic background: Roman, Greek, Egyptian, or any other. Although the privileges of the Roman and Greek upper class did not disappear immediately, the differences were levelled gradually. The other important change was the complete reorganization of the administrative structure of the empire under Diocletian (reigned 284–305), which suspended the particular status of Egypt as an imperial province directly controlled by the emperor's prefect. Diocletian also strengthened the position of Latin as the language of law in the eastern parts of the empire. Although this was confirmed by later emperors of the east in the fifth and sixth centuries, however, Greek remained the dominant language of legal practice in Egypt.³⁶

The reign of Diocletian also brought a definite end to the pharaonic religion and textual culture with its traditional writing systems (hieroglyphs, hieratic, and Demotic). The native Egyptian population, already completely accustomed to the use of Greek for their legal affairs, now also had to convert their private written communication into Greek. But, as we have seen above, the Coptic writing system, very probably developed by erudite pagan priests, was already at hand. The Egyptian population converted to Christianity, and the brutal persecution of Christians under Diocletian and his immediate successors led in the end only to a reinforcement of Christianization under the emperor Constantine (reigned 306–37). Coptic became, beside Greek, a script and language of the Christian Egyptians, and was predominantly used in the fourth and fifth centuries to translate the Bible and Christian literature from Greek into Egyptian, very likely an achievement of the strong monastic movement in Egypt.³⁷ The language of law remained Greek (and Latin), but Coptic now came slowly into use for private written communication between Egyptian-speakers.³⁸ The surprisingly fast appearance and

³⁵ Kramer (ed.), *Glossaria bilinguia*; Kramer (ed.), *Glossaria bilinguia altera*; Fournet, 'Multilingual environment', 421–30.

³⁶ Fournet, 'Multilingual environment', 421–30.

³⁷ Feder, 'Coptic translations'; Feder, 'Koptische Übersetzung'.

³⁸ Fournet, 'Multilingual environment', 430–7.

prominence of the Coptic translations of the Bible suggest skilled support by literate people: scholars experienced in lexicography, grammar, and translation. Among the oldest witnesses for the Christian use of Coptic in Egypt is the glossing of Greek manuscripts of the Bible, a new use for a tradition, which, as we have seen, goes back in Egypt to the glossing in Late Egyptian of the Classical Egyptian of sacred texts.³⁹

How relevant Latin could still be in a certain context is shown by a Latin–Greek–Coptic ‘conversation manual’ of the sixth century, obviously used by an Egyptian-speaking group or person, giving equivalents for expressions, which might be used, for example, in writing letters to authorities.⁴⁰ Interestingly, in some cases there is no Coptic equivalent given as the Greek was apparently enough for understanding the Latin terms.

Very instructive in this context is the example of Dioskoros of Aphroditopolis, a small landowner and notable of this provincial town, and a notary and curator of the nearby monastery founded by his father.⁴¹ He lived in the sixth century and managed the business affairs of his hometown. He must have been fluently bilingual in Coptic and Greek. Besides his own attempts to write Greek poetry he compiled bilingual Greek–Coptic glossaries. One example from his archive, containing equivalents for nouns, comes very close to a wordlist or dictionary and reflects his lexical training as a person of legal affairs who had to master Greek expressions. That this wordlist was meant for his personal use is shown by the way he sometimes noted the same Coptic equivalent for different but etymologically connected or synonymous Greek words shortly with ‘this again’.⁴² How widespread such wordlists and all kinds of exercises in Greek and Coptic (rarely also in Latin) were for education in Egypt in this era is largely and effectively demonstrated in Monika Hasitzka’s compilation *Neue Texte und Dokumentation zum Koptischunterricht*.

Islamization and Arabization: The End of the Native Egyptian Language (AD 642–1400)

With the beginning of the seventh century the world of late antiquity entered a deep crisis, which finally led to its terminal decline. After the emperor

³⁹ Bell and Thompson (eds.), ‘Greek–Coptic glossary to Hosea and Amos’; Fournet, ‘Multilingual environment’, 431.

⁴⁰ Kramer (ed.), *Glossaria bilingua*, 99–108; Hasitzka, *Neue Texte und Dokumentation zum Koptischunterricht*, 210–13 (no. 270); Fournet, ‘Multilingual environment’, 428–9.

⁴¹ Fournet, ‘Multilingual environment’, 439.

⁴² Bell and Crum (eds.), ‘Greek–Coptic glossary’; see also Hasitzka, *Neue Texte und Dokumentation zum Koptischunterricht*, 181–90 (no. 256).

Justinian's attempt in the sixth century to reconquer the whole Mediterranean and to reconstruct the old Roman power finally failed, the eastern part of the empire began to develop into a realm of its own, later called Byzantine by historians, which survived until 1453. In the former western part of the Roman empire the Germanic invaders founded new states, which together with the establishment of the Papal State at Rome itself became the basis of the Latin Middle Ages. But a new power rose in the east: Islam and the Caliphate. The last act of the old world began when the eastern Roman empire and its long-time enemy, the Persian or Sassanid empire, crossed swords for a final combat. The war ended in 628/9 without a groundbreaking success for either side, but weakened the power of both fatally. Less than thirty years after the death of the prophet Muhammad in 632, the caliphs had occupied the whole Near East, Egypt, and the Sassanid empire. The entire region was, from now on, part of the Muslim world.

Already under the Sassanid occupation (619–29) but increasingly after the Arab conquest (642), Coptic became, step by step, the dominant language for documentary texts in Egypt. Of course, the period of almost a thousand years when legal administration was dominated by Greek had deeply coloured legal practice in Egypt, and this was still visible in the formal shape of the documents and, for instance, in the continued use of Greek in otherwise Coptic written contracts. The structures of the Roman administration remained in place at first, and were supplanted at the highest level only by the administration of the Wali, the Arab governor of Egypt, in which Arabic assumed, in its turn, the role which Latin had played in the centuries before. Although the caliph 'Abd al-Malik (reigned 685–705) decreed the exclusive use of Arabic in the higher administration, the first entirely Arabic protocols do not appear before 732.⁴³ A welcome example is the archive of the governor Qurra ibn Sharīk (in office 709–15), which shows that Greek was still used as an intermediary between Coptic and Arabic, with the effect that Coptic documents were translated into Greek as before, and then from Greek into Arabic, for communication with the highest level of administration.⁴⁴

Nevertheless, the timespan from the late seventh to the eleventh centuries can be regarded as a time of flowering not only of Coptic documents but also of Coptic Christian literature in general.⁴⁵ However, in the mid eleventh century, at the latest, literary production in Coptic must have come to an end and became mainly reproductive. It is difficult to fix the moment when the

⁴³ Sijpesteijn, 'Arabic papyri', 458–63. ⁴⁴ Richter, 'Language choice in the Qurra dossier'.

⁴⁵ Sidarus, 'Littérature Copte à la première époque Arabe'.

Coptic language was not actively spoken any more – as a language of liturgy in the church, Coptic is still in use today (like Latin in the Catholic church). An indication that the final stage of Coptic began in the eleventh century might be the outset of a great process of, as it were, translating the tradition, during which the Bible and the Christian literature of Egypt were translated into Arabic. Around 1300 this work was completed with a revised Arabic Bible and Arabic versions of large parts of the dogmatic, patristic, and canonical heritage, commentaries on biblical books, the *Synaxarion* (festal calendar), and the history of the Christian patriarchs of Egypt. From the twelfth century onwards, liturgical manuscripts were, again, bilingual, with an Arabic version in the margin besides the more prominently presented Coptic text.⁴⁶

In view of the extinction of Coptic in the thirteenth and fourteenth centuries, an astonishing school of Coptic philology and lexicography emerged and developed two kinds of tools: Coptic–Arabic wordlists (Arabic *sullam*, Latin *scalae*, ‘ladders’) and Coptic grammars written in Arabic (Arabic *muqaddimmāt*).⁴⁷ John Samannudi was one of the early prominent scholars in lexicography. He wrote a *Scala ecclesiastica* headed by a grammatical introduction, which was in effect a Coptic–Arabic glossary of the biblical and liturgical books of the Coptic church, because it follows the order of those books and not an alphabetical or semantic sequence.⁴⁸ The best-known of these lexica was the *Scala magna* (Arabic *al-Sullam al-kabīr*), the ‘Great scala’, compiled by the then widely known Coptic encyclopedist Abū l-Barakāt ibn Kabar.⁴⁹ His work was also called *al-Sullam al-muqtarah* (‘invented vocabulary’). This classified vocabulary is divided into ten large chapters and comprises a total of thirty sections. The division of his glossary recalls immediately the pharaonic ‘Onomastica’, since it classifies the vocabulary according to semantic groups. Beginning with the names and attributes of God, the celestial hierarchies, the higher and then the lower cosmos, it passes to the animal, vegetable, and mineral worlds. After that the author inserts the terminology related to human beings: the body, the senses, the diseases and infirmities, the virtues and vices, the languages and the nations, the social and military hierarchy, the professions and the arts and crafts, and so on. After chapter 6 on geography and topography comes chapter 7 on the church, its cult, and its institutions. Amazingly, the old Egyptian tradition of the

⁴⁶ Richter, ‘Greek, Coptic, and the “language of the Hijra”’, 418.

⁴⁷ Sidarus, ‘Medieval Coptic grammars in Arabic’.

⁴⁸ Richter, ‘Greek, Coptic, and the “language of the Hijra”’, 418–19; Sidarus, ‘Coptic lexicography’, 127–8.

⁴⁹ Sidarus, ‘Coptic lexicography’, 132–3.

encyclopedic genre was still alive in a culture of Arabic-speaking Christian scholars more than a thousand years after the pharaonic culture had disappeared.

Probably one of the latest but most important scholars in this movement was Athanasius of Qus, who worked on wordlists and wrote a Coptic grammar in Arabic.⁵⁰ The *Scala* of John Samannudi and also the Coptic grammar of Athanasius of Qus in Arabic, and many other texts of this kind, were transmitted to European scholars as early as the sixteenth and seventeenth centuries. They learned the Coptic language by way of its description in Arabic, like the educated Copts who were still trying to keep their ancient language alive in the fourteenth century.⁵¹ When the German Jesuit Athanasius Kircher published his *Lingua aegyptiaca restituta* (1643), having learned Coptic from these wordlists and grammars, he was aware, almost two centuries before the decipherment of the hieroglyphs, that Coptic must be nothing other than the language of the ancient Egyptians.

⁵⁰ Sidarus, 'Medieval Coptic grammars in Arabic', 70–5.

⁵¹ Richter, 'Greek, Coptic, and the "language of the Hijra"', 428–9; Sidarus, 'Coptic lexicography', 137–40.

Ancient China

FRANÇOISE BOTTÉRO

As in many other cultures, the beginning of Chinese lexicography is rooted in the philological heritage. Interest in classical texts whose meaning had become difficult to understand engendered a significant number of explanations. Following or inserted into the texts, these explanations were subsequently collected to form the first glossaries and collections of synonyms. Later scholars would draw deep from within this rich exegetic tradition to create the first dictionaries.

Before Dictionaries: The First Wordlists

The practice of writing requires methods and then perhaps even manuals to learn how to write. In the case of ancient China, there was also the need to manipulate different styles of writing, depending on the type of documents to be produced. Nothing is known about how the earliest Chinese scribes were educated or what kind of documents they may have used to learn and practise.

Shǐzhòu Piān 史籀篇

The first primer for children of which there are records is the *Shǐzhòu piān* 史籀篇, traditionally attributed to the scribe of King Xuān 宣 of Zhōu, who reigned in the late ninth and early eighth centuries BC. The original text, which comprised fifteen chapters written in large seal script (*dàzhuàn* 大篆 or *zhòuwén* 籀文), has not survived.¹ By the Hàn dynasty (206 BC–AD 220), nine chapters were already lost, but it appears that some 225 graphs were chosen from the chapters which were still extant and included in *Shuōwén jiězì* (for which, see below). Based on the analysis of these characters and their graphic constituents, it is thought that the *Shǐzhòu piān* was probably

¹ Bān Gù, *Hànshū*, 30.1719.

composed under royal command around the fifth century BC, in an effort to introduce an orthographic norm.²

Cāngjié Piān 倉頡篇

A series of other wordlists was produced when the first emperor, Qín Shǐ huángdì 秦始皇帝, conquered the Chinese territory in 221 BC. As part of the unification of the Chinese empire, he tried to impose unification of writing on all the ancient states that had developed their own graphic variants. His prime minister Lǐ Sī 李斯 compiled a manual of characters called *Cāngjié piān* 倉頡篇. The original text, in seven sections, consisted of rhyming sentences of four characters, and its study was strongly recommended for those who wanted to get a position in government. Two other texts, one in six sections and one in seven, soon followed, forming a larger work, which was still known as *Cāngjié piān*. With the Hàn dynasty the *Cāngjié piān* was rearranged and augmented for pedagogical needs. Different versions, some including explanations, were also produced. Most of them were lost, but fragments have been found in different places in China, suggesting the importance and the wide diffusion of the text at that time.

Jíjiù Piān 急就篇

Among the ten or so (mnemonic) wordlists produced during the Western Hàn and recorded in the bibliography of the great historical work *Hànshū*, only the *Jíjiù piān* 急就篇 is still available today.³ This manual, conceived to help its readers learn a text of 2,016 characters rapidly, was written by Shǐ Yóu 史游, under the reign of Kǐng Yuán 元, in the second half of the first century BC. It consisted of thirty-two sections of sixty-three characters each, with different lengths of rhyming phrases. The words, very often disyllabic, were enumerated without being embedded in sentences. The vocabulary was organized thematically: family names, fabrics, colours, buying and selling, cereals, vegetables, metallic objects, manufactured objects, aquatic animals, women, servants, sleeping room objects, musical instruments, kitchen, food, the human body, weapons, charts, buildings, terms related to work in the fields, trees, animals, diseases, pharmacopeia, and terms from the religious and ritual domain. In the end, Shǐ Yóu gave a general presentation of government organization with officers' names and titles, a list of the texts officers were supposed to study, and the laws and regulations they should know. The *Jíjiù piān* ended with a paragraph to the glory of the Hàn.

² See Pān Yùkūn, “Shǐzhòu piān” niándài kǎo’. ³ Bān Gù, *Hànshū*, 30.1719–20.

From the study of the *Jíjiù piān* as well as that of the various fragments of the *Cāngjié piān*, we can understand how these manuals were composed, and how they could have been influential for the lexicographic tradition that would arise a century or so later. Not only did they collect words, with close or opposed meanings, but they also tended to put together characters written with the same semantic constituent (a graphic element which enters into the formation of multiple characters, and is supposed to give them its meaning). In the bamboo slips C33–C34 of the *Cāngjié piān*, discovered in Fù yáng 阜陽 in 1977, or in chapter 11 of the *Jíjiù piān*, for example, we can see no fewer than nine characters with the semantic constituent *hēi* 黑 ‘black’, and nineteen characters with the semantic constituent *jīn* 金 ‘metal’, in a row. This no doubt gave Xǔ Shèn (see below) the idea of radical classification, which would later play a role in Chinese dictionaries equivalent to that of the alphabet.

Before Dictionaries: The First Collections of Glosses and Synonyms

Next to wordlists, collections of glosses or synonyms were also produced before and after the development of dictionaries in the second century. Three of them are still extant: the *Ēryǎ*, the *Fāngyán*, and the *Shímíng*. The *Ēryǎ* was by far the most important lexicographic work of its time.

The Ēryǎ 爾雅 ‘Approaching Perfection’

The *Ēryǎ* is traditionally presented as the oldest Chinese dictionary. The author is unknown but, from the Eastern Hàn dynasty onwards, different dates have been proposed for its composition. It has been said, for instance, to have been written by disciples of Confucius (who died in the fifth century BC), or by the younger brother of the first duke of Zhōu (twelfth century BC). The content and the heterogeneous structure of the text suggest a much later date. The *Ēryǎ* gathers many expressions from the classics and pre-Qín (late third century BC) authors, but it also includes a certain number of terms and toponyms clearly linked to the Hàn. Thus, it should best be considered as a compilation regrouping different glosses, scholia, or texts written between the fifth and the first centuries BC, and compiled around the time of Emperor Hàn Wǔdì 武帝, whose reign ended in 87 BC.⁴

⁴ Zhōu Zǔmó, *Wén xué jí*, 675.

The text of the *Ēryǎ* which has been transmitted to the present day is 13,000 characters long and comprises 19 chapters. A preface is said to have existed, and a major early source counts twenty chapters, one of which was presumably the lost preface. One should distinguish two (or even three) parts in the *Ēryǎ*. The first three chapters present lists of words (nouns, verbs, adjectives, grammatical words, adverbs, and so on) with more or less the same meaning, whereas the last sixteen chapters resemble small encyclopedic treatises organized according to different themes and providing information concerning names and expressions related to each theme. In other words, the *Ēryǎ* combines three collections of synonyms (chapters 1 to 3) with sixteen topical glossaries (chapters 4 to 19).

The first chapter, *Shì gǔ* 釋詁 ('Explaining old words'), gathers 173 lists of words that can be found in ancient texts. Each of these lists is defined by a single and more common word that comes at the end of the list (that is, the bottom: Chinese texts were written from top to bottom): for example, 'Qí 𣎵 [𣎵 is a rare character] is like qí 汽 "vapour".' The lists can include as many as thirty-nine 'synonyms' (with words used in different contexts to refer to the same notion), the longest being the third list of words corresponding to the meaning 'big'. The text starts with words referring to the meaning 'beginning' and ends with those referring to the meaning 'death':

First, sprout [of a tree], head, basis, initiate, ancestor, primary, foetus, start, set, bud are [like] 'beginning' [chū 初, zāi 哉, shǒu 首, jī 基, zhào 肇, zǔ 祖, yuán 元, tāi 胎, chù 俶, luò 落, quán yú 權輿, shǐ yě 始也] . . .

Collapse, death of a prince, death, die, pass away, die are [like] 'die/death' [bēng 崩, hōng 薨, wú lù 無祿, zú 卒, cú luò 徂落, yì 殯, sǐ yě 死也].⁵

Next to formal, poetic, or ancient words, we find common as well as dialectal words. An example is a list of first-person pronouns, 'áng 印, wú 吾, yí 台, yú 予, zhèn 朕, shēn 身, fǔ 甫, yú 余, yán 言, wǒ yě 我也'.⁶ Here, áng 印 is a dialectal word corresponding to wú 吾 'we', as yí 台 corresponds to yú 予.⁷ Zhèn 朕 was an ordinary southern dialect first-person pronoun in the *Songs of Chǔ* (*Chǔ cí* 楚辭, fourth–third century BC), before it became an honorific first-person pronoun in the third century BC. Shēn 身 'body' is a pseudo-pronoun used by a speaker to refer to himself. Fǔ 甫 is a courtesy name that is explained as referring to the first person in the *Books of Rites* (*Lǐ jì* 禮記, fourth–second century BC). The pronoun yú 余, like yú 予 'we

⁵ *Ēryǎ*, 1.1, 1.173. For matter of convenience I refer to the numbers Xú Cháohuá provides in his *Ēryǎ jīnshù* for each list.

⁶ *Ēryǎ*, 1.43. ⁷ Xú Cháohuá, *Ēryǎ jīnshù*, 27.

[personally]’, is often used by poets. *Yán* 言 ‘to speak’ can be used to refer to the speaker(s) of the speech. The more standard *wǒ* 我 ‘we [the speaker’s group]’ is used as the *definiendum*.

Words have different meanings, and therefore they may appear more than once. In the next list, another meaning of *zhèn* 朕 and *yú* 余, which we have just seen as pronouns, is made clear as they are brought together with *gōng* 躬 ‘body, person’ in a list of words equivalent to *shēn* 身 ‘person, ego, self’.⁸ Moreover, words in the same list do not necessarily have the same meaning. In the next list, the six words *yí* 台, *zhèn* 朕, *lài* 賚, *bì* 畀, *bǔ* 卜, and *yáng* 陽 are all glossed with the character 予.⁹ But this character records in fact two different words: *yǔ* ‘give’ and *yú* ‘we’. Here, *yí* 台 and *zhèn* 朕, refer, as they did in the list discussed above, to the first-person pronoun. *Bǔ* 卜 is said to be used for ‘we’ in a verse of the *Book of Songs* (tenth–sixth century BC).¹⁰ As for *yáng* 陽, it stands for the homophonous dialectal pronoun *áng* 印 ‘we’, which was discussed above. In between these four words equated with *yú* ‘we’, however, *lài* 賚 and *bì* 畀 both mean ‘to give’, and are being equated with *yǔ* ‘give’. So, in this particular case, the *definiendum* has two meanings, and the list combines words with one meaning or the other. This potentially confusing list shows that the *Ēryǎ* was aimed at scholars who already had a good knowledge of ancient Chinese.

The second chapter, *Shì yán* 釋言 (‘Explaining words’), gathers 280 smaller lists of synonyms of more ordinary use, such as ‘*yǒng* 泳 is like *yóu* 游 “to swim”’ and ‘*móu* 謀 “to plan” is like *xīn* 心 “to think”’.¹¹ It is interesting to note that the second chapter ends with two words meaning ‘end’: ‘*mí* 彌 is like *zhōng* 終 “end”’. One is tempted to think that this second section could have been compiled as a development of the first one to form an independent set, and thus count as the third part of the *Ēryǎ*.

With its own independent structure, the third chapter, *Shì xùn* 釋訓 (‘Explaining meanings’), gathers 116 lists. The first seventy-five lists present disyllabic expressions, taken from the classics (for instance, the *Book of Songs* and *Book of Documents*, eleventh–third century BC) or other pre-Qín and Hàn texts, which are defined by a monosyllabic word or a sentence. For example: ‘*Màomào* 懋懋, *mómò* 模模, is *miǎn* 勉 “to make earnest effort”’.¹² The expression *màomào* 懋懋 appears in the *Book of Documents*.¹³ It had been

⁸ *Ēryǎ*, 1.44. ⁹ *Ēryǎ*, 1.45.

¹⁰ The verse is 卜爾萬壽無疆 ‘We predict for you a myriad years of life without limit’ (‘*Tiān bǎo*’ 天保, in *Shísānjīng zhùshù*, 412b); in fact, the subject ‘we’ has been dropped in this verse, and *bǔ* 卜 is simply a verb meaning ‘to divine, predict’.

¹¹ *Ēryǎ*, 2.97, 2.137. ¹² *Ēryǎ*, 3.28; Xú Cháohuá, *Ēryǎ jīnshù*, 136.

¹³ *Shísānjīng zhùshù*, 161b.

explained by the earlier Hàn commentator Kǒng Ānguó 孔安國 as *miǎn* 勉 ‘to make earnest effort’ (*miǎnli* 勉力 ‘to make effort’, *quǎnmiǎn* 勸勉 ‘to encourage’), just as it is in the *Ēryǎ*. In another passage from the *Book of Documents*, Kǒng Ānguó also explains *mào* 懋 as *miǎn* 勉 ‘to make earnest effort’.¹⁴ This example shows how glosses given by commentators on the *Book of Documents* or other ancient texts were culled to form the *Ēryǎ*. The end of the third chapter also gathers forty-seven expressions excerpted from ancient texts and explained: for example, ‘*shéi xī* 誰昔 is equivalent to *xī* 昔 “before, in the past”’.¹⁵

The other sixteen chapters of the *Ēryǎ* follow a different pattern since they are organized thematically, with the subjects of kinship, architecture, utensils, music, the heavens, the earth, hills, mountains, rivers, bushes and grasses, trees, insects, aquatic animals, birds, beasts, and domestic animals. Chapters 13 to 19 all gather essentially plant and animal names, formal as well as colloquial; as will be discussed in Chapter 6, some later works in the *Ēryǎ* tradition focused exclusively on this semantic domain.

In the sixteen topical chapters of the *Ēryǎ*, most of the formulae used to gloss words and expressions take the shape ‘X corresponds to Y’ or ‘X is called Y’ or ‘X is like Y.’ Thus, for example, ‘A *zōnggǔ* [“fishing net”] is called *jiǔyù*, a *jiǔyù* is a net for fishing’ and ‘*Zǎi* is like *sui* “year”. The Xià called it *sui* [“harvest”], the Shāng called it *sì* [“sacrifice”], the Zhōu called it *nián* [“harvest”], [the legendary emperors] Yáo 堯 and Shùn 舜 called it *zǎi* [“year”].’¹⁶ Subthematic ordering is used to classify words and expressions in these chapters.

A close look at chapter 8 of the *Ēryǎ* shows how the topic of the heavens (more precisely, ‘Explaining celestial [terms]’) is organized under twelve subdivisions. This chapter first provides, under the heading ‘Four seasons’, the different names for the sky according to the four seasons; then secondly ‘The auspicious signs’ according to the four seasons; and thirdly ‘The disasters’, with the different terms for hunger engendered by different kinds of bad harvest. Fourthly to seventhly, under ‘Names of Jupiter’, it lists the names of the ten and the twelve years of ancient Chinese chronology according to the position of Jupiter in the sky; then the five ‘Names of the year’ according to the different dynasties (this passage was quoted above); the different names for the moon according to its ten positions in the sky; and the names for the months. Eighthly, the text presents under

¹⁴ *Shísānjīng zhùshū*, 130b. ¹⁵ *Ēryǎ*, 3.113.

¹⁶ *Ēryǎ*, 6.4, ‘纓罟謂之九罟；九罟，魚罔也’ (Xú Cháohuá, *Ēryǎ jīnshù*, 178); *Ēryǎ*, 8.9, ‘載，歲也。夏曰歲，商曰祀，周曰年，唐虞曰載’ (Xú Cháohuá, *Ēryǎ jīnshù*, 201).

‘Wind and rain’ the different names of the wind according to the quarter from which it blows or to some of its particularities, and the names for frost, rainbow, hailstone, and thunderbolt, as well as different types of rain. Ninthly, under ‘Constellation names’, it provides names of different stars and constellations, grouped according to cardinal directions. The names for the different sacrifices are listed according to the four seasons and their beneficiaries (such as heaven, earth, mountains, rivers, constellations, and winds) in tenth place. Then we find the different names for hunting according to the four seasons, followed by two excerpts taken from the classical *Shījīng* (詩經 ‘Book of poems’), explained and commented on at length under the eleventh heading, ‘Military training’. The first excerpt discusses the importance of offering sacrifices before letting armies march, and the second the order of seniority when going to or coming back from war. The twelfth and last part describes the different kinds of banners and flags.

In the *Ēryǎ*, the lexicographic unit is the word, and words are often disyllabic. Some of them have different meanings, others refer to different words, and characters very often stand for homophonous words. Thus, polysemy is partially taken into account as well as homonymy. The *Ēryǎ* was no doubt compiled to help the reading and understanding of the classics and the ancient texts. But the non-homogeneous structure of the text and the lack of systematic ordering makes it difficult to count the *Ēryǎ* as the first dictionary. Yet, as the first attempt to collect semantic glosses and discuss words out of their context (excerpts set apart), the *Ēryǎ* played a very important role for the beginning of Chinese dictionaries.

The Fāngyán 方言 ‘Regional Words’

Another glossary of a different kind was produced at the beginning of the first century AD. It is called the *Fāngyán* 方言 (‘Regional words’), which is an abbreviation of the full title, *Yóuxuān shǐzhě juédài yǔshì biéguó fāngyán* 輶軒使者絕代語釋別國方言 (‘The imperial light carriage emissary explaining discursively the regional words in different states through the ages’).¹⁷ Compiled by Yáng Xióng 楊雄, the *Fāngyán* collected synonyms taken from different dialects and languages, gathered by court messengers who had been sent to various regions of China. It is the oldest known Chinese documentation on languages other than Chinese. The text, divided into 13 thematic chapters, contained more than 9,000 characters. Like the beginning of the *Ēryǎ*, the *Fāngyán* first presents lists of synonyms, before naming the area of use for

¹⁷ *Fāngyán jiào jiān* 方言校箋 (1956) is an edition.

each word. It also mentions ancient versus contemporary words, and words of common usage all over the Hàn territory or among a speech community, as well as close or slightly modified pronunciations between dialects.¹⁸ However, Yáng Xióng was in fact more concerned with the different ways to write words rather than with words per se.

The Shimíng 釋名 'Explaining Names'

The *Shimíng* 釋名 ('Explaining names') was apparently composed by Liú Xī 劉熙, in AD 200.¹⁹ The particularity of this topical glossary lies in the general use of paronomastic glosses (also called puns or phonetic glosses, *shēng xùn* 聲訓) in order to clarify the supposed etymology of some 1,500 words. This method consisted in giving a more or less homophonous word with the entry to show a semantic link between them. These supposed motivations at the basis of the creation of words usually corresponded to folk etymology, as in the following examples.

Yuè 'moon' is associated with *quē* 'lacking': it wanes after being full.

When a man begins to discontinue breathing, we talk about *sǐ* 'death'. *Sǐ* 'death' is associated with *sī* 澌 'disappear': it is to move towards disintegration, *xiāosī*.

Zēngzǔ 'great-grandfather': the ones below [i.e. the ones who died later] push the ones above [i.e. the ones who died earlier], so the position of the ancestors is removed and goes further up ['augments' [*zēngyì*] in the ancestral hierarchy].²⁰

Liú Xī arranged the words into twenty-seven thematic sections: heaven, earth, mountains, rivers, hills, roads, regions and kingdoms, shapes and bodies, appearances, seniority, kinship, expressions, food, silks, ornaments, clothes, palaces, beds and curtains, written documents, classics and arts, tools, and musical instruments.

Compared to the preceding lexicographic works, if Liú Xī also gathered many glosses from ancient texts and commentators, his interest in the everyday use of words was quite original. He introduced colloquial words that one would hardly find in any other extant glossary, with senses such as 'caress', 'laugh', and 'latrines'. He also discussed the polysemy and the homonymy of some words. *Wàng* 望, for example, is not only explained as 'to look into the distance' in chapter 9, its deviant use referring to a kind of beam is also

¹⁸ Harbsmeier, *Language and Logic*, 77.

¹⁹ *Shimíng huì jiào* 釋名匯校 (2006) is the standard edition.

²⁰ *Shimíng*, 1.3, '月, 缺也; 滿則缺也'; 27.1, '人始氣絕曰死, 死, 澌也, 就消澌也'; 11.6, '曾祖, 從下推上, 祖位轉增益也'.

explained in chapter 17, and the meaning ‘full moon’ is discussed in the first chapter.²¹ When using the formula ‘A is associated with A’ (A A 也), Liú Xī could record examples of homonymy, for instance ‘Bù “cloth” is associated with bù “spread”’(布布也).²² He showed, moreover, that different words could refer to the same thing not only between dialects, or different periods, but also in what may be called the standard language. Last but not least, he expressed a special interest in the pronunciation of spoken words and went as far as referring to the way sounds should be articulated.

But the *Shímíng* was too limited by the small number of entries Liú Xī could explain with his ad hoc phonetic glosses. It had little impact on later generations of dictionaries, which essentially focused on the meaning of characters in the classics.

The Invention of Dictionaries in China: *Shuōwén Jiězì*, and *Yùpiān*

The period of the *Ēryǎ*, the *Fāngyán*, and the *Shímíng*, from around 100 BC to around AD 200, overlaps with that of the first Chinese texts which can uncontroversially be called dictionaries, namely *Shuōwén jiězì* (around AD 100) and *Yùpiān* (AD 543).

Shuōwén Jiězì 說文解字 ‘Explain the Graphs to Unravel the Written Words’

Shuōwén jiězì 說文解字 (abbreviated as *Shuōwén*), which preceded the *Shímíng* by a hundred years, is the first dictionary of Chinese characters.²³ It was composed by Xǔ Shèn 許慎 around AD 100. For the first time, all the characters included in the work were presented according to a new system of classification invented by the author. This is a tremendous achievement in the history of this non-alphabetical writing system. Xǔ Shèn gathered as many as 9,353 characters and organized them according to 540 ‘classifiers’ or semantic constituents (also called ‘radicals’ *bùshǒu* 部首) such as ‘one’, ‘woman’, ‘jade’, and ‘aquatic animal’. It counted 14 chapters followed by a postface, and also included some 1,163 allographs. Each of the 540 radicals gathered between 1 (the radical only) and more than 400 headgraphs or characters. It appears that when

²¹ *Shímíng*, 9.29 and 12.112 (‘look into the distance’), 17.25 (‘kind of beam’), 1.73 (‘full moon’).

²² *Shímíng*, 14.14.

²³ Xǔ Shèn, *Shuōwén jiězì* (2006), is an edition – not of the lost original, but of the tenth-century recension by Xú Xuàn (see Chapter 6) which is used in its place.

many entries were collected under a single radical, a certain number of semantic series structured the text.²⁴ The headgraphs were presented in the small seal-style *xiǎo zhuàn* 小篆, an older style of writing on which Xǔ Shèn based his graphic analysis. For each entry, Xǔ Shèn provided a gloss as well as a graphic analysis into constituents, semantic, phonetic, or both. But he sometimes also gave supplementary information concerning meanings, pronunciations, allographs, or illustrative quotations, as well as encyclopedic material.

The originality of *Shuōwén* is that the lexicographic unit is the character, and no longer the word. Xǔ Shèn wanted to show the necessity of retrieving the proper way to write characters from older-style graphic forms. He was not interested in the basic meaning of words. When he glossed *wǔ* 五 ‘five’ as ‘[for example] the Five Elements’, he did not define it as a number, but instead provided an illustrative quotation, which from our modern perspective corresponds to the mistake of using the *definiendum* in the *definiens*. And then when he said in his graphic analysis that ‘[The graph 五] has “two” as a semantic constituent, with the Yin and the Yang intersecting between Heaven and Earth’, we can see that he wanted to make a connection between the Five Elements and the Yin and Yang theory.²⁵ In his dictionary, Xǔ Shèn tried his best to define the meaning of a graph according to its semantic constituent. In other words, characters with ‘heart’ as a semantic constituent were usually defined in terms of psychology, and those with, for example, the semantic constituents ‘woman’, ‘jade’, or ‘aquatic animal’, in terms of those constituents. The use of the older small seal-style graphs allowed Xǔ Shèn to recover graphic constituents before they were modified in modern script and to provide what he saw as a correct analysis of graphs, which could clarify their supposed original meaning in the classics. Thus, *Shuōwén* is not a dictionary of the meaning of words. It is a graphic etymological dictionary in which Xǔ Shèn tried to provide the meaning that best suited the graphic structure of the graph (written word) and its immediate constituents. Xǔ Shèn came close to composing a dictionary of the orthography of words.

This being so, we can understand why polysemy is scarcely touched upon in *Shuōwén*: the supposed original meaning of a graph was unitary. So, Xǔ Shèn usually only gave one gloss for each headgraph. He drew many of his glosses from the rich exegetic tradition, sometimes mentioning his sources (including the *Ėryǎ* and commentators on classical texts), but often keeping silent about them. Thus, we find all sorts of glosses in *Shuōwén*: near-

²⁴ Bottéro and Harbsmeier, *Chinese Lexicography on Matters of the Heart*.

²⁵ Xǔ Shèn, *Shuōwén jiězì*, 14B 7b, ‘五 五行也。从二陰陽在天地間交午也’.

synonymic glosses ('Zhōng "loyal" is [a way of] showing respect'); analytical glosses ('Lì "minor official" is someone who keeps order among others'); dialectal glosses ('In Yǎnzhou, they refer to cheating by saying *tuó*'); paronomastic glosses, usually followed by an explanation ('Jiǔ "fermented wine" is [evocative of] *jiù* "approach". It is the means by which one approaches the good or evil nature of man'); technical glosses ('A *zǔ* is an auspicious relief on a *cóng*-type jade'); geographic glosses ('Yǐng is the ancient capital of the Kingdom of Chǔ; it is located 10 li north of the south province of Jiānglíng'); and others.²⁶ Xǔ Shèn not only dealt with lexical items, but also included proper names in his work. For example, under the radical 'water', the 464 entries are divided into 2 sections. The first section deals essentially with the names of rivers, and the second with lexical items. The rivers are neatly arranged according to the four cardinal directions and the centre, with all the provinces of his time included. Thus, this section dedicated to rivers offers a kind of encyclopedic treatise on Chinese rivers. In other words, *Shuōwén*, which incorporates features of an organized thematic encyclopedia, is much more than a dictionary of orthography and graphic etymology.

Following the *Ēryǎ*, but in a more systematic way, Xǔ Shèn explained words out of their context (although in an orally based culture, the context was probably more present in the reader's mind and recognizable than it may seem to us). Xǔ Shèn also provided a method to organize the entries equivalent to the alphabet in European dictionaries. Yet looking for a character in *Shuōwén* was not an easy task, for the radical classification was so much intertwined with the Hàn philosophical context (for instance, the Five Elements, and Yin and Yang) and Xǔ Shèn's own worldview. For example, for the first character and radical, *yī* 一 'one', Xǔ Shèn referred to traditional Chinese cosmogony, which placed One at the very beginning of the creation of the Universe, and thus did not gloss it as a number but as the metaphysical beginning of all things (characters included): 'Initially, at the great beginning, the Way established through the One. By separation it created Heaven and Earth, and [thus] transformed so as to bring to completion the Myriad Creatures.'²⁷

In the postface, Xǔ Shèn provided the first account of the history of the Chinese script, explaining what preceded it and how it came to be invented by the scribe of the legendary Yellow Emperor. He also touched upon the 'Six ways of writing down (words)' *liùshū* 六書: a pedagogical method,

²⁶ Xǔ Shèn, *Shuōwén jiězi*, 忠敬也; 吏治人者也; 沆州謂欺曰沆; 酒就也, 所以就人性之善惡; 珎琮玉之璫; 郢故(古)楚都。在南郡江陵北十里。

²⁷ Xǔ Shèn, *Shuōwén jiězi*, 1A 1a, '惟初太始, 道立於一, 造分天地, 化成萬物'.

traditionally used to introduce writing to eight-year-old children. Four among these ‘six ways’ – ‘symbolize physical shape’ *xiàng xíng* 象形, ‘refer [pictorially] to something’ *zhǐ shì* 指事, ‘indicate shape and sound’ *xíng shēng* 形聲, and ‘associate ideas’ *huì yì* 會意 – referred to character creation, whereas the last two – ‘borrow [one graph for another]’ *jiǎ jiè* 假借 and *zhuǎn zhǔ* 轉注 – referred to the re-utilization of the others. But it is not clear what exactly *zhuǎn zhǔ* 轉注 or ‘refer (pictorially) to something’ *zhǐ shì* 指事 mean.²⁸ If Xǔ Shèn occasionally mentioned that a character ‘symbolizes physical shape’ *xiàng xíng* 象形 in his work, he never characterized characters in terms of *liùshū*. The *liùshū* were in no way imposed in the body of his dictionary, as some scholars would imagine a few centuries later (see Chapter 6). The importance of the *liùshū* lies in the fact that phonetic constituents in the characters were recognized for the first time. And this, no doubt, made Xǔ Shèn’s graphic analysis into semantic and phonetic constituents possible.

Shuōwén was not a book to consult; it was a text one would read from the beginning to the end or memorize, until indexes were added much later to facilitate its consultation. It had a tremendous influence on the dictionaries of later generations. We know that, among the lexicographic works compiled between the second century and the fourth that have not survived, some copied its organization. The *Zilín* 字林 (‘Forest of characters’) written by Lǚ Chén 呂忱, for example, gathered 12,824 entries and used the 540 radicals to classify the characters. It was composed to complete *Shuōwén*. The author sometimes added the pronunciation of the entries using methods such as *fǎnqiè* 反切 (for which, see Chapter 6), or *zhíyīn* 直音 (which consisted in providing a homophonous character). Yet none of these character dictionaries reached the level of the *Yùpiān* 玉篇, to which we now turn, in terms of lexicographic improvement.

The *Yùpiān* 玉篇 ‘The Jade Chapters’

One of the first real Chinese meanings dictionaries is the *Yùpiān* 玉篇, composed in 543 by the very learned and talented Gù Yěwáng 顧野王. The original text, which counted 16,917 headgraphs for 30 chapters, plus a preface, was already lost by the Sòng dynasty (960–1279), and was replaced by the augmented version of Chén Péngnián 陳彭年, *Dàguǎng yìhuì Yùpiān* 大廣益會玉篇 (1013), which included 22,561 entries.²⁹ Many versions of the *Yùpiān*

²⁸ See Qiú Xīguī, *Wénzìxué gǎiyào*, 100–1.

²⁹ Chén Péngnián, *Sòng běn Yùpiān* (1983), is an edition of this recension; see also Gù Yěwáng, *Yuánběn Yùpiān cǎnjiàn* (1985).

were produced as early as the seventh century, with more or fewer modifications. Among them Sūn Qiáng's 孫強 *Yùpiān* (674) and the *Dàguǎng yìhuì Yùpiān* are the most famous. But there were also a Buddhist and a Taoist as well as a Korean and a Japanese version of this work.

Luckily, more than a tenth of the original *Yùpiān* was found at the end of the nineteenth century in Japan, where it had been imported around the end of the ninth century. The study of these fragments (which include sixty-three radicals from seven chapters) shows little change in the organization of the dictionary between the original and the extant version, but, as we shall see, drastic cuts in the presentation.³⁰ Gù Yěwáng took over Xǔ Shèn's system of classification, but modified it more than it seems. Among the 540 radicals used by Xǔ Shèn, he suppressed 10, and added 12 new ones, which made a total of 542 radicals. As we can tell from the eleventh-century augmented version, the first and the last entries (or radicals) were faithfully reproduced, but the organization of the radicals in the chapters was completely modified to provide a thematic organization. It started with radicals related to the subject 'heaven', followed by 'earth', 'man', 'kinship terms', 'parts of the body', 'buildings', 'plants', and 'musical instruments', and proceeding through other categories including 'animals', to end with 'numbers'. Gù Yěwáng was not interested in the structure of the characters, as Xǔ Shèn had been, but in their meaning. Therefore, he only used the current *kǎishū* 楷書 style to write all the entries (including ancient allographs in the *zhòuwén* and *gǔwén* scripts taken from Xǔ Shèn). One can incidentally notice the importance of *Shuōwén* in the eyes of Gù Yěwáng, who often referred to it. The author of the *Yùpiān* also referred to other lexica or glossaries, many of which have disappeared – for instance, the *Ēryǎ*, *Fāngyán*, *Cāngjié piān*, *Pí Cāng* 埤蒼, *Guǎngyǎ* 廣雅, *Zìshū* 字書, and *Shēnglèi* 聲類 – for the way they defined words or wrote them differently.

The comparison between the Sòng version and the original version of the *Yùpiān* is very instructive if we are to understand the level of development of Gù Yěwáng's original work. In the case of the entry *diǎn* 典 'canon, law, decree, document, classic', for example, the Sòng version gives only the pronunciation with the *fǎnqiè* 反切 spelling method, and a very simple semantic definition. By contrast, the original *Yùpiān* offered a much more complete and elaborated definition of this term in the classics, focusing on polysemy. Gù Yěwáng first gave the pronunciation with the *fǎnqiè* method. Then, using various examples, he presented the different meanings of *diǎn* 典

³⁰ Bottéro, *Sémantisme et classification*, 96–105.

in the classics followed by commentators' glosses. The first example is from *The Book of Documents*: 'He had *diǎn* 典, he had patterns.' It is followed by a gloss by Kǒng Ānguó which says '(*Diǎn* 典) refers to the canons.' The second and third examples are from the *Rites of Zhōu* (third century BC), 'He is in charge of the six *diǎn* 典' and 'Two middle-rank servants are attributed to the rulers' wives 典婦', with another Hàn commentator's glosses explaining *diǎn* 典 in these passages as equivalent to *cháng* 常 'law' and *zhǔ* 主 'ruler, direct' respectively. Gù Yěwáng then introduced his own reading of a classical text: 'I, Yěwáng, observe that when "King Shùn commanded Bóyí 伯夷 to *diǎn* 典 his three rites" and "Xià to *diǎn* 典", *diǎn* 典 corresponded to this meaning', namely, 'direct'.³¹ Then he referred to a commentary on another text which defined *diǎn* 典 as *fǎ* 法 'law, statute'.³² He quoted the Ēryǎ for the gloss '*diǎn* 典 is like *jīng* 經 "canon, classic"', and *Shuōwén jiězì* for the statement that '*Diǎn* 典 is like the documents of the five Emperors. [The graph] has *cè* "bound documents" on a small table, where they are placed respectfully, as semantic constituents. Another explanation says "*diǎn* 典 are great documents".' Returning to the first person, he added 'I, Yěwáng, observe that in the *Book of Documents*, there are the Yáo *diǎn* 堯典 and the Shùn *diǎn* 舜典. According to Kǒng Ānguó's explanations "they can represent the eternal way/conduct of the preceding generations".' Finally, he turned to other ways of writing the character 'canon' and the character representing the word 'direct': 'The ancient graph is written 籒 and is classified under the bamboo radical. As for *diǎn* 𠄎 meaning "direct" it is classified under the radical pū 父.'

We can see that Gù Yěwáng gathered all the meanings of the character and the word *diǎn* 典 in the classics. He did not limit his presentation to concrete examples taken from the classics: he also referred to their commentators, added his own opinions, and reproduced glosses from older lexicographic works such as the Ēryǎ and *Shuōwén jiězì*. But what is probably more original is that he went as far as providing cross-references between entries within his dictionary. At the same time, his new thematic organization of the radicals facilitated looking for a character in the dictionary, and so did the number attached to the radicals as well as the total number of entries they gathered. The author of the *Yùpiān* intended his dictionary to be easy to consult. Thus, compared with previous lexicographic works, we can see in the original *Yùpiān* a clear step in the direction of a real dictionary. But the *Yùpiān* was

³¹ The texts cited are from the *Book of Documents*, in *Shísānjīng zhùshū*, 131b, 134b.

³² The original text is a passage in the *Book of Songs* which refers to 'The statutes of Wen Wang' (*Shísānjīng zhùshū*, 584c, 588b).

not a dictionary of contemporary meanings of words; it too was limited to the meanings or characters in the classics.

We can only regret that such a masterpiece was lost. *Gù Yěwáng's* dictionary had a strong influence on later works, such as *Lèipiān* 類篇 (1066), *Zìhuì* 字彙 (1615; see Chapter 6), and *Zhèngzìtōng* 正字通 (1680). Even in Japan, the famous Buddhist monk *Kūkai* 空海 copied it to write the *Tenrei banshō meigi* 篆隸萬象名義 (see Chapter 10), but suppressed the quotations from the classics as well as *Gù Yěwáng's* observations. For a long time the term *Yùpiān* was used in the general sense 'dictionary'.

A New Type of Dictionary: The Rhyme Dictionaries

With the growing importance of literary composition, different types of books providing the pronunciation of characters were produced. Some of them distinguished a certain number of rhymes or finals under which they classified the characters.³³ Most of the earlier texts have been lost, apart from the *Qièyùn* 切韻 (see Chapter 6), which was copied, recopied, and augmented, and had an immense influence, so as to create a new genre.

Conclusion

All the lexicographic works I have introduced here were compiled to help in the reading of classical texts or the composition of poetry. Thus, except for the *Shímíng*, none of them included a wide range of everyday words. The discovery of lexica dating from the *Táng* dynasty (618–907), at the beginning of the twentieth century in the desert of *Dūnhuáng*, shows that lists gathering everyday vocabulary according to thematic categories did in fact exist, and that dictionaries did not draw on these. Until the sixth century, monolingual dictionaries were the rule. The *Fāngyán* included non-Chinese items and thus constituted an interesting exception until the translation of the Buddhist canon from the fourth century onwards led to the compilation of certain bilingual lexica (Sanskrit/Prakrit/Pali–Chinese). However, these multilingual dictionaries dealt with a much smaller part of the total vocabulary than the older monolingual dictionaries.

With the increase in characters included in dictionaries from *Shuōwén* onwards, the need was felt to facilitate the retrieval of given entries. When

³³ Bottéro, 'Le développement des livres de rimes en dictionnaires'.

Xǔ Shèn invented the radical system of classification, he did not think of it as a way to find characters in his dictionary (although for all we know he may well have come to notice the advantages of the radical system for retrieval of lexical entries). He was primarily looking for a system to organize the characters which represented the structured realities of the world in his eyes. Gù Yèwáng considerably modified the system to make it easier to consult. The new system of classification on the basis of rhymes employed in the *Qièyùn* would offer a much more efficient way of retrieving lexical entries.

With the radical and rhyme systems of classification, lexical entries could not be words but had to be characters. However, since characters constitute the units of the Chinese writing system, and since characters write the roots of compound words, looking up characters was of great help for finding the meanings of even complex words consisting of more than one character.

Ancient India

LATA MAHESH DEOKAR

(With a Note on Tamil by Jean-Luc Chevillard)

The story of Sanskrit lexicography starts with the sacred texts of the Brahmanical tradition such as the *Ṛgveda* (c. 1500 BC), the earliest extant Indian literature. *Ṛgveda* contains hymns or prayers addressed to various deities for the fulfilment of desires. In the Vedic tradition, the utmost importance was given to the exact pronunciation of the hymns. It was believed that even a minor change in accent while chanting the hymns would not yield the merit which the sacrifice was supposed to bestow upon the sacrificer. This belief made it extremely necessary to commit the Vedas to memory with inordinate fidelity. In order to achieve this goal, several methods of recitation were invented. In a *saṁhitāpāṭha* 'continuous recitation', words are pronounced following the rules of euphonic combination, whereas in a *padapāṭha* 'word-by-word recitation', they are pronounced using their original form, thus suppressing the euphonic combination. In the *padapāṭha*, words are also analysed, thereby helping the interpretation of a verse. As Sumitra Katre rightly points out, the 'analysis of the continuous text into its constituent word forms was one of the major steps which led to the development of Indo-Aryan lexicography'.¹

In Vedic hymns, poets use a number of epithets of a deity along with his or her name. Such epithets are employed to praise the peculiar characteristics and heroic deeds of that deity. For example, Lord Indra is praised with a number of epithets such as *Maghavan* 'munificent one' and *Vṛtrahan* 'the slayer of [the demon named] *Vṛtra*'. Epithets complete the description of a deity and invoke the sense of devotion and respect for him or her in the minds of reciters. In a hymn addressed to a particular deity, the name and epithets of the deity occur in the same case, gender, and number. In late Vedic literature too, this tendency to use epithets continues. For example, in the famous hymn of the *Atharvaveda*, the poet invokes Mother Earth (*bhūmi*)

¹ Katre, 'Lexicography of Old Indo-Aryan', 2487.

with a number of epithets: ‘all-bearing, good-holding, firm-standing, gold-backed [-*vākṣas*], reposer of moving things [*jāgat*], bearing the universal [*vāiṣvānarā*] fire’.² At a later stage, we find that the epithets of a deity are recorded as synonyms of his or her name. In the *Nighaṇṭu* (to which I shall return shortly) and in the *Nāmaliṅgānuśāsana*, popularly known as *Amarakośa* after its author Amarasimha (to which I shall return in Chapter 7), some of the above-mentioned epithets are listed as synonyms of the earth. Epithets are, however, not synonyms in the sense of words having the same meaning. Rather they are co-referential in nature, referring to the same individual or thing. The Sanskrit grammatical tradition has a category of *paryāyavacanas* ‘convertible terms’, which is not quite the same as the category of synonyms, and epithets are best suited to this category.³

The Brāhmaṇa texts comment on the Vedic hymns. They provide meanings of some of the expressions found in the Vedas and also supply their folk etymologies. While doing so, they often explain a word by merely providing its synonym; for instance, the word *rātri* ‘night’ is explained by its synonym *kṣapā*.⁴ Elsewhere, a word is explained metaphorically with the help of another word when their meanings share some common characteristics. For example, due to the similarity of the duties of fire and a draft animal – namely that fire carries oblations to gods and a draft animal carries a load – *agni* ‘fire’ is called ‘a draft animal of gods’.⁵ Sometimes such a metaphorical synonymy results from superimposing the quality of one object on another. For instance, the expression *sutarmā nauḥ* ‘easily ferrying boat’ is used to describe a sacrifice, the hide of a black antelope, and speech, for all three help a person to cross all kinds of evils and reach heaven through a sacrificial performance.⁶ Similarly, since animals are moving (*jāgat*), they are associated with the verse metre named *jagatī*, which is derived from the word *jāgat* ‘moving’.⁷ Another principle governing such synonymic paraphrases is the principle of inner synonymy called *bandhutā*. According to this principle, the identity between two given words is established on the basis of their association. For example, *ghṛta* ‘clarified butter’ is associated with fire – butter is clarified by heating – and is then identified with a word for fire (*agni*).⁸ However, such metaphorical synonyms or those used on the basis of the principle of inner synonymy did not reach the status of full synonyms which could be lexicalized.

² *Atharvaveda*, 2.662 (12.1.6). ³ *Kāśikā*, I.1.68.2 and I.1.68.3. ⁴ *Aitareya Brāhmaṇa*, 9.

⁵ *Aitareya Brāhmaṇa*, 11. ⁶ *Aitareya Brāhmaṇa*, 10. ⁷ *Aitareya Brāhmaṇa*, 69.

⁸ *Taittirīya Brāhmaṇa*, 47 (III.8.14).

The Upaniṣads, the culmination point of Vedic literature, recognize synonymy as a phenomenon in a natural language. For instance, in the *Aitareya Upaniṣad*, *saṁjñāna*, *ājñāna*, and *viññāna* are recorded as synonyms of *prajñāna* 'knowledge'.⁹ In the Upaniṣads, we also come across a 'reciprocatory mode of stating synonymy'.¹⁰ For instance, the *Aitareya Upaniṣad* says 'this which is known as the heart, this mind'.¹¹

It is clear from the above discussion that the earliest literature of the Vedic tradition shows awareness of the concept of synonymy. As time elapsed, the need to prepare a list of obsolete words used in the Vedic texts and to explain the relation between synonyms must have arisen from the fact that the language of the hymns differed from that of the next generations, and the cultural context of many words became obscure. This, in fact, became the prime motivation for the development of Indian lexicography.

In order to explain the 'rare, unexplained, vague, or otherwise difficult terms' that occurred in the sacred Vedic literature, *Nighaṇṭus* 'wordlists' were compiled.¹² Scholars believe that there must have been many such wordlists, but nearly all of them have been lost. One survivor is the subject of a commentary, the *Nirukta* of Yāska, to which I shall return shortly. According to its twentieth-century editor Lakshman Sarup, the composition of the *Nighaṇṭu* can be traced back to an old tradition of sages.¹³ As to the question of the usefulness of these lists of words, Rudolf von Roth was of the opinion that, in earlier times, there was no real need for exhaustive commentaries for, in his view, 'a simple catalogue of the gods and the objects of worship as we have it in the *Naighaṇṭuka* sufficed as a guide to oral instruction'.¹⁴

The *Nighaṇṭu* is written in prose. It deals with all four parts of speech of the Sanskrit language, namely, nouns, verbs, preverbs, and indeclinables. It is divided into five sections. The first three sections (*naighaṇṭukakāṇḍa*) deal with synonyms. These sections are chiefly based on the *Rgveda*, the foremost among the Vedas, which alone demanded a really philological exposition. The fourth (*naigamakāṇḍa*) section treats homonyms, and the last (*daivatakāṇḍa*) deals with deities.

The *Nighaṇṭu* is, as I have just remarked, the subject of a commentary by a scholar known as Yāska; this commentary is called the *Nirukta*. Its date is uncertain, but may be about 500 BC. Its first section, in seventeen chapters, deals with physical objects such as earth and air, and objects of nature like

⁹ *Aitareya Upaniṣad*, 58 (5.2). ¹⁰ Dhadphale, *Synonymic Collocations*, 85.

¹¹ *Aitareya Upaniṣad*, 58 (5.2). ¹² Quotation from Vogel, *Indian Lexicography*, II.

¹³ In *Nirukta*, [part 2], introduction, 14. ¹⁴ Roth, *Introduction to the Nirukta*, 32.

cloud and dawn. Words related to man, his limbs, objects and qualities associated with man, and the like are treated in twenty-two chapters of the second section. The third section treats abstract qualities such as heaviness and lightness in thirty chapters. The fourth treats homonyms. Yāska divides the fifth section into six chapters. The first three chapters include the deities of the earth region and deified objects found therein, the fourth and the fifth chapters include the deities of the mid-region, and the last includes those of the highest region. This division seems to have influenced the arrangement of later lexica. The end of a particular list of synonyms is shown by repeating the last word of the list preceded by the word *iti* 'thus'. This arrangement shows the earliest attempt to group words methodically. Although it is hard to prove in each and every case, commentators of the *Nirukta* have tried their best to justify the sequence of synonyms listed in the *Nighaṇṭu*. Durga, a very early commentator on the *Nirukta*, justifies the propriety of the order of synonyms mainly on the basis of the descriptions of sacrifices. He also considers other associative and causal factors. For instance, according to him, words for 'strong' come after the verbs meaning 'eating' because those who eat become strong.¹⁵

It seems that Yāska is the first to define the concepts of synonymy and homonymy. At the beginning of the fourth section, Yāska says: 'Synonyms [literally "many words which have one meaning"] have been explained [in preceding sections]. Now therefore we shall take homonyms [literally "single words which have many meanings"] in their respective order.'¹⁶ Yāska further says that words having the same function are synonymous.¹⁷ Another discussion clearly points to the fact that interchangeability is yet another criterion of synonymy: while explaining the existence of numerous epithets of a deity, Yāska says that 'each receives many appellations on account of his supereminence, or the diversity of his function'.¹⁸ About Yāska's method of commenting upon the *Nighaṇṭu*, Lakshman Sarup observes that, 'At first he attributes a particular meaning to a particular word, and then supports his assertion by quoting a passage, generally from the Vedic literature, in which that word is used in that particular sense'.¹⁹ Sarup has shown the links between various Brāhmaṇa texts and the *Nirukta*, thereby making it evident that many of the etymologies found in the *Nirukta* have their basis in Vedic literature.²⁰

¹⁵ *Durgācāryakṛtavarṇanāmetam Niruktam*, 1.28. ¹⁶ *Nirukta*, [part 2], 56 (4.1).

¹⁷ *Nirukta*, [part 1], 48 (4.9). ¹⁸ *Nirukta*, [part 1], 135 (7.5) = [part 2], 115.

¹⁹ *Nirukta*, [part 2], introduction, 56. ²⁰ *Nirukta*, [part 1], 246–79.

Besides the *Nighaṇṭu* commented upon by Yāska, there exists another Vedic *Nighaṇṭu* in the tradition of the *Atharvaveda*, called the *Kautsavyanighaṇṭu*. It broadly follows the system of the earlier *Nighaṇṭu*. The word *nighaṇṭu*, however, did not remain in vogue to refer to the Sanskrit lexica in general. But in the later period it came to be used frequently for the Āyurvedic lexica, probably to invoke authority and sacredness.

Almost coeval with Yāska is the foremost grammarian of Sanskrit, Pāṇini, who wrote a grammar of the Sanskrit language in about 4,000 aphorisms, the *Aṣṭādhyāyī*. To his grammar are attached a lexicon of verbal roots and a list of nominal stems undergoing specific grammatical operations. The former contains 2,000 verbal roots classified into 10 classes depending upon their class markers. Within each class, the verbal roots are arranged according to their final consonant. The verbal roots, which are not accompanied by their meanings, are usually given as stems, except for a few cases where they are listed in the nominative. Although tradition ascribes both these wordlists to Pāṇini, and they have for many centuries been transmitted with meanings attached to the roots, scholars such as Sumitra Katre are of the opinion that the meanings are later additions, perhaps as late as AD 300.²¹

The gap separating the *Nirukta* of Yāska and the *Amarakośa* of Amarasimha, the first classical Sanskrit lexicon which is extant in its entirety, is approximately 1,000 years: as we have seen, the *Nirukta* may be from about 500 BC, and the *Amarakośa*, with which Chapter 7 will begin, is dated around AD 500. Only fragments of the Indian lexica of the thousand-year period between these two works survive. The earliest which can be dated with some certainty is Weber fragment no. 6, unearthed in East Turkestan, a manuscript of the sixth or seventh century AD preserving a text which appears to be rather earlier. It is extant only in eight folios and preserves parts of a synonymic lexicon.²²

Since Amarasimha does not name his predecessors, and mentions their works only in very general terms as 'other works', we do not even know the names of the pioneers of classical Sanskrit lexicography.²³ Amarasimha's commentators mention lexica such as Bhāguri's *Trikāṇḍa* and Vyāḍi's *Utpalinī* as his sources.²⁴ Hugga alias Ugra alias Rudra, Kātya, the author of the *Nāmamālā*, Vācaspati, Vararuci, Vikramāditya, the author of the *Samśārāvarta*, and Vopālita are some other lexicographers whose works

²¹ Katre, 'Lexicography of Old Indo-Aryan', 2488. ²² Vogel, *Indian Lexicography*, 18–19.

²³ Amarasimha, *Amarakośa*, I.1.2.

²⁴ Subhūticandra, *Kavikāmadhenu on Amarakośa*, 99–100; Sarvānanda, *Ṭīkāsarvasva*, 1.3; Rāyamukūṭa, *Padacandrikā*, 1.4–5; Bhānuji Dīkṣita, *Vyākhyāsudhā*, 2.

have not come down to us in their entirety. Robert Birwé considers Vyāḍi's *Utpalinī*, Vācaspati's *Śabdārṇava*, and Vikramāditya's *Samśārāvarta* to be later than Amarasiṃha.²⁵ Claus Vogel, on the other hand, believes that Vyāḍi preceded the *Amarakośa*.²⁶ He argues against Birwé that Vācaspati flourished before the fifth century AD, but agrees with him that Vikramāditya can be dated after the fifth century.²⁷ All these lexica have survived only in the form of citations scattered in later Sanskrit literature – scholars such as Birwé and Vogel have collected and edited the fragments of some of them – and, whatever their exact dates, they can be treated together here.

What we know of them is a matter of inference from fragments. Regarding Vikramāditya, for instance, Vogel remarks that, 'The allusion . . . to the origin of the fabulous river Jambu suggests that the *Samśārāvarta* too was rather detailed, and that Vikramāditya possibly was a Hindu by faith'.²⁸ We know that Vācaspati's *Śabdārṇava* dealt with synonyms, homonyms, primary derivative affixes, and gender.²⁹ Similarly, the available fragments of Vyāḍi's *Utpalinī* and Vikramāditya's *Samśārāvarta* are from two sections, namely, synonymic and homonymic. The Weber fragment mentioned above preserves parts of a synonymic dictionary. The evidence suggests that the makers of synonymic dictionaries dedicated a substantial portion of the text to synonymous words, and that they dealt with homonyms to a smaller extent. Additionally in these dictionaries we find small sections devoted to indeclinables, rules on gender, adjectives, and miscellanea.³⁰ It is possible to say that the division of the *Nighaṇṭu* was at the basis of this kind of classification found in later lexica.

Pali and Prakrit

Both Gautama the Buddha and Vardhamāna Mahāvīra, the respective founders of the śramaṇic religions, namely Buddhism and Jainism, chose to teach in the languages of the masses. The Buddha allowed his disciples to learn his teachings in their own language and admonished them not to deviate from common parlance. As a result, we find the teachings of both preserved in a number of popular languages, which linguists generally refer to by the name Middle Indo-Aryan. This term represents the languages which are mid

²⁵ Birwé, 'Fragments', 527.

²⁶ Vogel, *Indian Lexicography*, 17; cf. Vogel, *Zum Aufbau*, 15, where he places Vyāḍi in the fifth century AD.

²⁷ Vogel, *Zum Aufbau*, 10, 19. ²⁸ Vogel, *Indian Lexicography*, 18.

²⁹ Vogel, *Indian Lexicography*, 16. ³⁰ Vogel, *Indian Lexicography*, 16–18.

way between Old Indo-Aryan and New Indo-Aryan. From the formal point of view, this designation is correct. From the chronological point of view, however, one cannot say that the Middle Indo-Aryan languages came into existence after the Old Indo-Aryan ones. As later research has shown, both Old Indo-Aryan, which includes Sanskrit, and Middle Indo-Aryan, which includes various Prakrits, were used side by side. Prakrit languages varied from region to region and were named accordingly. Some of the more important among them are Śaurasenī, Mahārāṣṭrī, Jain Mahārāṣṭrī, Ardhamāgadhī, and Apabhraṃśa. Each of these languages is associated with a specific religious group or a certain stratum of society. For instance, followers of the Śvetāmbara school of Jainism used Ardhamāgadhī for their canon and Jain Mahārāṣṭrī for non-canonical literature, whereas those following the Digambara school used a particular form of Śaurasenī, which later came to be known as Jain Śaurasenī, for their holy scriptures. Mahārāṣṭrī was used in epics and lyrics and was also included in Sanskrit plays. Female characters and socially inferior characters of Sanskrit plays were depicted as using Śaurasenī and Māgadhī. As noted by Elisabeth Strandberg, 'As the Prakrits became literary languages, their original regional application was extended to a more or less theoretically established, partly socially stratified all-Indian use'.³¹

Pali, the language in which the canon of Theravāda Buddhism is preserved, is 'a composite, perhaps partly artificial, language believed to represent more than one regional idiom'.³² It appears to be closely related to Vedic Sanskrit but is not directly descended from it. The commentators of the Pali canon refer to it as Māgadhī, the language of the Magadha region, where the Buddha primarily gave his teachings. Due to its somewhat peculiar characteristics and its Buddhist affiliation, Pali enjoys a unique status among the Middle Indo-Aryan languages, and hence has often been treated differently from the Prakrit languages by the philologists of the past as well as the present.

Sanskrit was considered an identity symbol of Brahmanic culture. Sanskrit grammarians glorified it as divine speech and a means to earn merit and salvation. They attached a high social status to it and looked upon the Prakrit languages as inferior. In their grammars, they treated only *śabdas* 'words', that is, the Sanskrit items, and left out *apaśabdas* 'any form of language not Sanskrit'. Thus, they distinguished between Sanskrit words and the so-called

³¹ Strandberg, 'Lexicography of Middle Indo-Aryan', 2498.

³² Strandberg, 'Lexicography of Middle Indo-Aryan', 2498.

non-standard Middle Indo-Aryan words which existed during their times. This derogatory approach towards the Prakrit languages is also reflected in the words that are used to refer to these languages, such as *prākṛta* and *apabhraṃśa*. Of these, the former has been interpreted variously as ‘original, natural’; or ‘ordinary’; or ‘uncultivated, vulgar’; or ‘provincial, vernacular’. The word *apabhraṃśa* means, first of all, ‘falling down or away, a fall’; secondly, ‘a corrupted word, corruption; hence an incorrect word’; and thirdly ‘a corrupt language, one of the lowest forms of the Prakrit dialect used by people such as cowherds’. Since all these languages were considered to be inferior to Sanskrit, the scholarly community did not ascribe much importance to them. When, however, independent literary compositions were written down in these languages, Sanskrit lexicographers started to take a note of Prakrit dialects in the form of regional variants.

The first lexica of Prakrit and Pali languages were compiled in the tenth and thirteenth centuries respectively, that is to say, much later than the composition of the Jain and Pali canons. If one glances through both these canons and their respective commentarial literature, it becomes evident that lexicographical discussions formed an integral part of this literature. This might explain why writers in these traditions did not feel the compilation of a lexicon necessary at an early date. I shall return to these lexica in Chapter 7. Here I will discuss the emergence and gradual development of lexicographical material in both traditions.

Pali

After attaining enlightenment at the age of thirty-five, the Buddha spent the remaining forty-five years of his life wandering across northern India, teaching people the truth he had found. He wanted his teachings to be communicable to people from all walks of life. He gave more importance to content than expression. The Buddha mainly taught in the regions of Magadha and Kosala, that is, modern Bihar and the eastern part of Uttar Pradesh. This necessitated the use of various provincial language varieties (*janapadaniruttis*) in these regions. The Buddha’s rule of thumb in this regard was that ‘[o]ne should not insist on local language, and one should not override normal usage’.³³ He advised the monks that since in different provinces people use different words such as ‘a dish’, ‘a saucer’, or ‘a basin’ to refer to the same thing, they should use those different words so that people could understand at least one of them. The Buddha himself used such strings of synonyms in his

³³ *Middle Length Discourses of the Buddha*, 1080.

discourses, for instance, ‘*Siṅgi, suvaṇṇa* or *kañcana*, which is also called as *jātarūpa* and *hāṭaka*’.³⁴ Such verses carry the flavour of a type of versified synonymic lexicon, which became popular at a later period.³⁵ Such synonymic collocations are found in both metrical and prose passages of the canon.

Apart from understandability, M. G. Dhadhale has put forth six more reasons why the Buddha must have used strings of synonyms.³⁶ The first of these is to make meaning clear or to add stylistic grace to speeches. For instance, while explaining old age, the Buddha says: ‘And what, *bhikkhus*, is ageing-and-death? The ageing of the various beings in the various orders of beings, their growing old, brokenness of teeth, greyness of hair, wrinkling of skin, decline of vitality, degeneration of the faculties: this is called ageing.’³⁷ Such varied expressions in the Buddha’s speeches reveal the many related and fine aspects of the concept of old age. This, in turn, helps in getting nearer to the exact sense of the concept. It can also be considered as an approach to the concept of old age from different points of view. The second reason is to express the thought with great precision. For instance, Ambaṭṭha says to the Buddha that people of the warrior class ‘do not honour, respect, esteem, revere, or pay homage to Brahmins’.³⁸ All these words convey the meaning collectively. The third is to express subtle distinctions or delicate shades of meaning with the help of approximate synonyms, for instance, ‘sorrow, lamentation, pain, grief, and distress’.³⁹ The fourth is to encompass effectively the intended meaning, for instance, ‘whoever has developed the four roads to power, practised them frequently, made them his vehicle, made them his base, established them, become familiar with them, and properly undertaken them’.⁴⁰ The fifth is to emphasize a particular thing or an idea, for instance, ‘unwearying, zealous, and resolute’.⁴¹ The last is euphony, for instance, ‘can see Brahmā with his own eyes, talk with him face to face, and consult with him’.⁴²

³⁴ *Āṅguttaranikāya*, 1.215.

³⁵ It should be noted here that the Buddha used strings of synonyms not only of nouns, but also of verbs, for instance, ‘*ācikkhanti desenti paññāpenti paṭṭhapenti vivaranti vibhajanti uttānīkaronti*’ (*Dīghanikāya*, 2.105) = ‘teach it, declare it, establish it, expound it, analyse it, make it clear’ (*Long Discourses of the Buddha*, 246–7).

³⁶ Dhadhale, *Synonymic Collocations*, 3ff.

³⁷ *Saṃyutta-Nikāya*, 2.2 = *Connected Discourses of the Buddha*, 1.534.

³⁸ *Dīghanikāya*, 1.91, ‘*na brāhmaṇe sakkaronti na brāhmaṇe garukaronti na brāhmaṇe mānenti na brāhmaṇe pūjenti na brāhmaṇe apacāyanti*’ = *Long Discourses of the Buddha*, 113.

³⁹ *Dīghanikāya*, 1.36 = *Long Discourses of the Buddha*, 85.

⁴⁰ *Dīghanikāya*, 2.103 = *Long Discourses of the Buddha*, 246.

⁴¹ *Dīghanikāya*, 1.177 = *Long Discourses of the Buddha*, 157.

⁴² *Dīghanikāya*, 2.237 = *Long Discourses of the Buddha*, 307.

Such repetition of synonyms is perhaps the most outstanding characteristic of the Pali canon. Even a cursory glance at the Buddha's teachings reveals that the Pali canon contains a huge collection of synonymic collocations. In some canonical texts of a commentarial nature, such collocations are used as glosses on a particular word. For example, the word *kāmayamānassa* 'the one who is desiring' is explained with a number of synonyms, namely, *icchamānassa*, *sādiyamānassa*, *patthayamānassa*, *pihayamānassa*, and *abhiyappamānassa*.⁴³ Sometimes even figurative expressions culled from discourses are used as glosses. For example, *paññā* 'wisdom' is glossed with figurative expressions such as 'wisdom as a sword', 'wisdom as a height', 'wisdom as light', 'wisdom as glory', 'wisdom as splendour', and 'wisdom as a precious stone'.⁴⁴ The length of such synonymic clusters varies from two to eight and even more. These can be further classified into dyads, triads, tetrads, pentads, and so on. They seem to have been arranged keeping in mind the number of syllables in a word as well as the phonetic affinity between the two synonyms, for example, *phasso phusanā samphusanā samphusitattam* 'contact: touching, the being brought into contact, the state of having been brought into touch'.⁴⁵ In his book *Style and Function*, Mark Allon has noted this peculiar feature of the arrangement of synonyms in ascending order of their number of syllables, calling it the principle of waxing syllables.⁴⁶ According to K. R. Norman, 'The beginnings of Pali lexicography probably lie in the system of explanation by means of synonyms'.⁴⁷ The same tendency of synonymic exegesis is also visible in later canonical works and the early commentarial literature.

In the Pali canon, while listing the accomplishments of a learned Brahmin, the word *nighaṇḍu* 'a collection of words or a treatise on lexicography' is used. The fifth-century commentator Buddhaghosa explains it as 'a collection of nouns, the science of explaining the synonyms [of words], beginning with *rukkha* "a tree" and so on'.⁴⁸ This indicates that Buddhaghosa was aware of some lexicon that had a list of words beginning with the word *rukkha*. However, no such lexicon prior to Buddhaghosa has so far come to light. Interestingly, just such a list is found in a late twelfth-century grammatical text, the *Saddanīti*, composed by a Burmese scholar, Aggavaṃsa.

Peṭakopadesa and the *Nettipakaraṇa* are the two Pali texts (neither of which is later than the fifth century AD) that use synonyms as one of the methods of

⁴³ *Mahāniddesa*, 2. ⁴⁴ *Dhammasaṅgaṇi*, 11 = *Buddhist Manual of Psychological Ethics*, 18.

⁴⁵ *Dhammasaṅgaṇi*, 9 = *Buddhist Manual of Psychological Ethics*, 6.

⁴⁶ Allon, *Style and Function*, 191. ⁴⁷ Norman, *Pāli Literature*, 166.

⁴⁸ Buddhaghosa, *Sumaṅgalavilāsinī*, 1.247.

conveying the exact sense of a canonical expression. The latter also provides lists of synonyms collected from various canonical works. For instance, it provides more than fifty synonyms for *nibbāna* 'liberation', nearly half of which are taken from an *Uddāna* 'list of words' from the *Saṃyuttanikāya*.⁴⁹ Both these texts preserve lists of epithets of the Buddha and synonyms of certain concepts. Sometimes, the number of epithets or synonyms found in these two texts exceeds the number recorded in the only Pali lexicon: the *Abhidhānappadīpikā*, composed in Sri Lanka in the thirteenth century.

Buddhaghosa's commentaries on the Pali canon are full of lexicographical discussions. He defines *adhivacana*, *nirutti*, and *paññatti* as three different modes of exegetical explanations resulting in synonymy. According to him, in the *adhivacana* type of exegesis a word is explained by providing an alternative expression. For example, 'luck-bringing' is explained as 'wealth-bringing'. Here 'wealth' is a near-synonym of 'luck'. Elsewhere the word *adhivacana* also signifies metaphorical expression, which can be treated as another category of synonymy. For example, the word *raja* ('dust') is metaphorically used to signify 'greed'.⁵⁰ *Nirutti* 'etymology' is a mode of exegesis in which a concept is analysed linguistically. Such a linguistic analysis, in turn, provides a justification for using a particular word to represent the concept under discussion. For example, 'mental formation' (*saṅkhāra*) is explained as 'something which forms' (*abhisankharoti*). This linguistic analysis justifies the use of the word *saṅkhāra* to represent 'mental formation'. In the third mode of exegesis called *paññatti* 'description', a particular notion is described in different ways. For example, 'thought' is explained by using three different words, namely, *takka*, *vitakka*, and *saṃkappa*.⁵¹

In his *Aṭṭhasālinī*, Buddhaghosa points out two main functions of synonymic collocations, namely, to meet the needs of persons of different calibres and as an embellishment of the discourse.⁵² Thus, while explaining the reason behind the existence of 'five different *suttas* almost verbatim but for the difference in their (synonymous) titles (namely, *Ābhāsutta*, *Pabhāsutta*, *Ālokasutta*, *Obhāsasutta* and *Pajjotasutta*)' Buddhaghosa says:

It is true that the sense could have been conveyed by any of them singly. However, all of them are used with a view to meeting the inclinations [*ajjhāsaya*] of various persons. The meaning when diversified [literally,

⁴⁹ *Saṃyuttanikāya*, 3.320. ⁵⁰ Buddhaghosa, *Jātakatṭhakathā*, I.117.

⁵¹ Buddhaghosa, *Aṭṭhasālinī*, 51.

⁵² Buddhaghosa, *Aṭṭhasālinī*, 47; see Dhadphale, *Synonymic Collocations*, 5.

‘divided’] becomes ‘well-specified’ as the persons of different calibre can understand them in their own way.⁵³

The phenomenon of synonymy is explained in the commentary to the *Mahāniddesa* as extending to cover the cases of roughly synonymous words which differ only in some of their sounds; words which are not fully synonymous but are derived from the same verbal stem with different prefixes; and words which are not fully synonymous but have related meanings.⁵⁴ For instance, the words *kodho*, *kujjhanā*, and *kujjhitattam* all mean ‘anger’. Since these words differ only with respect to a few syllables they often result in an alliteration. The verbs *ijjhati* and *samijjhati* both mean ‘succeeds’ and are derived from the same root *idh-*; while *pañdiccam* (‘erudition’), *kosallam* (‘proficiency’), and *nepuññam* (‘cleverness’) have different but related meanings. Again, the *Mahāniddesa* classifies synonyms on the basis of manifoldness of expression, characteristics, performance, and negation. For example, in the first type of synonym, anger (*kodha*) is expressed by two different expressions, *vyāpāda* and *dosa*. In the second type, although *rūpa* (‘form’), *vedanā* (‘feelings’), *saññā* (‘perceptions’), *cetanā* (‘volitions’), and *viññāṇa* (‘consciousness’) are synonyms as far as all are aggregates of human personality, they differ from each other on account of their specific characteristics. In the third type, the four categories of right effort (*sammappadhāna*), though synonymous in so far as each is an exertion, differ from each other on account of the different functions each one performs. In the fourth type, *kodhagarutā* (‘weightage to anger’) is synonymous to the negation of its opposite, that is, *saddhammagarutā* (‘weightage to sublime righteousness’).⁵⁵ This kind of classification of synonyms is highly characteristic of the Pali lexicographical literature. Thus, in the case of the Pali literature, the need of a lexicon was satisfied by its rich exegetical literature to a great extent. Even when a lexicon like the *Abhidhānappadīpikā* was composed, scholars primarily relied on the exegetical literature rather than the lexicon for the explanation of primary texts.

Prakrit

Mahāvīra Jina, a senior contemporary of the Buddha, also taught in the local dialects rather than in Sanskrit. The tendency of using synonymic collocations is visible in his teachings as well. In the Jain canonical literature, we

⁵³ Buddhaghosa, *Aṭṭhasālinī*, 1.48 = Buddhaghosa, *Expositor* 1.196–7, quoted in Dhadhphale, *Synonymic Collocations*, 5.

⁵⁴ *Saddhammapajjotikā*, 1.19. ⁵⁵ *Saddhammapajjotikā*, 1.19–20.

come across words such as *egatṭha* ‘having one meaning’, *nānāghosa* ‘various expressions’, and *nānāvamjaṇa* ‘different manifestations’.

So, for instance, in one major canonical text, the *Bhagavatī*, Gautama Svāmin asks Mahāvīra, ‘Lord, do these nine words *calamāṇa* “that which is being moved”, *calia* “that which is moved” etc. have the same meaning but various expressions and different manifestations or do they have different meanings, different expressions and different manifestations?’ Mahāvīra answers:

Amongst these, [the words] *calamāṇa* and *calia*, *udīrijjamāṇa* ‘that which is being uttered’ and *udīria* ‘that which is uttered’, *vedijjamāṇa* ‘that which is being known’ and *vedia* ‘that which is known’, and *pahijjamāṇa* ‘that which is being abandoned’ and *pahīṇa* ‘that which is abandoned’ are synonyms. The remaining five have different meanings, different expressions, and different manifestations.⁵⁶

Similarly, the *Oghaniryukti*, a text about the details of a monk’s life, lists eight synonyms of *uvahi* ‘fraud, circumvention’.⁵⁷ These denote various stages and characteristics of *upadhi*. Like Buddhaghosa, the author of another canonical text, the *Jambūdvīpaprajñaptiṭīkā*, also holds that ‘synonymous words are used for the benefit of disciples coming from various countries’.⁵⁸ The *Bṛhatkalpabhāṣya* says that a sage should know many languages (dialects) so that he will be beneficial to the masses. It gives a list of four purposes of synonyms. These are to replace one word with its synonym so as to fit in a verse; to achieve brevity; to make the understanding of a text easier; and to exhibit the richness of vocabulary of a teacher.⁵⁹ According to Somaṇī Kusumaprajñā, in ancient times, each subject was taught with the help of twelve methods. Among these twelve, synonymous words played an important role. A student was not expected to learn a dictionary by heart. Rather, during the course of a discourse on a particular topic several synonyms were used so that the student would learn them by heart.⁶⁰

To conclude, Indian lexicographical activity of the period before AD 500 tended not to result in free-standing wordlists, with the exception of a few Sanskrit lexica, some of doubtful date, many of which are now lost. It was, however, rich in discussions regarding the phenomena of

⁵⁶ *Bhagavatī*, 1.12.

⁵⁷ *Oghaniryukti*, 666; see Kusumaprajñā in *Ekārthak kośa*, Introduction, 15.

⁵⁸ *Jambūdvīpaprajñaptiṭīkā*, 33. ⁵⁹ *Bṛhatkalpabhāṣya*, 1229.

⁶⁰ See Kusumaprajñā in *Ekārthak kośa*, Introduction, 15.

synonymy and homonymy, and produced scattered strings of synonyms in Sanskrit, Prakrit, and Pali. Its further developments will be surveyed in Chapter 7.

Tamil

The Classical Tamil lexicographical tradition has its most explicit origin in the seventeenth chapter, called *Uriyiyal*, of a versified treatise called *Tolkāppiyam*. This is the sandhi form of *Uri Iyal*, where *Iyal* is (here) ‘chapter’ and *uri* stands for *uriccol*, a word to which we shall return. The chapter comprises 131 metrical lines, expressing 98 rules (or *cūttirams*), among which 82 consist of a single metrical line, and the longest, which is the first, consists of eight.

The *Tolkāppiyam* as a whole contains twenty-seven chapters, in three books, *Eluttatikāram*, *Collatikāram*, and *Poruḷatikāram* (conventionally cited as TE, TC, and TP respectively), of nine chapters each.⁶¹ Its total length is 1,610 *cūttirams*, comprising 4,013 metrical lines, and taking approximately five and a half hours for a complete recitation. Recitation, and not just reading, probably did take place: it is generally believed that such texts were memorized at the time when they were in use, although the practice is no longer alive for Tamil (it continues for Sanskrit in some places).⁶² The *Tolkāppiyam* can probably be dated roughly in the first half of the first millennium AD, although later and earlier dates have been proposed. According to an early tradition, it takes its name from that of its author, *Tolkāppiyaṇ* (nowadays politely referred to as *Tolkāppiyaṇār*, with honorific plural ending *-ār* added).⁶³ It must have been meant as a reference guide, or as a training manual, for those who desired to become professional *pulavars* (‘scholar-poets’), as is visible in the frequent occurrence of statements of the pattern ‘scholar-poets say [that] P’ (*P enmaṇār pulavar*).

The *Tolkāppiyam* contains elements of information which would be described today as relevant for phonetics, morphology, syntax, semantics, and other kinds of linguistic enquiry, and were included because they pertain

⁶¹ The editions cited here are *Tolkāppiyam Eluttatikāram*, with the commentary by Nacṇārkkiniyar (1937); *Tolkāppiyam Collatikāram*, with the commentary by Cēṇāvaraiyar (1938); *Tolkāppiyam Poruḷatikāram*, with the commentary by Teyvaccilaiyār (2003); and *Tolkāppiyam Poruḷatikāram*, with the commentary by Iḷampūraṇar (2003). The text is divided differently in the versions transmitted with different commentaries, and so the commentary is specified in references: TC18c below refers to the eighteenth *cūttiram* of *Tolkāppiyam Collatikāram* as commented on by Cēṇāvaraiyar.

⁶² The stated duration of a recitation of the *Tolkāppiyam* is based on the recording published as *Tolkāppiyam Murrōtal*.

⁶³ The name is discussed by Irākavaiyaṅkār, *Ārāyccit Tokuti*, 97.

to the composition of poetry in a language variety referred to as *Cen-Tamiḷ* ‘perfect Tamil’ (compare the Sanskrit phrase *saṃskṛtā vāk* ‘refined speech’, which gives the Sanskrit language its name, meaning ‘refined’). Within the framework of *Cen-Tamiḷ*, two ‘ways’ of Tamil are distinguished, *ceyyuḷ-āru* ‘the way of poetical composition’, and *valakk-āru* ‘the way of usage’: so, for instance, TC18c explains that non-specifying (decorative) epithets are used in *ceyyuḷ-āru* but not in *valakk-āru*. Regarding the latter term, a *cūttiram* (TP638i) from the twenty-seventh chapter makes it clear that it is only the superior usage which is under consideration. Indeed, the general impression gained through the reading of the *Tolkāppiyam* is that the central preoccupation of the group of professional *pulavars* was poetical composition (*ceyyuḷ*), rather than anything like what we would call ‘general linguistics’, as appears for instance from the fact that the longest chapter in the *Tolkāppiyam* deals with metrics, and that most of the topics dealt with in the third book describe various conventions pertaining to a poetical corpus (to which new poems can be added by students of the *Tolkāppiyam*). As for the first two books, it can be argued that, although *Eluttatikāram* deals with phonetics/phonology (and especially with sandhi) from a point of view like that of modern linguistics, the information which it contains is primarily of interest to a practitioner of metrics. Similarly, *Collatikāram*, in which the *Uriyiyal* is the penultimate chapter, deals mostly in a general manner with morphology (its main topic) and with syntax, but even in this book there are several passages where the focus is clearly literary usage.

Coming now to the place of lexicography within the *Tolkāppiyam*, the briefest manner of characterizing the *Uriyiyal* (which represents 3 per cent of the *Tolkāppiyam* as a whole, and which can be recited in eleven minutes) is to say that it is a sample dictionary, providing the meanings of 120 difficult – or ‘infrequent’ (*payilāta*) – words.⁶⁴ These are collectively labelled as *uric col*, an expression which can be literally translated as ‘proper words’ or ‘appropriate words’, although there is more than one explanation of the label. The easiest explanation – mentioned as the opinion of ‘others’ by the medieval commentator Cēṇāvaraiyar – is that they are ‘appropriate for poetry’.⁶⁵ (It would be tempting to explore the related possibility that the lexica descended from the *Uriyiyal*, for which see the section on Tamil in Chapter 7, were, like synonym dictionaries in some other traditions, devised as a tool for the poet

⁶⁴ For a detailed description, see Chevillard, ‘Rare words’, 302–12.

⁶⁵ This opinion seems to be traceable to the first *cūttiram* in the *Uriyiyal* chapter of the *Namūḷ*, a thirteenth-century grammar, which precedes the *Tolkāppiyam* commentary by Cēṇāvaraiyar.

to find the appropriate word.) However, this explanation does not seem to be the preferred one for most *Tolkāppiyam* commentators (they may of course be mistaken, since they were writing long after the composition of the primary text). In the course of its eight metrical lines, the first *cūttiram* in the *Uriyiyal* (TC297c) evokes, densely, the following three ideas. First, there is a triad of possible manners of originating of the *uric col*: they express a sound (*icai*), an idea (*kurippu*), or a quality (*panpu*). Secondly, their status as far as the opposition between verb and noun is concerned is indeterminate (or ambiguous): ‘They can be concretely undistinguishable from either noun or verb.’ Thirdly, there are various possibilities – one-to-many and many-to-one – for the appropriateness of the relationship between word and meaning (the possibility of a one-to-one relationship was probably considered too obvious to mention): ‘It can happen that one [*uric col*] is appropriate for several values, or that several [*uric col*] are appropriate for one [and the same] value.’ Finally, the practical way of dealing with *uric col*, which is what occupies the greatest part of the *Uriyiyal*, is touched on. It is not easy to give a crisp rendering in English of what the commentators on this passage seem to be driving at when using the expression *uric col*, especially since two (or three) simultaneous lines of explanation are followed by them, constrained as they are by the wording of the original. The commentator Teyvaccilaiyār states helpfully that the purpose of the *Uriyiyal* is to deal with the equivalent of what Sanskrit grammarians call *tātu* (the Sanskrit word is actually *dhātu*), in other words, verbal roots.⁶⁶ In French, one might say that the chapter called ‘*Uriyiyal*’ deals with ‘*Le propre du mot*’: the sense which remains once one has removed the various suffixes. Let us explore this point further.

The general context inside which such a characterization appears to be a step is corpus-based word analysis. The underlying practical question being examined by the redactor of the difficult passage under discussion here seems to be the following: if we imagine a corpus which has been divided into words, categorized into two main categories, nouns and verbs, and if we remove from the individual words the various suffixes (such as case marks and gender-class marks), or more precisely *ītaic col* ‘interstitial word(s)’, by which the nouns and the verbs can be identified and which have been tentatively enumerated elsewhere in the *Tolkāppiyam*, what can we say about the remaining (radical) part?⁶⁷ The implicit answer found in

⁶⁶ Teyvaccilaiyār does not seem to mention the nominal stems, covered by the Pāṇinian Sanskrit technical expression *prātipadika*.

⁶⁷ Categorization as nouns and verbs is indeed the dominant scheme in the *Tolkāppiyam*, as evidenced by a *cūttiram* such as TC158c: ‘The knowledgeable say that [items] fit to be

the second line of the passage under discussion is that we are left with something which (semantically) originates either in a sound, or in a notion, or in a quality, and that this something is neither a noun nor a verb, although it could potentially become either of those. This point of view is in accordance with what is seen in the ultimate *cūttiram* of *Eluttatikāram*, the first book of the *Tolkāppiyam*, where the same triad (*icai*, *kurippu*, *paṇṇu*) is mentioned, in connection with items which are referred to as *kuṛaic cor kīlavi* 'linguistic expression which is a [morphologically] incomplete word'. Those items are mentioned as part of a (conclusive) enumeration of the topics which have *not* been dealt with in *Eluttatikāram*. We must view this negative statement as a sign of the fact that the *Tolkāppiyam* was somehow considered as part of a work in progress.⁶⁸

As a final statement for this introductory section, I must add that if the seventeenth chapter of the *Tolkāppiyam* is the main precursor of what will become the autonomous Tamil lexicographic tradition, it is not the only one, and several other parts of the *Tolkāppiyam* should be mentioned – for instance, a sequence of twenty-four *cūttirams*, totalling forty-four metrical lines, where the designations for the males, the females, and the young of various animals are enumerated (TP545i to TP568i) – but such an endeavour would take us too far.

called words are those two: name and action.' However, the following *cūttiram* (TC159c) immediately asserts the existence of two other types of *cols* 'words', namely *īṭaic col* and *uric col* (and each has its own chapter). A general characterization of the *īṭaic col* 'interstitial word(s)' (or 'particles'), stating that they are not autonomous and accompany nouns and verbs, is found in TC249c, followed immediately by an enumeration of the types of particles.

⁶⁸ Another illustration of the same point is that a discussion of one of the kinds of particle, verbal tense marks, is called for, but never takes place. A complete description of Tamil verbal morphology was in fact produced many centuries later, by Western missionaries (for whom see Chapter 30). Obviously, the need was not felt before that.

The Greco-Roman World

ROLANDO FERRI

Scholars specializing in Greco-Roman lexicographical texts are very insistent on distinguishing between lexica, where the organizational principle is alphabetical, and glossaries, following the order of a source text from which all lemmata are culled, whether selectively or not. The third important category of lexica is that of the *onomastica*, organized by semantic areas and not alphabetical.¹

Almost all Greek lexica extant in the direct tradition date from the Byzantine period (for which see Chapter 12) as a result of the fact that later lexicographers tended to supersede earlier works and caused their disappearance. However the indirect tradition, and, from the end of the nineteenth century onwards, the evidence provided by papyrological discoveries help to shed light on the history of a much older, and wide-reaching, tradition, with a high degree of methodological awareness, and a carefully thought-out range of typologies for textual presentation on the written page.

The major input to the study of the Greek vocabulary came from the great flourishing of Alexandrian scholarship from the third century BC onwards – a date which, as we shall see, is also that of the earliest extant documents of lexicographical activity. Alexandrian scholars, however, often criticize earlier students of Homeric vocabulary under the collective name of *glōssographoi*.² These were presumably authors of collections of obscure words in Homer which were used for educational purposes (Homer was the foremost school text, especially at the elementary and intermediate levels).

There are eighty-one records in the *Mertens-Pack*³ online database of paraliterary papyri under the heading ‘glossaries and wordlists’, with the earliest finds going back to the third century BC.³ The oldest type of lexicon

¹ See Goetz, ‘Glossographie’, and Tolkiehn, ‘Lexikographie’.

² Lehrs, *De Aristarchi studiis Homericis*, 43–6.

³ An edition of an early lexicographical papyrus is included in Ucciardello, *Hypomnemata papiracei e lessicografia*.

must have been Homeric, either in the form of a list of select words no longer understandable for the student and glossed with a more modern equivalent, or in the form of a continuous word-for-word facing 'translation' of a passage into more current Greek. The latter class goes under the name of *scholia minora*, to distinguish them from proper lexis because they follow the order of the source text.⁴ *Scholia minora* are not normally included in studies of lexicography, but the distinction is slight, and the educational purpose was certainly the same. A typical specimen of this class of texts is a papyrus of the Michigan collection, *P. Mich. inv.* 1588, perhaps dating from the first century AD.⁵ It contains the beginning of the *Iliad*, with original and paraphrase – or more exactly translation – placed in two parallel columns. Here the famous initial word of the *Iliad*, *mēnin*, is matched, in the right-hand column, by three explanatory translations, *cholon*, *orgēn*, and *thymon*, all meaning 'anger', but more ordinary in register and supposedly more familiar to a student. The second line has the second word, *aeide*, matched in the parallel column by the imperative *aide*, followed by the explanation *ho estin humnei*, 'which means sing'.

As first described in a seminal work by the Greek-American scholar Mark Naoumides in 1969, from the earliest period, Greek lexicographic works adopt one of two prevailing page layouts in the organization of their materials. The first is one in which the text is a compact continuous unit, but the lemma always stands at the beginning of a new line, regardless of the point where a preceding entry had ended, and sometimes protrudes out from the writing space of the papyrus column, or *selis*, a position the ancient sources call *ekthesis*. Lemmata are also highlighted by a horizontal line placed underneath the writing line, the *paragaphos*, and they are sometimes separated from the explanation (called sometimes *metaphrasis*) by a raised dot (·). (In the Byzantine period, this layout was abandoned, with the text written as a continuous run, where lemmata are found in all positions within the writing space, though the use of colours and larger initials sometimes makes up for the loss of visibility associated with the new format.)⁶ Among early lexicographical papyri, this format occurs in *P. Oxy.* 1801 (first century AD), an alphabetical lexicon (in AB- order) of words occurring

⁴ *Scholia minora in Homerum: An Alphabetical List*, ed. Lundon, is an alphabetized edition bringing multiple lists together; *Scholia minora in Homerum*, ed. Montanari et al., is a database giving access to the individual lists.

⁵ First published in Renner, 'Three new Homeric'; a more recent edition (by F. Montanari, 2007) is in *Scholia minora in Homerum*, ed. Montanari et al., under *P. Mich. inv.* 1588.

⁶ On the same themes, cf. Esposito, 'Fragments of Greek lexicography'.

mostly in comedy.⁷ Most lemmata found in *P. Oxy.* 1801 occur in later lexicographers, for example in Hesychius, but they are treated here in a much fuller format, with quotations, and often including diverging interpretations offered by earlier scholars, perhaps commentators on comic texts. For example, two entries read:

bembix [spinning top]: a spinning movement. Aristophanes [says] 'be like spinning tops'. Some [interpret this] as a figure of tragic dancing . . .
bereschetoi fools: a word made up by Aristophanes.⁸

The second textual layout prevails in word-for-word paraphrases and later in bilingual glossaries, and has the lemma and its 'translation' laid out as two parallel columns on the page – a layout which takes up a lot of space but has the immediate advantage of visibility and highlights the important elements of each entry. This is, for example, the format of *P. Mich. inv.* 1588 described above. Eleanor Dickey has argued that this feature was an innovation of the bilingual books of the Romans, but in my view the Greek monolingual tradition, which is earlier, and is organized exactly like later bilinguals, cannot be left out of this picture.⁹

The earliest of the lexicographical papyri, *P. Hibeh* 175, dates back to the period of the flourishing of Alexandrian scholarship, 260–240 BC, and is evidence of the lexicographical work of the circles of the great librarians, whose need for accurate editions and commentaries naturally led to producing specialized essays focused on language and vocabulary.¹⁰ *P. Hibeh* 175 is indeed a collection of Homeric and epic words, in alphabetical order and arranged in a bicolumnar format. Unfortunately, only two fragments are legible, of words beginning in *de-* and words beginning *eu-*, with glosses in more current Greek: so for example, in fragm. 2, 18–20, we read, in three subsequent lines,

⁷ This is the 'Theon' glossary, so called because Hesychius mentions this grammarian as an author of *lexeis kōmikai*. Most entries are transcribed and translated, with an ample commentary, in *Commentaria et lexica Graeca in papyris reperta*, under their respective authors (mostly Alexis and Aristophanes).

⁸ I quote the text here from Austin, *Comicorum Graecorum fragmenta in papyris reperta*, 342–3: '*bembix* peristrophē. Aristophanēs "bembikes engenesthōn". enioi de schēma ti tragikēs orchēseōs . . . *bereschetoi* anoētoi: peplastai par' Aristophanei'. Here and elsewhere, when quoting papyrological evidence, I omit most diacritics for supplements and uncertain letters for the sake of clarity.

⁹ See Dickey, 'Columnar translation'.

¹⁰ The first of the librarians, Zenodotus of Ephesus, wrote a book of *Homeric glosses*: it was presumably a collection of supporting material for the critical edition of Homer ascribed to this scholar. On this topic see generally Montana, 'Hellenistic scholarship', 102; Tosi, 'Lessicografia'; Tosi, 'Typology of Greek lexicography'.

Eukairon	tachu	(‘well-timed’)
Euchos	kleos	(‘glory’)
Euthunei	kubernai	(‘he directs’).

The same area in which *P. Hibeh* 175 was excavated brought to light *P. Hibeh* 172, written at the same period or earlier. It is a wordlist of poetic nominal composita, which Rudolf Pfeiffer was inclined to see as related to the famous lexicon of Philitas the Coan, the *Ataktoi glōssai*, ‘rare words with no arrangement’: this was, however, only a list, with no explanatory material provided.¹¹

Another scholar-poet from the early Alexandrian period who was active as a lexicographer was Callimachus of Cyrene, author of the *Ethnikai onomasiai* (‘Names given by different populations’).¹² No fragments are reconstructable, but it was certainly a work dealing with regional differentiation within Greek, with perhaps the occasional foray into what ‘barbarians’ also called a given object. It was almost certainly a precursor of the genre later called *onomastica*, containing lists of words arranged by semantic fields: a later source, the Byzantine *Suda* (for which see Chapter 12), also registers the headings of some of its separate sections, for example ‘Denominations of the months by nations and cities’ (‘*mēnōn prosēgoriai kata ethnē kai poleis*’), ‘Names of fish’, ‘Names of winds’.

The founding father of ancient lexicography must have been the late third-century Aristophanes of Byzantium.¹³ Aristophanes seems to have composed a work perhaps titled *Lexeis Attikai* (‘Attic words’; a title only extracted from one of the three extant independent manuscript witnesses, *Parisinus* gr. 1630), as well as a number of thematic essays pivoting on the analysis of special fields of lexicographic interest, such as the denominations for age groups in humans and animals (‘*peri onomasias ēlikōn*’), and the ‘appellatives for degrees of affinity’ (‘*peri sungenikōn*’).¹⁴ What we have is difficult to evaluate, owing to the fragmentary nature of the texts. The nature of the text seems to be similar to that of Pollux’s *Onomasticon* (for which see below), a format evidently

¹¹ Pfeiffer, *History of Classical Scholarship*, 94. It is, however, odd to find entries like l. 110 *kampuloprunnos*, ‘with a curved stern’, which are not found in any poetical text, but only as an interpretative gloss for the Homeric word *korōnis*, ‘curved’, in scholia on Homer *Iliad*, 2.392 (Scholia D in *Iliadem* [2014], 114), and in Hesychius.

¹² Callimachus, *Fragmenta*, item 406; Schoenemann, *De lexicographis antiquis*, 44; Blum, *Kallimachos*, 135–6.

¹³ See H. Erbse in *Lexica Graeca minora* (1965), xiii.

¹⁴ All these works are preserved only in fragments and in a number of indirect sources, especially the twelfth-century Homer commentator Eustathios of Thessalonica.

favoured at this early stage, comprising a sequence of chapters organized by semantic fields and themes, and encyclopedic in nature: we only have sections relating to human anatomy, animals, kinship, and archaisms.¹⁵ It is, however, difficult to establish if Aristophanes had in mind an encyclopedic project similar to Pollux's or if these essays only serendipitously foreshadow later developments. Aristophanes, however, was alien from the linguistic biases of later Atticist lexicographers, who strove to teach correct (that is, Attic) usage. In his essay on the 'Words suspected not to have been used by the ancients', Aristophanes was able to find good linguistic precedents for a number of idioms other grammarians of his time censured as incorrect.

Aristophanes was also a pioneer in the attempt to identify geographical and diachronic variant forms in Greek, even if the exact framework for the study and listing of these forms remains unclear.¹⁶ He was also alert to register variation: in his study of the terms used for defining kinship, he paid great attention to various types of address forms marking respect, proximity, deference, or distance, as well as to diachronic and diatopic variations of usage:

there are different address forms, some more playful, others endearing . . . those who talk to older people may call them just 'father', even if they are strangers to them, and 'mother' older women, making a guess about the age, whereas even older women they address with *maia* ['good mother'] or *tēthē* ['grandmother'].¹⁷

The sources of Aristophanes were mostly written and literary, but he was able to draw on epigraphic texts, and certainly referred occasionally to contemporary spoken language, the so-called *koine*, the speakers being sometimes more precisely identified as *hoi Alexandreis*, 'the Alexandrians', that is the residents of Alexandria of Egypt, where Aristophanes spent a large part of his life. Aristophanes was also attentive to polysemy, both in the *Lexeis* and in his other fragments, as a natural consequence of his engagement with literary and textual criticism throughout his life.¹⁸ Aristophanes sees polysemy as the

¹⁵ There is an edition in Miller, *Mélanges de littérature grecque*, 427–34.

¹⁶ Nauck, in his edition of Aristophanes, *Fragmenta* (1848), 76–7, tentatively attempted to identify two competing arrangements, one similar to Pollux's, and encyclopedic in scope, and one regional, listing special Greek words used by different Greek nations.

¹⁷ W. J. Slater, paraphrasing Aristophanes, *Fragmenta* (1986), item 241: ἡὸς εἰσι προσφῶνῆσαι διαφωροὶ παῖνιόδεσται τῖνες καὶ ὑποκοριστικαί, ἐπαγεῖ, ἡοῖον *appa*, *pappa*, *mamma*, *mammē*, *mammia*, *tetta*, *atta*. houtō de, phēsi, kai pateras men diarrhēden tous presbuterous kalousi, kan ōsin allotrioi, mēteras de tas presbuteras, tēn hēlikian eikazontes, tas de eti presbuteras maías kai tēthas'.

¹⁸ Callanan, *Sprachbeschreibung bei Aristophanes*, 90.

evolution from a core, etymological meaning, but he considers also various extensional meanings, usually the results of metonymical or metaphorical processes. Also remarkable is his interest in diatopic and diaphasic variation. So, for example, he registers the different meanings of the term *steganomion* in the Attic writers and in the usage of his contemporaries, without any hint of censure: ‘the *steganomion* . . . in our usage is the place where [people] take their meals, but in Attic usage it signifies the rent paid for lodgings’.¹⁹ Along similar lines, he writes ‘many believe the word *epistatēs* to be used only for “a stand for a pot”, as well as “beggar” as in Homer. But they ignore that this is also the word for “trainer”; I am not completely sure general usage is demonstrable for other sorts of teachers.’²⁰ Aristophanes’ inheritance must have been very wide-reaching for the traditions of ancient linguistics.

In Rome, the influence of Hellenistic scholarship made itself felt from the second century BC onwards, when successive waves of Greek scholars moved to Rome and held lectures.²¹ The most relevant for grammatical and lexicographical study must have been Philoxenos, a grammarian who lived between the end of the republic and the age of Augustus (who reigned 31 BC–AD 14), and was probably a contemporary of the great Roman scholar Varro Reatinus. He was greatly interested in the Latin language, which he maintained to be derived from Greek via the Aeolic dialect; he was also interested in word derivation and etymology, all pursuits very close to the core interests of lexicographers.²²

Archaic Latin texts were also used in teaching, at all levels from elementary to advanced, and the need arose to use study tools for the explanation of obsolete or difficult vocabulary, on the same lines as collections of glosses were used for Homer and the other poets in Greece.²³ The earliest extant lexicon in Rome is Verrius Flaccus’ *De significatione uerborum*, a large compilation spanning perhaps sixty books. Verrius, however, cites several earlier

¹⁹ Aristophanes, *Fragmenta* (1986), item 7, ‘to *steganomion*, ho dêloi, phēsi, par’hēmin men ton topon en hōi estiōntai, par’Attikois de ton misthon tou pandochēiou’.

²⁰ Aristophanes, *Fragmenta* (1986), item 35, ‘*epistatēs* epi monou tou chutropodos dokei tois pollois tatesthai, kai tou metaitou, hōs par’Homērōi. agnoousi de hoti kai ho paidotribēs houtō kaleitai. to de kai epi tōn allōn kathēgētōn tattein ou phēmi pantōs tēn chrēsin deiknunai’.

²¹ Zetzel, *Latin Textual Criticism in Antiquity*, 10–26. Roman scholars also travelled east to meet eminent Greek intellectuals such as Dionysius Thrax: cf. Pfeiffer, *History of Classical Scholarship*, 266.

²² Philoxenos discussed contacts between the Latin and Greek languages in his ‘Peri tēs tōn Rōmaion dialektou’, preserved in his *Fragmenta* (1976), items 311–29.

²³ Cf. Deufert, *Textgeschichte und Rezeption der plautinischen Komödien*, 117–20.

scholars as linguistic authorities: among them is one Santra, author of *De antiquitate uerborum*, which was perhaps a lexicographic work.

Verrius lived at the end of the republic and was so famous as a teacher that Augustus invited him to be tutor to his own two nephews. Verrius' work is extant only in an epitome, by one Sextus Pompeius Festus; in turn, this later epitome was even further shortened by the great Lombard scholar Paul the Deacon.²⁴ Festus' epitome, however, is still extant in its original form, in just one badly preserved early medieval manuscript and its humanistic copies, which preserve only the letters from M onwards.²⁵ The *De uerborum significatione* concentrated mostly on antiquarian and obsolete vocabulary, using a vast array of literary sources to elucidate the possible meanings of a word. The lemmata were ordered in an alphabetical sequence: indeed, each surviving letter has two groups of alphabetical sequences (the first in ABC- and the second in AB- order), perhaps a remnant of different sources imperfectly amalgamated, or the reflection of two stages of Verrius' work. The meanings were illustrated listing several quotations from a canon of earlier authors, usually following one another in a fixed order – for example, first Accius (if the lemma occurs in Accius), then Afranius, and so on – as if Verrius had gone over a preordained list of lemmata many times, adding the appropriate citations with each new author he was reading.²⁶

The vocabulary included in Festus' epitome of Verrius is mainly religious and antiquarian, even if Festus declares, oddly, that he has omitted from his abridgement 'words long passed and obsolete' (*intermortua et sepulta uerba*), a definition which in some ways describes the entire scope of the dictionary. Verrius is generally very full in registering multiple and evolved meanings, as well as etymology, always a primary concern of ancient lexicography as a tool to uncover the 'true' meaning of a word. Here is an example of his etymological concerns:

monstrum, as Aelius Stilo has shown, comes from *moneo*, as if it were **monestrum*. Similar is the view of Sinius Capito, who interprets *monstrum*

²⁴ See R. A. Kaster, in Suetonius, *De grammaticis* (1995), 190–6, and Pieroni, *Marcus Verrius Flaccus* 'De significatu uerborum'.

²⁵ For a thorough re-examination of the manuscript tradition of the work and its title, see Di Marco, 'Per un riesame della tradizione umanistica di Festo'.

²⁶ This is the so-called *Lex Lindsay*, because the first to identify this procedure was W. M. Lindsay, in his study of the sources of a later Roman lexicographer, Nonius Marcellus; on the *Lex Lindsay*, see now Gatti's introduction to Nonius Marcellus, *De conpendiosa doctrina* (2014), vol. I. This practice was probably also followed by Greek authors of lexica, but more research needs to be done on this point.

because it ‘reveals’, *monstrat*, the future, and is a warning about the intentions of the gods.²⁷

The interesting detail here is the use of the ‘reconstructed’ form *monestrum* to represent in visible terms the evolution from the suggested etymology to the existing lexeme.

If we compare the first extant letter, M, with the wordlist of a modern dictionary, we can form an adequate perception of Verrius’ selection criteria. There are in Verrius fifteen entries between *manare* and *mansuetum*.²⁸ The corresponding interval in the modern *Thesaurus linguae Latinae* includes 282 lemmata. Verrius has ignored all high-frequency lexemes, such as *mando* ‘to send’, *maneo* ‘to stay’, and *mane* ‘morning’, and even culturally relevant entries such as *mancipium* ‘slave’. On the other hand, he has included many proper names, of deities (*Mamers*) or historical figures (*Mamiliorum*); other lemmata are important for the determination and understanding of ancient rituals (*manalem*, *manubiae*), folklore (*manias*), or simply religious language. Indeed, the aim of Verrius’ work was to illustrate the meanings of more recondite words, or the early meanings of other slightly commoner words. Also interesting is the entry *molucrum*, where we find again the presence of abundant source referencing, and a comparison with a Greek word of similar meaning:

molucrum is not only used for a kind of broom used for sweeping the millstones – the Greeks call it *mulēkoron* – but also for a swelling of the belly which affects girls even if not pregnant, which Afranius mentions in his *Virgo* . . . Cloatius also in his books about sacred rites: ‘They call *molucrum* a square wooden tool used for placing sacrifice victims’. Aelius, in his *Annotations on the Carmina Saliaria*, says that it is called with this name because it is placed under the millstone. Aurelius Opillus calls *molucrum* the place where the grain is ground.²⁹

²⁷ Pompeius Festus, *De verborum significatu* (1913), 122 line 6, ‘MONSTRVM, ut Aelius Stilo interpretatur, a monendo dictum est, velut monestrum. Item Sinius Capito, quod monstret futurum, et moneat voluntatem deorum.’

²⁸ They are *manare*, *manalem*, *manubiae*, *manias*, *manceps*, *manduci*, *mancina*, *Mamercus*, *Mamuri*, *Mamers*, *martialis*, *Mamiliorum*, *Mamilia*, *mamphur*, and *mansuetum*.

²⁹ Pompeius Festus, *De verborum significatu* (1913), 124, ‘MOLVCVRVM non solum quo molaē verruntur dicitur, id quod Graeci μὴλκωρον appellant, sed etiam tumor ventris, qui etiam virginibus incidere solet: cuius meminit Afranius in *Virgine* . . . Cloatius, etiam in libris sacrorum: “Molucrum esse aiunt. ligneum quoddam quadratum, ubi immolatur.” Idem Aelius in explanatione carminum Saliarium eodem nomine appellari ait, quod sub mola supponatur. Aurelius Opillus appellat ubi molatur.’

Greek is indeed often compared with Latin, perhaps in the wake of influential Greek lexicographers such as Philoxenos, who had written on the contacts between the Latin language and Aeolic dialect of Greek: hence the remark that *petoriturum* ‘kind of four-wheeled carriage’ is thought by some to come ‘from the Oscan language, since there the word *pitora* means “four”; others think it is from Greek, but from the Aeolic dialect’.³⁰ The Verrius–Festus dictionary was clearly considered authoritative even by Greek-speaking scholars interested in Latin: the later Latin–Greek dictionary going by the name of ‘Glossarium Philoxeni’, extant in one manuscript of Carolingian date, but certainly late antique, quotes *Pompēios*, that is Verrius–Festus, as its source several times.³¹

Roughly contemporary with Festus is the most extensive alphabetical author lexicon which has reached us, the ‘Homeric Lexicon’ of Apollonius Sophista, a scholar not otherwise identified, but active certainly in the second century AD.³² Only one medieval manuscript is extant, but the text is also transmitted by several papyri, which points to early popularity. It is in ABC- or, for limited sections, AB- order. The lemmatization is sometimes disorderly, and most entries have a simple structure similar to that of traditional *scholia minora*, that is a one-to-one verbal correspondence. For some few verbal and nominal entries, some attention is given to polysemy, though only in the form of juxtaposition of different meanings, with an exemplification in one or two Homeric lines. Close in time to Apollonius is the Apion lexicon, also extant in one manuscript, with the Greek title of ‘*Ἀπιόνος γλῶσσαι Ὁμηρικοί*’. The relationship between the two is disputed, but they well represent, together and in conjunction with the evidence of papyri, a pool of Homeric scholarship which was a very important ground for the development of lexicography in antiquity. Homer, for example, forming a diverse corpus, gave ample opportunity to experience polysemy and meaning extension, which these glossaries carefully register, or simply juxtapose:

akeiomenon means ‘healing’ and ‘repairing’; with the meaning *healing* [cf.] ‘healing the wound’; with the meaning *repairing* [cf.] ‘mending the ships’.³³

³⁰ Pompeius Festus, *De verborum significatu* (1913), 228, ‘*PETORITVM . . . alii Osce, quod i quoque pitora quattuor vocent, alii Graece, sed aiolikōs dictum*’.

³¹ For example, ‘Glossarium Philoxeni’ in *Corpus glossariorum Latinorum*, 2.8, line 21, ‘*ador nikē hōs Pompēios*’ (‘*ador* [glory] victory, as Festus says’). Cf. Grafton, *Joseph Scaliger*, 1.151; Dionisotti, ‘Greek grammars and dictionaries’, 6–10.

³² See Haslam, ‘Homer lexicon of Apollonius Sophista’.

³³ Apollonius Sophista, *Lexicon Homericum* (1833), 19 line 10, ‘*akeiomenon iōmenon kai episkeuazonta. epi men tou iōmenon “helkos akeiomenon”, epi de tou episkeuazonta “nēas akeiomenon”*’.

agoreuein properly means 'to speak in the assembly'; by extension, however, it simply means 'to speak'.³⁴

The beginning of the common era is also the time at which papyrus documents containing bilingual lexica start to appear.³⁵ These documents are sometimes marginalized in studies of Greek and Roman lexicography, on account of their practical nature, but this is a misconceived decision, influenced by the inferior status the ancients themselves assigned to these books and their authors. Indeed, lexicographers in the Greek monolingual tradition never acknowledge the existence of bilingual dictionaries, out of a preoccupation with the purity of the language, and a reflection of the low profile enjoyed by the genre is the total anonymity of the authors: no scholars of note are on record producing bilingual lexica or glossaries, and we also have explicit evidence that teachers of foreign languages were regarded with some disdain and were poorly paid.³⁶

The importance of this class of lexica in the history of lexicography is, however, very great, and has been entirely passed over hitherto. Bilingual dictionaries were probably the first to include all registers of the languages systematically: not just literary quotations from canonical authors, but also everyday phrases taken from ordinary language, because learners cannot limit their acquaintance with the language they are learning to *glossemata*, difficult and rare vocabulary used by poets. When the two traditions of monolingual-antiquarian and bilingual-practical lexicography came together, which happened at the end of antiquity with the appearance of large alphabetical bilingual books, dictionaries became truly comprehensive inventories of one language or the other.

The earliest such documents have the structure of *onomastica*, lists of words arranged by semantic fields, but we also find fairly comprehensive alphabetized lexica from a relatively early date. Particularly interesting is *P. Sorb.* 2069, a document from the early third century AD.³⁷ The recto of the papyrus contains several columns of an alphabetic Latin–Greek dictionary (the letters *S-T-U*, which probably ended the roll). What is particularly interesting about this book is both the nature of the vocabulary chosen for

³⁴ Apollonius Sophista, *Lexicon Homericum* (1833), 4 line 12, '*agoreuein kuriōs men en ekklēsia legein, katachrestikōs de psilōs to legomenon*'.

³⁵ Collections of bilingual dictionaries are in Kramer, *Glossaria bilinguia*; Kramer, *Glossaria bilinguia altera*; and Scappaticcio, *Artes Grammaticae in fragmentis*.

³⁶ Seneca the Elder, *Controversiae* 9.3.14, '*numquam magnas mercedes accepisse eos qui hermeneumata docerent*' ('People who teach interpretation never get paid a lot').

³⁷ See Dickey and Ferri, 'New edition of the Latin–Greek glossary', for an edition, and Dickey, 'Creation of Latin teaching materials', esp. 189–90, for discussion.

lemmatization, and the conspicuous attention to polysemy and homonymy in the selections of the lemmata: the guiding principle in the choice of the lemmata seems to have been to highlight words which can be confused because they are homographs or near-homographs, for example *tibicen* ('*aulos* player') and *tubicen* ('trumpeter'), *uentus* ('wind') and *uentum* (past participle of *uenio*, 'I come'), and *uero* (ablative of *uerus* 'true') and *uero* 'indeed' (adverb). Here is an extract:

top *koruphe ilinx* top of the head *koruphe kephalēs* whirlpool of rivers *ilinx potamōn* the other forms are inflected as in [Latin] 'calx' *ta loipa hōs to asbestos* [but] the word is masculine *arsenikōs* the meaning is the same *to auto sēmainei* as with spelling 'uortex' *kai ouortex*.³⁸

The odd sequence, with short Latin phrases alternating with short Greek translations in the same line, is the result of adaptation from an original vertical layout in parallel columns into a space-saving continuous run. First comes the lemma, *uertex*, with its possible Greek translations, then come the exemplifications for both meanings, and finally some declensional information. Of particular interest here is the presence of a transliterated form, to take account of an archaic spelling in Latin, with root -o- instead of -e-, *uortex*. Information about grammar took account of apophonic changes in the course of the paradigm.

P. Sorb. 2069 also includes metalinguistic expressions signifying register: for example, the lemma *tonsa* 'shorn, shaved' is correctly explained with *kekar-mena* 'shorn', but a metaphorical meaning *kōpē* 'oar' is also registered, which is then explained as *glōssēmatikōs*, literally 'using an antiquated or obscure expression'. We will find similar expressions to indicate rarity and frequency in Hesychius, who uses *spaniōs* 'seldom' – an interesting link between the monolingual and bilingual traditions.

Interest in homographs and near-homographs (also called *paragogae* in Latin), and generally in the malapropistic and solecistic use of language, existed also in the monolingual tradition. It is seen in the work of Herennius Philo of Byblos, who compiled a collection of words with different meanings, the principle of which was to distinguish exactly the meaning of words sometimes used inappropriately in the common language. The work survives in the form of a later epitome, Ammonius' *Peri homoion kai diaphoron*

³⁸ Dickey and Ferri, 'New edition of the Latin–Greek glossary', col. 7, lines 115–19, '*uertex koruphe ilinx uertex capitis koruphe kephalēs uertex fluminum ilinx potamōn cetera ut calx ta loipa hōs to asbestos* masculine *arsenikōs* idem significat *to auto sēmainei* et *uortex* *kai ouortex*'.

lexeōn. In this example, we see that Herennius Philo included dialects in his lexicon:

pus, *pei*, *poi*, *pō* have different meanings in Doric; *pus* and *poi* indicate motion to a place, whereas *pei* is for being in, and *pō* for movement from ... therefore people speaking Doric and saying *pei poreuēi* ‘where [in what place] are you going?’ make a mistake, since the correct phrase is *pus poreuēi*.³⁹

Contemporary with *P. Sorb.* 2069 is *P. Oxy.* 1802, an alphabetical glossary (fragments of the letter M are preserved, but no other entries) also concentrating on non-Greek vocabulary. It is, however, of an entirely different nature, including dialectal Greek alongside Aramaic, Anatolian, and Iranian, presumably culled from antiquarian and historical sources. The entries contain no grammatical and inflectional information, and do not even identify clearly the exact language from which the lemmata are taken. In an interesting pair of lemmata, the word *mētrai* is lemmatized in two subsequent entries, correctly corresponding to two different words with different etymologies.⁴⁰ The first *mētrai* means ‘a species of bees’, and Aristotle’s *On the parts of animals* is given as the source. The second *mētrai* has this explanation: ‘in Tarsus and in Soli the writing tablets on which they register houses are called *mētrai*’. Aristotle, *Constitution of Soli*, is adduced as a source. This second *mētrai* is clearly non-Greek, perhaps Cilician, but the main interest lies with the bureaucratic meaning of the word, treated as if it were naturalized in the Greek language.

Study of the citational habits of Roman authors such as Verrius and Nonius has suggested that lexicographers started by compiling alphabetical lists of lemmata, perhaps based on a corpus of canonical literary texts. Two such wordlists may indeed have survived. *P. Oxy.* 3660 is an alphabetical list of words culled from unidentified sources, mostly historical and perhaps rhetorical, and has been interpreted as the preparatory stage of a major Latin lexicon, possibly even bilingual.⁴¹ The purpose of *P. Hibeh* 172, described above, may have been similar. Authors of bilingual dictionaries pillaged monolingual Greek dictionaries, which provided ready-made lists, even if

³⁹ Ammonius, *Liber de adfinium vocabulorum differentia* (1966), 423, ‘*pus kai pei kai poi kai pō* diapherei para tois Dōrieusi. to men gar *pus* kai *poi* tēn eis topon sēmasian dēloi· to de *pei* tēn en topōi· to de *pō* tēn ek topou ... hōsth’ hoi Dōrizontes kai legontes “*pei poreuēi*” hamartanousi· deon gar “*pus poreuēi*”’.

⁴⁰ *P. Oxy.* 1802, fragment III.iii, lines 4–7, edited in Schironi, *From Alexandria to Babylon*, 60.

⁴¹ First published by H. M. Cockle in *Oxyrhynchus Papyri*, LII, item 3660; reprinted in Scappaticcio, *Artes Grammaticae in fragmentis*, 462–3.

the literary and highfalutin vocabulary was not perhaps the first port of call for someone who needed to communicate effectively in daily routines. Other useful tools were the bilingual translations of school authors, such as Vergil and Cicero, which had the same bibliological format of the lexica, in two parallel columns. In 1979, Herwig Maehler published a papyrus of a bilingual translation of Vergil, in which the expression *mare ueliuolum* 'a sea furrowed by sails' was rendered by a Greek compound *armenopetes* – unparalleled elsewhere in Greek literature and visibly an ad hoc calque of the Latin created by the translator, to represent best the meaning of a rare Latin literary adjective, icastically catching the sea 'swept by sails'.⁴² However, the Greek compound appears also in a later Greek–Latin dictionary, the 'Glossarium Cyrilli', thus proving the point that bilingual translation was a source of bilingual lexicography.⁴³

In the Roman imperial period, the majority of extant Greek lexicographical works were driven by an Atticist bias: that is, they aimed at spreading and preserving knowledge of the 'Attic' variant of Greek, against contemporary incorrect or simply unrefined usage (*koinē homilia*, *sunētheia*, *chrēsis*). This purist movement seems to have become important in the Augustan period: one of the earliest writers in this line was the Alexandrian Irenaeus, also known by his Roman name Minucius Pacatus.⁴⁴

Indeed, Atticism is a major factor in the vast increase in the number of lexicographical papyrological finds from the second and third centuries AD.⁴⁵ This must correspond to a greater popularity of the genre, and a market demand for linguistic purity-correctness, in the face of a rapidly evolving language situation. Most lexicography from this period is preserved only in fragments cited by Byzantine authors, or in epitomes.⁴⁶ For many of these works several subsequent editions are mentioned. Among the earliest writers in this vein are Aelius Dionysius of Halicarnassus, and Pausanias. They were not Attic nationals, and in fact Aelius made a point of emphasizing the Attic qualifications of his native city, in Caria, in fact a colony of mixed Dorian and Attic ancestry.

Phrynichus, called the Arab, was the author of at least two important collections, extant only in later abridgements, the *Ecloga* and the *Praeparatio*

⁴² Maehler, 'Zweisprachiger Aeneis-Codex'; the papyrus is *P. Berolin. inv. 21138*; see also Fressura, *Vergilius Latinograecus*.

⁴³ *Corpus glossariorum Latinorum*, 2.245, line 7.

⁴⁴ Schmid, *Atticismus*, 1.204–10; see also Erbse, *Untersuchungen zu den attizistischen Lexika*.

⁴⁵ Esposito, 'Fragments of Greek lexicography', 258.

⁴⁶ For the period in general, see Matthaïos, 'Greek scholarship in the imperial era and late antiquity'.

sophistica, the latter probably the more important of the two, interesting also because it was a phraseology book, based on an Attic canon. The preface of the *Eklogē* (*Eklogē Attikōn rhēmātōn kai onomatōn*) ends with the warning ‘whoever wants to use language according to the approved and ancient manner [*archaiōs kai dokimōs*] should keep in mind the following’.

Phrynichus in fact championed a very restrictive canon of approved authors, limiting the range to little more than Aristophanes, Plato, Demosthenes, Thucydides, Xenophon, and the three great tragic poets. An interesting feature of this work was the censure of Menander, accused of polluting his native Attic idiom with unacceptable vocabulary, in this case for using the variant *kataphagas* ‘gluttonous’ in the place of the correct form *katōphagas* used by the dramatist Aristophanes – a passage showing how fine-grained the distinction between pure and debased Attic usage sometimes was.⁴⁷ A typical specimen is ‘*ergodotēs* [contractor] does not occur, but the verb *ergodotein* does in the more recent comedians: they are not, however, an authoritative source’.⁴⁸ Phrynichus contrasts the usage of the ancients or Attic writers with that of the uneducated, or the majority:

exormenizein [means] to put out flowers, which the majority say *ekballein*. Indeed, *ormena* [‘sprouts’] is the Attic word for the flowers of vegetables. But the majority and the ignorant call them *asparagoi*.⁴⁹

To an attentive reader, however, his analysis of vocabulary is much more nuanced than simply censure and prescription, and is not deaf to differences between the poetical and the standard, everyday register: ‘The poet says *ereugesthai*, “and he vomited while drunk”, but the ordinary person will say *erugganō*.’⁵⁰

The *Praeparatio* is perhaps the earliest example of an Atticistic syntactic lexicon, focused on the government of verbs and more generally on idiomatic verbal and nominal constructions such as the *schema Atticum* (agreement of neutral plurals and third-person singular verbal predicates). An unascrbed

⁴⁷ Phrynichus, *Ekloge* (1974), 400; see Sonnino, ‘Frammenti della commedia greca’, 166–8.

⁴⁸ Phrynichus, *Ekloge* (1974), 322, ‘*ergodotēs* ou keitai, to de *ergodotein* para tini tōn neōterōn kōmōidōn, hois kai autois ou peisteon’.

⁴⁹ Phrynichus, *Praeparatio sophistica* (1911), 67–8, ‘*exormenizein*: to exanthein, hoper hoi polloi ekballein legousin. *ormena* gar kaleitai hupo tōn Attikōn ta tōn lachanōn exanthēmata. hoi de polloi kai amatheis tauta asparagous kalousin’.

⁵⁰ Phrynichus, *Ekloge* (1974), 42, ‘*Ereugesthai* ho poiētēs: “ho d’ereugeto oinobareion”, all’ ho politikos *erugganein* legetō; the word *politikos* is one of the less common definitions for something approaching a label for ‘ordinary register’ in an ancient Greek native speaker’s text. On the issues of standard and register in Greek and Latin generally, see Clackson, *Language and Society in the Greek and Roman Worlds*.

descendant of this tradition is known from the adaptation made by Priscian of Caesarea, the most famous of Roman grammatical writers, who taught in Constantinople at the beginning of the sixth century. Priscian used this lexicon for the final, eighteenth book of his *Ars*, adding similar Latin phrases for each Greek entry: it is the single source of numerous Attic fragments otherwise unknown.⁵¹

Some titles of lost lexicographical works contain references to ‘usage’, such as that of Telephus of Pergamon’s *Peri chrēseōs ētoi onomatōn esthētos kai tōn allōn ois chrōmetha* (‘On the usage, or names, of clothing and other objects we make use of’). However, no fragments allow us to check to what extent the spoken language was an object of direct study rather than of censure. Interest in more colloquial and realistic language was also promoted by Atticism, because comedy – especially that of the dramatist Aristophanes – was thought to embody the true spirit of the Attic language; indeed, comic texts transmitted a large number of words for cloth items, footwear, and kitchenware, as well as proverbial expressions and some obscenities. Presumably a large part of this vocabulary was beginning to become confused or obsolete in the Hellenistic world. There were also specialized lexica concentrating on technical domains, such as anatomy (in a work by Rufus of Ephesus), political institutions, games, swear words and curses, and theatrical masks.

The Roman biographer and encyclopedic writer Suetonius Tranquillus, among his other works, wrote lexicographical tracts in Greek which provide an interesting example of cross-cultural erudition.⁵² From his *Peri blasphemōn kai pothen hekastē* (‘On swear words and their sources’), which Suetonius declares he culled from Homer, comedy, the orators, and other literary sources, see the following extract: ‘*Lupa*: is what Italians call a she-wolf. The word is also used ironically in reference to prostitutes, because they are bent on prey like those animals.’⁵³ Suetonius here may have drawn on his native language, although writing in Greek he adopted a detached stance, avoiding the glottonym ‘Latin’ (which the Greeks tended to call ‘the language

⁵¹ The syntactic Atticistic lexicon is well analysed in the many contributions collected in Martorelli, *Greco antico nell’Occidente carolingio*.

⁵² Wallace-Hadrill, *Suetonius*, 44; Suetonius, *Peri blasphemōn*; *Peri paidiōn* (1967) is an edition. Cf. *Suda* (1928–38), entry tau 895, ‘Trangkullos, ho Seuētonios chrēmatisas, grammatikos Rōmaios. egrapse ... peri dusphēmōn lexeōn ētoi blasphemōn, kai pothen hekastē’.

⁵³ Suetonius, *Peri blasphemōn*; *Peri paidiōn* (1967), 50, ‘*Loupa*· legetai de houtōs para Italiōtais hē lukaina· ho dē onoma metēktai asteiōs eis hetairidōn prosēgorian· to zōion te gar harpaktikon, hē loupa eitoun lukaina, kai hai hetairides de homoiotroi’.

of the Romans'), but he is not untrue to his statement to have drawn from literary sources, as the Latinism is often discussed by Greek historians of various periods.

Pollux is in many ways one of the most interesting examples of a lexicographer at work for the early period.⁵⁴ What we have is an abridgement of a longer work, an epitome which replaced the original work at least as early as the ninth century, in ten books, and not extant in its original form in any evidence concerning Pollux, no matter how ancient – the predestined fate of such a long work. Every one of the ten books has a preface addressing the imperial heir-designate Commodus Caesar, and briefly outlining the contents of each book. The likely date of composition is assumed to have been between AD 166 and 176, the date at which Commodus assumed the title of Augustus. We have some information about Pollux from a close source, the sophist Philostratus of Athens, who lived in the first half of the third century, in his *Lives of the Sophists*. Philostratus offers a qualified appreciation of his talents as a public speaker, calling him an average Atticist, but remarking that if we consider only 'his studies in words [*ta onomata*, literally "his nouns"] it seems that his tongue had been well trained in the Attic dialect'.⁵⁵ The title *Onomasticon*, Pollux explains, alludes to the presence of lists of words considered as de facto synonyms because speakers can treat them as basically interchangeable when delivering a set piece on, for example, the praises of a benefactor, a king, or a city council.⁵⁶ The work is in fact more than a tool for rhetorical declamation, and it is rightly considered an important step in the history of ancient lexicography.

A table of contents was placed at the beginning of each book in some manuscripts, at a very early stage of the transmission of the work; it may even have been authorial, at least in part.⁵⁷ It is, taken together, a very useful blueprint for gauging the reception and understanding of Pollux's work at a very early stage. Ancient or Byzantine, the language of the index has points of similarity with the paratexts of several other bilingual vocabularies.

⁵⁴ There is very little modern secondary literature on Pollux: for a general presentation, see Tosi, 'Polluce', and for a very good discussion of Pollux's epistles prefaced to all books and the in-fighting in the Atticist camp they sometimes presuppose, cf. Radici Colace, 'Polluce nell'Onomasticon'.

⁵⁵ Philostratus, *Lives of the Sophists* (1921), 239 (= 2.12).

⁵⁶ *Suda* (1928–38), entry πi 1951, describes the work as 'sunagōgē tōn diaphorōs kata tou autou legomenōn' ('a collection of different words used for the same thing').

⁵⁷ On the topic of indexes and tables of contents in ancient works, cf. Schröder, *Titel und Text*, 93–115.

Following the ancient Index, we find that Book 1 includes, in this order, ‘the names and appellatives used in reference to gods and other divine entities, the honours paid to them, their holy sites and peculiar lands, their shrines, and the seats of artists [i.e., makers of divine images] and the gods’ servants’; then come sections on ‘divination, worship and piety, the gods who live above the skies, the appropriate behaviour when approaching a shrine’, and so on.⁵⁸ The Index is in fact more an index of subjects, selective, and somewhat arbitrary in inclusion, and varied in the form of the entries. With paragraph 40 starts a section identified as *basilika onomata*, ‘words used for kings’, which highlights an important aspect of Pollux’s work, the composition of an inventory of promptings for rhetorical epideictic speeches, as shown for example by this passage: ‘when praising a king call him *father, meek, benevolent, blessed with foresight, easy to approach, humane, generous . . .*’. The structure of individual books is not always very carefully planned: this eulogy of kings is followed by a section on words describing ‘those who are not slow’ (which could still be said in praise of kings) and their opposites; then comes a section on dyeing textiles, purple, and traders; then on meteorology, the seasons, and expressions of time; then the house, ship-making and sailing, weaponry, chariots, agriculture, and hippiatrics. Of the other books, of particular interest are the second, which concentrates on human anatomy; the third on affinity, degrees of kinship (one is reminded here of Aristophanes of Byzantium), and social relationships; the sixth on food and meals; and the tenth on utensils, tools, and vessels.

Pollux seems to have made an effort at making continuous reading less demanding by inserting aetiological narratives for the origins of a custom or a word, to make the longer lists of nouns, adjectives, verbs, and adverbs less dreary. He usually develops a topic by expanding the root term with its derivatives, nominal, verbal and adverbial, simple and composita, and other etymological or pseudo-etymological cognates. In book 6, for example, the first core lemma is *polis*, ‘the city’, and immediately thereafter Pollux goes on: ‘if we wish to consider the city and its parts, we will say in this way: *oikistēs, oikizōn, polizōn*, not however *polistēs, ktizōn* and *ktistēs*’. All these words mean ‘founder’, but some were approved, because they had prestigious precedents in Attic prose or poetry, some less so. Some attention is also paid to extended and metaphorical usage and to idioms, as in 6.12: ‘from the word *agroī* [“fields”] come the Attic phrases *agronde elthein* [“to go upcountry”], and *agrothi oikein* [“to live in the country”], and *agrothen elthein* [“to come from

⁵⁸ Pollux, *Onomasticon* (1900–37), 2.249.

the countryside"] . . . *agroikos* is the boor, and he who lives in the country; *agrios* has both meanings, though preferably the boor'.

There is no sustained contrasting of Attic and later usage, as identified in other Atticist sources by the antinomic pair of *Attikoi* and *Hellēnes*. The prevailing opposition is between past and present usage, and it is not clear if by the latter Pollux means 'present' in Attica or in the Greek-speaking world. Pollux also attempts some distinction of register, when he does not approve the choice of a word, or when one word is more polite than another. For example at 10.45 he describes the words *diphros* 'litter' and *diphriskos* 'little cart' as polite alternatives (*euphēmoteron*) to *lasanon* 'chamberpot'. Another interesting example is at 3.18, where he condemns the word *pappepipappos*: 'for I don't approve of the word, even if some have used it, for the third grandfather: it is indeed terribly vulgar [*idiōtikon*]'. Sometimes, he prefers to keep an open mind, as in 10.25 when, listing various literary words for the collective 'furniture' which he prefers, he also relates that the prevailing usage is now *endomenia*, which he does not approve (*ouk epainō*), but he also adduces the evidence of 'a document listing winners at the Olympic games, and the prizes they received, where the word *endomenia* is used'. Another fascinating feature of Pollux is his attention to the evolution of word meanings, which we can exemplify at the change of meaning of the word *parasitos*, from 'attendant of the main performer of a sacrifice' to 'parasite' (6.35).

Pollux's lexicon is not a wordlist arranged in the vertical format as in other ancient dictionaries, with lemma and explanation succeeding each other every line or every second line, and the visually typical characteristics of indenting when a lemma extends over several lines of the column, to highlight the lemma. It is, as I have said, a more discursive text, in which the writer also sets out doubts, personal discoveries, and other people's opinions. However, the text as we have it may have been visually different in its day: many unconstrued lists may have been laid out as vertical lists at some early period, alternating with more discursive sections written continuously.

A particularly interesting book is the tenth, devoted to utensils and vessels in daily use. In the prefatory epistle to the book, Pollux clarifies his motives for putting together such a book, with a polemical strain: previous literature on the subject (he mentions Eratosthenes, author of *On the names of tools*) is incomplete, or inadequate in dealing with this area of the dictionary. Also of interest is the claim that present and contemporary usage is used as a basis for linguistic assessment, and placed on the same level as the authority of the fifth–fourth-century Attic writers:

And I think that, when you can see it for yourself, the present book will appear to you preferable to all as to usefulness. For, even if none of the others is without its uses, this does indeed dwell also on the most usual and current vocabulary. This is indeed the reason why I have adduced a great number of linguistic authorities, for the greatest part of the words was in need of a backup or even courage. And don't be surprised if we have included words we use even today [*ei de tina tōn nun eirēmenōn kan toutōi gegraptai, mē panu thaumasēis*]: as I collected the greatest number of words for vessels, I thought it was necessary to draw not only from the ancients but also from our own usage.

So far we have encountered author glossaries (mostly on Homer, but also on genres such as comedy or lyric poetry, or the orators, or the medical writers), and specialized lexica, such as lexica of Attic usage (vocabulary or syntax of the Attic writers), the lexica of synonyms and the lexica of quasi-synonyms, and finally *onomastica*. Between the first and second centuries AD, a new class of lexical works appears, that of the universal lexica, in which all the previous subgenres are fused together. These are huge compilations whose first intention was probably to unify the specialized lexica and glossaries of earlier authors, rather than to provide a comprehensive inventory of the Greek language. As said earlier, the idea that lexica should include all words of a language came about only gradually, and in my view mainly through the influence of bilingual Latin–Greek dictionaries, which were becoming more and more comprehensive as the centuries went by.

The first to write one of these universal (monolingual) books was Pamphilus of Alexandria, in the first century AD, whose mammoth work is reported by the *Suda* to have gone by the title of *Peri glōssōn ētoi lexeōn* ('Glosses and words'), in ninety-five books.⁵⁹ Pamphilus seems to have used both an alphabetical ordering principle, perhaps for the part on literary glosses, and the classified order used in *Onomastica*, in a second part. Over the course of the following century, one Iulius Vestinus epitomized the lexicon.⁶⁰ His epitome was again reduced by Diogenianus, in a work by the title of *Periergopenētes*, that is the 'Curious paupers', which seems to designate 'those who are without means but keen on learning'. The main reason why this list of works is provided here is that the compiler of our single most extensive universal lexicon of the pre-Byzantine period, Hesychius (for whom see also Chapter 12), quotes Diogenianus as his main source, in a preface which is so illuminating and at last lends a personal voice to all this chain of lost shadows that it is worth dwelling on, in some detail.

⁵⁹ *Suda* (1928–38), entry pi 142. ⁶⁰ *Suda* (1928–38), entry omicron 835.

The title preserved in the single manuscript transmitting the work, Venice, *Marcianus graecus* 622, may be spurious, but highlights interestingly only a part of Hesychius' heritage: 'Hēsychiou grammatikou Alexandreōs sunagōgē paṣōn lexeōn kata stoicheion ek tōn Aristarchou kai Apīōnos kai Hēliodorou' ('Collection by the grammarian Hesychius of Alexandria of all the words, in alphabetical order, from [the glossaries of] Aristarchus, Apion, and Heliodorus'), in which the dictionary is situated at the end of a line of Homeric glossography, certainly a recognizable thread, but not the most interesting in the dictionary.

In the prefatory matter, addressing an unknown Eulogius, Hesychius stresses as his principal merit that of completeness of the collected materials: others before him have collected glosses of individual authors, Homer, or the lyric poets and the tragic poets, but none all of them together. Only Diogenianus collected them all regardless of generic distinctions, and added the rhetorical writers; he left out, however, writers of medicine and historians. Another merit of Diogenianus was the alphabetical ordering of all the lemmata in ABC- or ABCD- order – a great help for his readers – and the inclusion of *paroimiai*, literally 'proverbs', but in fact also idiomatic expressions generally. In short, Diogenianus offered a very useful educative tool for all interested in the best use of language. The only drawback of Diogenianus, in Hesychius' view, is the failure to provide such idiomatic or proverbial expressions with a commentary, whereas for the other lemmata there is no indication of source (author and work title where the word is used). Finally Diogenianus did not explain words with multiple meanings, or explained them only cursorily, whereas it would have been necessary to illustrate each different meaning of the lemma with appropriate passages. In the last part of the preface, Hesychius boasts that he has included all words found in his predecessors, and has added a great number of new ones he could not find in these works; he does not specify where he took these new words from. Hesychius also sets great store by orthographical correctness, according to the precepts of Herodian, choosing, among homonyms, the spelling appropriate for the meaning required.⁶¹ He boasts that he has included full information about the texts referred to in the entry or subentry, even when a particular word was little or rarely used, and that he has omitted only those words which were clear and unnecessary.

⁶¹ On Herodian, a second-century Alexandrian grammarian who wrote, among other works, a *Peri katholikēs prosōdias* ('The rules of accentuation'), see Probert, *Ancient Greek Accentuation*, 23–6.

This preface contains many possible suggestions for querying Hesychius' text even from a modern standpoint. I have already underlined the stress placed on completeness, which is, however, more a matter of the combination of all the written sources rather than the idea that a lexicon should give a complete picture of the language repertoire. Also interesting is the reference to frequency and rarity of usage, even if Hesychius makes it only to boast that he has quoted authors and titles for rare words and meanings too, and the candid statement about the omission of 'words which were clear'. In lemmata for which more than one meaning is provided, commoner meanings are sometimes labelled *to sunēthes legomenon* 'the usual sense', with no further attempt at a definition: the absence of explanation for commoner meanings brings us back to the origins of Greek lexicography, the explanation of difficult *glossemata*. When he speaks of 'rarity', it is also uncertain if Hesychius refers to the rarity of a specific lexeme or to the rarity of meaning assumed within the range associated with an entry.

True to his initial claims, Hesychius pays great attention to polysemy. The polysemic entries are generally complex and ambitious, as in the following examples:

agkon [elbow]: has the usual meaning, and the bend of a wall: 'scaled an angle of the lofty wall' [*Iliad*, 16.702]; *agkones* is also the name of the parts holding the arms of the lyre.⁶²

salpinx [trumpet] . . . instead of *kēru*x ['herald']; some think it may denote a bird of some kind, a war machine, and the sea trumpet, but in Archilochos it is the *strombos* [a kind of shell]. They also interpret it as the temple of Athena Salpinx at Argos.⁶³

Finally, a fascinating feature of Hesychius is the way in which he discusses Latin words (in fact, only 'the Romans' are mentioned, not their language) in several entries. Greek and Latin were profoundly intertwined at this time, but most lexicographers adhered to strict canons of linguistic purism and all but pretended to ignore 'barbaric' languages such as Latin. An interesting example is the statement that 'the Romans call heavens *kailus* [*caelus*]'.⁶⁴ This may strike us at first as a rare intrusion of the spoken language (in Vulgar Latin the

⁶² Hesychius, *Lexicon*, vol. I (2018), entry alpha 485, 'angkōn· to te sunēthōs legomenon, kai hē tou teichous kampē· "ep' ankōnos bē teicheos hupsēloio" kai tēs kitharas de ta anechonta tous pēcheis ankōnes legontai'.

⁶³ Hesychius, *Lexicon*, vol. III (2005), entry sigma 125, 'salpinx· . . . anti tou kēru;x; tines de ornin poion kai organon polemikon kai thalassian salpinga, para Archilochōi de ton strombon. ekdechontai de kai Salpinggos Athēnas hieron para Argeiois'.

⁶⁴ Hesychius, *Lexicon*, vol. II (1966), entry kappa 228, 'kailous ouranos Rōmaioi'.

masculine was replacing the neuter, and writers complain of the substitution of *caelus* for *caelum*). More likely, however, it is a piece of inherited Roman erudition (*Caelus* is the Roman god corresponding to *Uranus* in Greek), perhaps from a bilingual *onomasticon* where attempts at co-ordination between the two competing sets of pagan gods were very common. Hesychius also registers other Latinisms in common usage at his date, as when instead of the Greek word *kuamos* 'bean, pulse' he uses *phaba*, a blatant Latinism which must have entered Greek at an early date in the imperial period, in an entry '*phaba*: a great fear [?] and the common pulse'.⁶⁵ There is no need, however, to exaggerate the personal contribution of the lexicographer and his observation of current language, and he may have drawn these words from technical, perhaps specialized medical and botanical sources: plants were often identified by the names given to them in various languages, to give a more precise identification of a species at a time when no scientific taxonomy existed.

⁶⁵ Hesychius, *Lexicon*, vol. IV (2009), entry phi 6, '*phaba*: megas phobos. kai to sunēthes osprion'.

PART II



THE PRE-MODERN WORLD

China, c. 600–c. 1700

NATHAN VEDAL

The story of Chinese lexicography, and indeed of China itself, from the years 600 to 1700 begins and ends in a period of unification. The short-lived Sui dynasty (605–18) unified the Chinese empire after centuries of division. In the wake of a new imperially sponsored examination system came dictionaries aimed at creating a unified standard for exam usage. By the early eighteenth century, the Manchu Qīng dynasty had eliminated most of the vestiges of rebellion from the previous dynasty and was set on expanding its territory. Part of its imperial project was linguistic, as the new multilingual empire staked its claim as an authority on Chinese, among other languages. During this period of more than a thousand years, states and dynasties came and went, and the territory claimed as China fluctuated drastically. A cultural identity, however, came to be maintained, in large part on the basis of a textual tradition. The centrality of texts and language for participating in this culture is reflected in turn by the important place lexicography occupied in the world of Chinese scholarship.

This chapter will focus primarily on three types of lexicographic activity: rhyme dictionaries; dictionaries of script form; and glossing dictionaries. Such a categorization existed in pre-modern China as a way of dividing up the broader field of philology (*xiǎoxué* 小學, ‘the lesser/elementary learning’). While there did exist scholars expert in all three categories, many specialized in only one, a fact observed by the sixteenth-century polymath Hú Yínglín 胡應麟.¹ This tripartite categorization has limitations. It may imply that these three endeavours were entirely disconnected, when in fact there was often a close relationship between them. However, as the scholars themselves framed their work primarily in relation to one of these three categories, it is useful for the purposes of introducing the main trends in Chinese lexicography of this period.

¹ Hu, ‘Huayang boyi’, 501.

By the sixteenth and seventeenth centuries, a tradition of writing in a form of vernacular Chinese was well established.² Nevertheless, lexicographic activity was focused for the most part on the literary language. Like Latin in the West (for which, in this period, see Chapter 14), this language, referred to here as Literary Chinese, was based on a written and spoken language of antiquity, and functioned as the linguistic currency of elite writing. This written language was largely monosyllabic, and dictionaries primarily explicated terms at the syllable level, represented by a single Chinese character, rather than the multisyllabic terms that characterized the spoken languages of this period. The Chinese syllable contains an ‘initial’ consonant (termed ‘zero-initial’ when the syllable begins without a consonant), ‘final’, and tone. The final always contains a main vowel, and is sometimes preceded by a medial segment or followed by a consonant coda, or both.

The period from 600 to 1700 encompasses multiple stages in the development of the Chinese language, the highly differentiated regional variations of which I will refer to as *topolects*.³ Despite possible distinctions, the language of the seventh through twelfth centuries can be broadly referred to as Middle Chinese (*zhōnggǔ Hànyǔ* 中古漢語). A notable set of phonological distinctions documented from around the fourteenth century is generally characterized as Early Mandarin, with those of the fifteenth through eighteenth centuries sometimes termed Middle Mandarin. Together they are considered Early Modern Chinese (*jīndài Hànyǔ* 近代漢語). Old Chinese (*shànggǔ Hànyǔ* 上古漢語) refers to the language of the eleventh century BC to the seventh century BC; as the language of the classics, it was of considerable concern to scholars in the period covered in this chapter.⁴

There was no single word for ‘dictionary’ in this period of Chinese history. Terms such as *zìshū* 字書 (‘character book’) and *yùnnshū* 韻書 (‘rhyme book’) were commonly adopted to refer to what we would consider dictionaries. As a general rule, *zìshū* referred to dictionaries organized according to the structure of characters, while *yùnnshū* were organized phonologically, but the two terms were often used interchangeably. At the end of the period covered in this chapter, there appeared an influential dictionary with the term *zìdiǎn* 字典 in the title. While this initial usage was to be understood literally as ‘the classic of characters’, this term would eventually be

² Idema, *Chinese Vernacular Fiction*; Mair, ‘Buddhism and the rise of the written vernacular’.

³ Mair, ‘What is a Chinese “dialect/topolect?”’

⁴ For more on the periodization of the Chinese language, see Norman, *Chinese*, 23, and Wang, *Hanyu shi gao*, 44–9.

adopted as the word for ‘dictionary’ in the modern Chinese language. The dictionaries discussed in this chapter were printed xylographically, unless otherwise noted.

Political Factors Influencing Dictionary Production

Texts of the Chinese classical tradition considered linguistic standardization as a task for righteous government. As an oft-quoted text in this tradition claimed, the sage kings of antiquity ensured that all their subjects ‘wrote in the same script’ (*shū tóngwén* 書同文).⁵ The imperial courts of many Chinese dynasties took linguistic matters to be of integral importance to their governance. This manifested itself primarily in two venues: imperially commissioned lexicographic works and requirements for the civil service examinations.

The early seventh-century Sui dynasty initiated a set of examinations, which could provide access to government service. Later in the seventh century, the Tang dynasty continued to standardize these exams. Although they would not serve as the primary means of achieving an official position for several centuries, lexicographic works geared towards these exams began to appear even at this early stage. For example, *Gānlù zìshū* 干祿字書 (‘Character book for seeking official emolument’) was a privately produced dictionary of script forms from the end of the seventh or the early eighth century. This text designated three categories of script form: ‘vulgar’ (*sú* 俗), ‘common-use’ (*tōng* 通), and ‘orthodox’ (*zhèng* 正). The last of these was intended to serve as a model for script form in civil service examination compositions.⁶ Through the eighth and ninth centuries, scholars continued to compile models of correct script forms expressly for the purpose of examination standards, such as *Wujīng wénzì* 五經文字 (‘Characters of the Five Classics’) and *Jiǔjīng zìyàng* 九經字樣 (‘Character models for the Nine Classics’). Another influential work in the examination field, known as *Lǐbù yùnlüè* 禮部韻略 (‘An overview of rhymes from the Ministry of Rites’), was compiled on imperial order in the eleventh century. This text established the acceptable rhyme scheme for poetry composition in the examinations. Although not from the Tang dynasty, it came to be seen in later periods as an important key to understanding the medieval rhyming system.

The imperial courts of many Chinese dynasties commissioned a dictionary for promulgation in the school system, as well as to represent a standard for the examinations. These works were typically compiled by a set of court

⁵ Zheng Xuan, et al. (eds.), *Liji zhushu*, j. 53, p. 13a. ⁶ Liu, *Ganlu zishu zilei yanjiu*.

scholars and printed at an imperial printing house. Some of them were influential for later generations. The eleventh-century *Guǎngyùn* 廣韻 ('Expanded rhymes'), for instance, was an expanded version of the seminal seventh-century rhyme dictionary *Qièyùn* 切韻 ('Divided rhymes'). As the earlier text was lost for most of the rest of Chinese history, the *Guǎngyùn* served as the point of access for understanding the medieval rhymes of the *Qièyùn* for most Chinese scholars into the twentieth century. The Kāngxī court of the early Qīng dynasty commissioned the *Kāngxī zìdiǎn* 康熙字典 ('Classic of characters of the Kāngxī emperor'), which was widely circulated and cited in scholarly works throughout the dynasty. Other imperial dictionaries, however, were barely noticed in their own time. For example, Míng dynasty scholars frequently made reference to contemporary neglect of the early Míng court *Hóngwǔ zhèngyùn* 洪武正韻 ('Correct rhymes of the Hongwu emperor').⁷

Despite claims for authority, lexicographic projects compiled at the court or explicitly linked to state enterprises ultimately comprised a small portion of the philological activity of this period. Many important innovations in dictionary production occurred within literary contexts, as the popularity of new genres necessitated guides for their composition. Lexicographical activity conducted by and in consultation with the Buddhist clergy exposed Chinese scholars to phonological methods developed in a Sanskrit context. Classicists, too, developed new lexicographic methods to bolster their interpretations of ancient texts.

Rhyme Dictionaries

One of the most influential works for the modern study of historical Chinese linguistics has been the *Qièyùn*, originally compiled in 601. This work was composed by Lù Fǎyán 陸法言 in consultation with several scholars, in order to set a standard for reading pronunciations of literary works. Considerable ink has been spilled in debates over what the linguistic basis of this standard was.⁸ Such debates are primarily a modern concern. In fact, the original text of the *Qièyùn* was lost early on, and a relatively early manuscript version was rediscovered only in the first half of the twentieth century. Numerous rhyme dictionaries have been unearthed in the caves of Dūnhuáng 敦煌, which

⁷ See, for instance, Lu, *Shuyuan zaji*, 123.

⁸ Chen, *Guangyun yanjiu*, 151–77; Baxter, *Handbook of Old Chinese Phonology*, 36–7.

were sealed in the eleventh century.⁹ While notable for the picture they give of lexicography in the medieval period (including early manuscript versions of the *Qièyùn*), they were not accessible until the past century and therefore did not greatly influence later lexicographers. As I have noted, the eleventh-century *Guǎngyùn*, itself based on the rhyme system of the *Qièyùn*, served as the closest access point to the original text for most of Chinese history. The *Qièyùn* inherited many of its most famous characteristics from an earlier tradition of phonological scholarship. Nevertheless, its standardization of these elements and conceptual influence on later scholarship have earned it a significant place in the history of Chinese rhyme dictionaries.

Most relevant to the present reader will be two characteristics of the *Qièyùn*: first, a four-part tonal division of Chinese characters and, secondly, the glossing of a single syllable (one Chinese character) by means of two characters. Neither of these methods was original to the *Qièyùn*. Their formalization, however, is generally attributed to this work. The text was organized according to the concept of *yùn* (韻 ‘rhyme’), meaning in this context the main vowel, consonant coda (if present), and tone. Within a particular rhyme category, homophonous characters sharing the same initial would be grouped together.

The four-part tonal division was originally codified in fifth-century literary circles for the purposes of poetry composition. The divisions are ‘level’ *píng* 平, ‘rising’ *shǎng* 上, ‘falling’ *qù* 去, and ‘entering’ *rù* 入. The first three are characterized by tonal contour, the nature of which may possibly be inferred from their names. The ‘entering’ tone is typified by its ending in a stop consonant, [p], [t], or [k] in Middle Chinese (Modern Mandarin does not contain an ‘entering’ tone). The overarching structure of the *Qièyùn* was governed by tones: composed of a total of five *juàn* (卷 ‘fascicles’), it included two *juàn* of level-tone characters, and one *juàn* each for the other three tones.

Fǎnqiè (反切, occasionally *fānqiè* 翻切) is a method employed throughout the *Qièyùn* for glossing the pronunciation of a syllable. Documented as early as the late second century in the context of classical commentary (there are examples in Chapter 3), *fǎnqiè* remained a common method of indicating pronunciation until the twentieth century. According to this method, the pronunciation of a single Chinese character, representing one syllable, is indicated by means of two characters. The first *fǎnqiè* character represents the initial of the glossed character, while the second character represents the

⁹ Bottéro, ‘*Qièyùn* manuscripts from Dūnhuáng’. See also works contained in Zhou, *Tang Wudai yunshu jicun*, and Pan, *Yingya Dunhuang yunji biele*.

final. For example, *tóng* 同 can be glossed in *fǎnqiè* as *tú* 徒 + *hóng* 紅 (this formula in Chinese would be written *tú hóng qiè* 徒紅切). While an important innovation in the history of Chinese phonology, *fǎnqiè* spellings can present ambiguities. The most significant issue in *fǎnqiè* was the lack of standardization. Differentiations of tone and the presence or absence of medial vowels in the final (-*iang*, for instance, as opposed to -*ang*) can be inferred, but are not explicit and are subject to speculation.¹⁰

Until a set of discoveries in the twentieth century, the content of the *Qièyùn* came down to us primarily through the *Guǎngyùn* ('Expanded rhymes'), a redaction made in the Song dynasty court under the editorship of Chén Péngnián 陳彭年 and printed in 1008. As its name implies, the *Guǎngyùn* was more than a reprint of the *Qièyùn*, and modified certain phonological categorizations, while expanding content. It has nevertheless been shown to reflect the content of the *Qièyùn* in broad strokes and is primarily notable for serving as the closest reflection of the *Qièyùn* system for most of Chinese history. The rhyme system of the *Qièyùn*/*Guǎngyùn* functioned as an authority for poetic composition in many genres, owing to the belief that it represented the phonology of the great medieval poets. It too experienced modifications at the hands of many later scholars, in search of simplifying its content. Hán Dào zhāo 韓道昭, for instance, simplified the 206 rhyme groups of the *Guǎngyùn* into 160 groups in his early thirteenth-century *Sìshēng piānhǎi* 四聲篇海 ('Sea of writing in the four tones'). *Lǐbù yùnlüè*, mentioned above as an examination standard, shared the 206 rhyme groups of the *Guǎngyùn*, but would also undergo a simplification in the thirteenth century, in this case to 106 rhyme groups. This simplification, which came to be known as the *píngshuǐ* 平水 rhymes, would form the basis for poetic education and rhyming technique for later generations.

In 1374, the founder of the Míng dynasty, Zhū Yuánzhāng 朱元璋, commissioned an imperial rhyme dictionary, *Hóngwǔ zhèngyùn*. This was framed as a revision of the revised version of *Lǐbù yùnlüè*. It further reduced the number of rhyme groups to seventy-six. Although it was intended to set a dynastic standard for literary composition, contemporary observers noted that little heed was paid to the dictionary outside certain official contexts. The rhyme system of *Hóngwǔ zhèngyùn* can be traced back to the *Guǎngyùn*. Later phonological dictionaries experimented with rhyme categories entirely removed from those of the *Guǎngyùn*. Increasingly in the sixteenth and seventeenth centuries in particular, new rhyming systems emerged based

¹⁰ Branner, 'Suí-Táng tradition of Fǎnqiè phonology'.

on contemporary vernaculars. Although dictionaries organized according to these new systems were still primarily concerned with the literary language, their phonological basis was no longer directly linked to the foundational medieval works of the tradition.¹¹

An important development in the history of phonological dictionaries occurred in the fourteenth century, again in a literary context. Composed in order to standardize rhymes for composition of arias in a style of northern drama, *Zhōngyuán yīnyùn* 中元音韻 ('Tones and rhymes of the central plains') presented a significantly different phonology from the *Qièyùn* tradition. This text contained only three tones, the level, rising, and falling. Its author, Zhōu Déqīng 周德清, eliminated the entering tone, which had largely disappeared from many northern dialects by this time, and distributed characters that had formerly been categorized as entering tone among the other three tones, based on his spoken language. He also divided the level-tone characters into two divisions, based on whether or not their initial was voiced.¹² His approach was fiercely debated in later centuries, with many opposing his reduction of the four-part rhyme scheme. However, the text was widely adopted in literary circles and inspired the production of numerous other dictionaries on its model for centuries to come, such as *Qiónglín yǎyùn* 瓊林雅韻 ('Elegant rhymes of the jade-white forest', 1398), *Zhōngzhōu quán'yùn* 中州全韻 ('Complete rhymes of the central states', probably seventeenth century), and *Yùnxué lízhū* 韻學驪珠 ('Precious pearls of rhyme study', printed 1792).

Another important area of phonological research in China was the study of 'ancient pronunciation' (古音 *gǔyīn*). The rhymes of antiquity, ascribed to texts from the first millennium BC, were important to later scholars primarily for two reasons: recitation of classical texts, which relied on subscribing to a theory of how their rhymes worked, and poetic composition on the model of ancient texts. Prior to the twelfth century, although scholars recognized that ancient poetry no longer rhymed when read in their contemporary language, attempts to deal with ancient rhymes were primarily undertaken ad hoc. For instance, the eighth–ninth-century poet Hán Yù 韓愈 was heralded among later poets for his use of ancient rhyme patterns in his compositions.¹³ However, his attempts must have been largely impressionistic, as there existed no systematic dictionary or phonological study of Old Chinese phonology from which to draw.

¹¹ Geng, *Ming Qing dengyunxue tonglun*, 140–5.

¹² Ning, *Zhongyuan yinyun biaoqao*; Stimson, *The Jongyuan in yunn*.

¹³ See appraisals in Li, 'Lidai xuezhe dui Han Yu shi zhi pingjia'.

The first dictionary to analyse ancient rhymes systematically was *Yùnbǔ* 韻補 ('Rhyme supplement') by Wú Yù 吳域. His contribution was to create a set of rhyme classes which he posited to exist in antiquity. Although he included some characters in more than one rhyme class, his system nevertheless reduced the range of options for adjusting the pronunciation of a word. Zhū Xī 朱熹, the great twelfth-century codifier of Neo-Confucian thought, was widely believed to have consulted this text for the pronunciation glosses in his own commentaries. This association served to bolster the authority of the text until the eighteenth century, when it became more widely known that this was likely not the case. The next significant work devoted to systematically discussing ancient rhymes appeared in 1606, when Chén Dì 陳第 published *Máoshī gǔyīn kǎo* 毛詩古音考 ('An analysis of the ancient sounds in the Mao poems'), a dictionary of ancient character pronunciations and corroborating evidence. Designed in part as a refutation of Wú's work, this text was one of the first to clearly recognize that the entire phonological system of Chinese antiquity could have been different from that of the writer's own time. As such, his work was highly influential on the increasingly sophisticated phonological studies of Old Chinese in the centuries to come. While Wú Yù envisioned a set of readings for each character to be adopted depending on the context, Chén Dì argued that each character in antiquity had only one pronunciation. Later scholars would build on this, creating rhyme groups to abstractly classify the sounds of Old Chinese.¹⁴

Alongside rhyme dictionaries organized by and dedicated to explicating the pronunciation of characters was another important genre of phonological description: rhyme tables (*děngyùn tú* 等韻圖). Rhyme tables were not dictionaries per se, but rather tools for describing the pronunciation of characters. They do not contain definitions and provide only one representative character for a particular syllable. They were often associated with dictionaries, however, being included as part of the lexicographical paratext, or separately as a form of phonological commentary on the content of particular dictionaries. The first extant rhyme tables come from the eleventh century, although the terminology associated with them appears to have been developed at the latest in the tenth. Rhyme tables in essence comprise a matrix of initials, tones, voicing, medial vowels, and rhymes. The format of rhyme tables was not standardized, although they generally shared several characteristics. The head of each column was typically an initial, while rows were organized according to elements of the final (including rhyme, medial

¹⁴ Baxter, *Handbook of Old Chinese Phonology*, 154–5.

vowel, and tone). In the medieval tradition, a set of thirty-six initials was adopted to encompass all existing Chinese initials; this system remained in use among many late imperial scholars. Similarly to *fǎnqiè*, syllables glossed in the body of a rhyme table represented the combination of an initial (at the head of its column) and a final. The tables provide more detailed description of both the final and initial than *fǎnqiè*, by virtue of the tables themselves being grouped according to specific phonological characteristics. Nevertheless, the phonological nature of certain technical descriptions remains debated.¹⁵

By the early seventeenth century, it was not uncommon for dictionaries to contain a rhyme table outlining the phonological system proposed by the compiler. For example, Wáng Yīngdiàn's 王應電 *Tóngwén bèikǎo* 同文備考 ('A complete examination of the shared script') concluded with a rhyme table titled *Shēngyùn huìtōng* 聲韻會通 ('Comprehensive collection of initials and rhymes'). Most famously, the *Kāngxī zìdiǎn* of 1716 contained a set of seventeenth-century rhyme tables. Such dictionaries continued to use *fǎnqiè* to gloss pronunciation within individual entries. Especially for dictionaries such as *Tóngwén bèikǎo* and *Kāngxī zìdiǎn*, which were not organized according to phonological categories, the inclusion of rhyme tables may have served to provide generalizable principles for the use of characters in literary composition.

There was great variation in lookup methods among dictionaries organized phonologically. A significant number followed the *Qièyùn*/*Guǎngyùn* model: characters were grouped according to tone, and further separated into distinct rhyme groups on the basis of shared finals. Within a particular rhyme group, those characters sharing the same initials (in other words, homophones) were also grouped together. Although this model remained popular, later dictionaries also employed alternative categorization methods. *Zhōngyuán yīnyùn*, for instance, grouped characters first according to a rhyme group and then by tone. In this way, *t'ūng* 通 and *tǔng* 懂 were classified as belonging to the same rhyme group (*tūng* 東), despite the fact that one was level-tone and the other rising-tone. By contrast, in the *Guǎngyùn*, *t'ūng* 通 would be assigned to a level-tone rhyme group (*tūng* 東), and *tǔng* 懂 to a rising-tone rhyme group (*tǔng* 董).¹⁶ Other methods were experimented with in later periods, such as the use of homophones

¹⁵ Branner, *Chinese Rime Tables*, 4–7.

¹⁶ Reconstructions for early Mandarin and Middle Chinese, which in this case map onto each other well, adapted from Ning, *Zhongyuan yinyun biao*, and Branner and Yi, *Yintong*.

rather than rhyme as the lookup unit in the seventeenth-century *Xiéshēng pǐnzì jiān* 諧聲品字箋 ('Notes on characters classified according to homophonous sound').¹⁷

Dictionaries of Script

Dictionaries of Chinese script took many forms in the period from 600 to 1700. Most were characterized by an organizational system based in some way on the shape of written characters. Some were intended to explain the nature of ancient scripts, which continued to be relevant in both artistic and classical studies. Others were prescriptive, indicating the correct form of characters and discrediting variant forms. Many hearkened back to the progenitor of Chinese lexicography, *Shuōwén jiězì* 說文解字 (see Chapter 3). Experimentation with organizational systems was particularly fruitful in this domain and gave birth to one of the major lookup methods still in use today in Chinese dictionaries.

Dictionaries of the script from the sixth to the eighth centuries focused primarily on establishing standard contemporary forms for use in official settings. Works such as *Gānlù zishū* and *Wǔjīng wénzì*, discussed above, were explicitly designed as resources for examination essays and other official compositions. Especially from the tenth century onwards, however, an increasing number of dictionaries came to focus on explicating ancient script forms. The uses of ancient scripts were varied. On the artistic side, archaic writing assumed an important place in calligraphic practice.¹⁸ For some moral thinkers, ancient characters reflected important ideological characteristics of antiquity. Certain classicists argued that versions of classical texts written in ancient script were more authentic than standardized versions. While such ideas were not held universally, this combination of artistic, scholarly, and political values instigated a newfound attention to the lexicography of old scripts. Even dictionaries that did not employ old script characters as headwords came to cite such forms for justification of the selection of a particular standardized form.

Shuōwén jiězì itself was reprinted in countless editions throughout this period. In addition, a tradition of commentarial dictionaries on *Shuōwén jiězì* flourished. Lǐ Yángbīng 李陽冰, an eighth-century calligrapher, prepared a new edition of *Shuōwén jiězì* to make available ancient script forms he felt

¹⁷ Söderblom Saarela, 'Shape and sound', 193–4.

¹⁸ For an in-depth study, see Bai, *Fu Shan's World*.

had been abandoned by his contemporaries. This edition was disparaged by later scholars on account of Lǐ's idiosyncratic editing processes. In the late tenth century, the court scholars Xú Kǎi 徐鉉 and Xú Xuàn 徐鉉 also produced important publications on the text. Xú Kǎi composed a work titled *Shuōwén jiězì xīzhuàn* 說文解字繫傳 ('Appended commentary on *Shuōwén jiězì*'), an in-depth commentary on the entries of *Shuōwén jiězì*, which also acted as a refutation of Lǐ Yángbīng's earlier edition. As the first great theoretical commentary on the nature of *Shuōwén jiězì*, Xú Kǎi's tenth-century text came to assume an important place in the tradition of Chinese lexicography. Combined with his brother's critical edition, it served as the entry point for discussions of *Shuōwén jiězì* for centuries to follow. A significant reordering of the text occurred in the twelfth century when Lǐ Tāo 李燾 compiled *Shuōwén jiězì wúyīn yùnpǔ* 說文解字五音韻譜 ('Rhymed arrangement of *Shuōwén jiězì* according to five tones'). Arranged according to the phonological characteristics of rhyme and tone, rather than classifier, this edition was also widely printed and was occasionally mistakenly attributed to the earlier Xú brothers.¹⁹

Commentaries on *Shuōwén jiězì* were justified by the high degree of ambiguity present in the original text. The terminology of the 'six ways of writing down [words]' (*liùshū* 六書) by which *Shuōwén jiězì* categorized characters was firmly established in the Chinese philological tradition. However, the actual meaning of certain terms was unclear and frequently debated.²⁰ Following the work of the Xú brothers, several other significant dictionaries were produced, which attempted to address the organizational and conceptual issues of *Shuōwén jiězì*. The first such text, *Liùshū gù* 六書故 ('Fundamentals of the six ways of writing down'), by the thirteenth-century scholar Dài Tóng 戴侗, organized characters into nine thematic categories: numbers, heaven, earth, man, animals, plants, crafts, miscellaneous, and indeterminate. Scholars such as Xú Kǎi had previously attempted to discern a pattern in the ordering of entries in *Shuōwén jiězì*. Dài Tóng created an explicit system of thematic organization by which to locate characters. Related contemporary dictionaries, such as *Liùshū tǒng* 六書統 ('Rules governing the six ways of writing down') and *Liùshū zhèng'è* 六書正譌 ('Correcting errors in the six ways of writing down'), as well as later works such as the late fourteenth-century *Liùshū běnyì* 六書本義 ('Original meaning of the six ways of writing down'), similarly attempted to revise the

¹⁹ Yongrong, Ji Yun, et al., *Qinding Siku quanshu zongmu*, j. 43, p. 7b–8a.

²⁰ See Qiu, *Chinese Writing*, 156.

classifications in *Shuōwén jiězì*. Some of these dictionaries focused on simplifying the script-based organization originally present in *Shuōwén jiězì*. Others, such as the late sixteenth-century *Liùshū zǒngyào* 六書總要 ('Comprehensive overview of the six ways of writing down'), followed Dài Tóng in devising thematic categories by which to class characters.

The commentarial tradition on *Shuōwén jiězì* in this period culminated in Zhào Yíguāng's 趙宦光 *Shuōwén chángjiān* 說文長箋 ('Extended notes on *Shuōwén*'). This text was not a direct commentary on *Shuōwén jiězì*. Instead, it was alternate parts dictionary and treatise, providing new analyses and categorization systems inspired by the terminology in *Shuōwén jiězì*. Zhào Yíguāng's reputation suffered at the hands of the prominent late seventeenth-century philologist Gù Yánwǔ 顧炎武.²¹ This text, however, remained influential for eighteenth-century lexicographers. In particular, Zhào's citation of variant character forms, informed by his reportedly comprehensive reading practices, served as an important resource for later scholars. *Kāngxī zìdiǎn*, for instance, frequently cited variant character forms contained in *Shuōwén chángjiān*, as well as its definitions and character analyses.

One of the hallmarks of dictionaries written within the tradition of *Shuōwén jiězì* was organization according to 'classifier' (*bùshǒu* 部首, sometimes termed 'radical'). *Shuōwén jiězì* contained 540 such classifiers, which function as graphic components shared among a substantial number of characters. Characters categorized according to a particular classifier all shared this component in their composition. Subsequent dictionaries, beginning from the sixth century, experimented with modifications of the 540-classifier system. Attempts were made throughout Chinese history to reorganize or simplify the classifiers of *Shuōwén jiězì*, as well as to identify new ones.²² The most significant innovation for the present reader is the creation of a 214-classifier system in the early seventeenth century, appearing in Méi Yīngzuò's 梅膺祚 *Zìhuì* 字彙 ('Collection of characters'). First printed in 1615, this text was an enormously influential dictionary both in terms of format and content (for its popularity in Japan, see Chapter 10). More important than its simplification of the number of classifiers, a lexicographic activity with ample precedent, was its pioneering lookup method. The classifiers within *Zìhuì* were ordered according to the number of strokes in each. Similarly, characters falling under a particular classifier category were ordered according to how many strokes they contained, not including the classifier. Other dictionaries, such as the early thirteenth-century *Sishēng piānhǎi*, had previously experimented with stroke-based ordering, but

²¹ Gu, Rizhi lu, j. 21, 30a–35a. ²² See Bottéro, *Sémantisme et classification*.

only in combination with an additional process of phonologically based lookup.²³

For example, in order to look up the character 時 (*shí* ‘time’), one would first identify the classifier 日 (*rì* ‘sun’) and proceed to count the strokes in the remaining portion of the character (寺), in this case six. The character 時 can then be located in the section of the dictionary dedicated to the 日 classifier, listed under characters with six additional strokes. The classifier 日 itself was grouped with other classifiers containing four strokes. The advantages of this lookup system were substantial. As logographs, Chinese characters do not convey precise phonetic information. With this method, one may look at a character, the pronunciation of which is entirely unknown to one, and locate it efficiently in a dictionary, based on its physical construction. The existence of classifier-based lookup systems since *Shuōwén jiězì* had allowed for such an approach previously. However, the organizational principle of stroke-counting as implemented in *Zihui* greatly facilitated readers’ ability to quickly locate characters.

The impact of Méi Yīngzuò’s method was immediate. In addition to the widespread use of the text documented by contemporary observers, it was highly influential on two subsequent dictionaries compiled in the following hundred years: *Zhèngzì tōng* 正字通 (‘Comprehensive rectification of characters’) and *Kāngxī zìdiǎn*.²⁴ *Zhèngzì tōng*, first printed in 1685, inherited Méi Yīngzuò’s system, but added considerably more content in terms of characters included and material cited in the lemmata. Like *Zihui*, it was printed in several editions and appears to have enjoyed widespread usage.²⁵ It was framed, along with other lexicographical works from the period such as the 1666 *Zihui bǔ* 字彙補 (‘Supplement to *Zihui*’), explicitly as a revision and expansion of *Zihui*.

Kāngxī zìdiǎn, first printed in 1716, is perhaps the best-known pre-modern dictionary in China today, and still is considered an authoritative reference work. Compiled by a committee of 30 court scholars over the course of 6 years, *Kāngxī zìdiǎn* contained more than 47,000 characters, organized according to the classifier system of *Zihui*. The 214 classifiers delineated by Méi Yīngzuò have in fact come to be known as the Kangxi classifiers, their earlier past often unrecognized. Beyond the organizational system, the court compilers of *Kāngxī zìdiǎn* also relied heavily on *Zihui* and *Zhèngzì tōng* for the content of lemmata. They nevertheless removed content from the earlier

²³ Bottéro, *Sémantisme et classification*, 152–7. ²⁴ Wu, ‘*Zihui*’ *bianzuan lilun yanjiu*, 15–16.

²⁵ Furuya, “‘Zhengzi tong’ banben ji zuozhe kao’.

dictionaries which they held to be inaccurate, and included new characters, citations, and critical commentary. Printed in numerous subsequent editions, and eventually revised in 1831, *Kāngxī zìdiǎn* represented a culmination of lexicographic scholarship in the Chinese tradition. As opposed to dictionaries directly related to *Shuōwén jiězì*, the headwords in *Kāngxī zìdiǎn* and its predecessors were not in ancient script, but rather in the standard contemporary script. Ancient script forms were nonetheless frequently referenced, either to justify the use of a particular contemporary form or to explicate the meaning of a character. Dictionaries in the *Zìhuì* lineage provided substantial content within lemmata, primarily culled from literary usage. Synonym or short descriptive glosses functioned as definitions, which were illustrated by means of citations from throughout the classical and literary tradition.

Another important, if lesser-known, approach to character analysis came to prominence during this period and had considerable lexicographic influence. As mentioned above, *Shuōwén jiězì* described six principles of character composition: each Chinese character was supposedly created according to one of these principles. One particular principle, ‘joined-meanings’ (*huìyì* 會意), was characterized by the fact that each component part of the character had semantic importance. For example, the ancient form of the character *kàn* (看 ‘to look’) is composed of two characters, ‘hand’ and ‘eye’, depicting the action of someone looking into the distance. Only a very small number of Chinese characters were identified as *huìyì* in *Shuōwén jiězì*, or in most later texts in the tradition. However, beginning in the eleventh century, a new method of glossing characters emerged, which attempted to interpret all Chinese characters as if their component parts told a story.

Wáng Ānshí 王安石, the most influential statesman of the eleventh century, composed a dictionary and required that it be promulgated throughout the school system. This text, titled *Zìshuō* 字說 (‘Explanations of characters’), was presented to the emperor in the late eleventh century and was subsequently instituted as part of the civil service examination curriculum. Wáng’s mission in *Zìshuō* mirrored his efforts in other classical commentaries to uncover the systematic coherence he believed to have governed society in antiquity. By revealing the way that component parts of ancient characters related to each other, Wáng believed he could reconstruct central elements of this system. He further established relationships between characters on the basis of this semantic content in order to draw a social or moral lesson. For example, he identified a food-related component within the characters for two different official positions. Wáng claimed that the presence of a food-related character within these two different characters (one meaning

‘minister’, the other ‘staff’) indicated that ‘both [characters] mean to nurture the people’. He further deduced that ‘whatever the king establishes is simply in order to nourish the people’.²⁶ The text was mocked even in its own time for what contemporaries considered to be an absurd extension of the *huìyì* principle. It was removed from the examination curriculum following Wáng’s fall from power and is no longer fully extant today.²⁷ Nevertheless, Wáng’s method of analysing characters did not disappear.

Perhaps the most striking text in the lineage of dictionaries related to *Zishuō* was a work by the sixteenth-century scholar-official Wèi Jiào 魏校, titled *Liùshū jīngyùn* 六書精蘊 (‘Mysterious essence of the six ways of writing down’). Wèi Jiào subscribed to an important sixteenth-century intellectual trend known as ‘the learning of the heart-mind’ (*xīn xué* 心學), which held that the path to moral enlightenment lay in uncovering the mind’s innate ability to differentiate good and evil. Wèi Jiào believed that the key to unlocking this ability lay in reconstructing the minds of the ancient sages, which could be accomplished by understanding the way in which they had constructed written characters. Although the nature of the definitions differs significantly from those in Wáng Ānshí’s *Zishuō*, Wèi alluded to the earlier text in his preface and seems to have seen some connection between his work and Wáng’s.²⁸ Like Wáng Ānshí, Wèi Jiào frequently interpreted each component part of a given character as bearing semantic importance. In particular, he identified many characters as illuminating the meaning of key terms from ‘the learning of the heart-mind’.

Wèi Jiào’s text was organized not by classifiers, as was often the case in dictionaries of the script, but rather by overarching thematic categories, such as heavens, earth, human relations, and so on. The ordering of the classifiers themselves in *Shuōwén jiězì* appears to reflect a particular cosmological worldview. Within a specific classifier category there are also signs of an intentional sequence, but a shared classifier alone is sufficient to justify inclusion within the category. In *Liùshū jīngyùn*, as well as similar contemporary texts, a character must have a clear semantic relationship to the overarching category in order to be included. This method of categorization based on meaning, rather than the sound or shape of characters, had

²⁶ ‘皆以養人為義，則王所建置，凡以養人而已’。Cf. Bol, ‘Wang Anshi and the *Zhouli*’, 241.

²⁷ For reconstructions based on extant citations in Song dynasty encyclopedias, dictionaries, and commentaries, see Huang, *Wang Anshi ‘Zishuo’ zhi yanjiu*, and Zhang, *Wang Anshi ‘Zishuo’ ji*.

²⁸ Wei, *Liushu jingyun*, xu, 6a; Rusk, ‘Old scripts, new actors’, 84–8.

precedent in works within the *Shuōwén* tradition, such as *Liùshū gù*, discussed above. Such thematic categorization schemes are likely ultimately derived, however, from the *Ēryǎ* tradition of lexicography, detailed below.

The *Ēryǎ* Tradition and Defining Meaning

As discussed in Chapter 3, the *Ēryǎ* and later works composed under its inspiration formed an integral part of the classical canon and the lexicographical tradition. *Ēryǎ*-style works were classified as works of ‘glossing’ (*xùngǔ* 訓詁). Their focus was primarily on conveying the meaning of a term, rather than information about its pronunciation or written form. Compared to lexicographic works focused on phonology and script, the *Ēryǎ* inspired relatively few later works.

The eleventh and twelfth centuries were a highpoint of *Ēryǎ* studies and the production of dictionaries in its model. Two major commentaries on the text were printed, as well as several *Ēryǎ*-style dictionaries. The original text of the *Ēryǎ* was originally linked with explicating terms in the classical *Shījīng* 詩經 (‘Book of poems’) and provided glosses for many obscure plants and animals referred to in its poems. As a result, the *Ēryǎ* came to be associated with natural historical knowledge, and later works in the tradition often focused exclusively on this domain. The two major Song works of *Ēryǎ*-style lexicography include only definitions for terms of botanical and zoological interest. One, titled *Píyǎ* 埤雅 (‘Expanded [*Ēr*]yǎ’), provides detailed, if occasionally fanciful, descriptions of *Shījīng* natural history. The other, titled *Máoshī míngwù jiě* 毛詩名物解 (‘Explanation of things in the Mao poems’), contains very similar content to *Píyǎ* with the addition of a section on various kinds of millet. The works are clearly related, although the exact nature of this relationship is unclear.²⁹ Both texts also bear a relationship to the *Zishuō* of Wáng Ānshí discussed above (*Máoshī míngwù jiě* was compiled by Wáng’s son-in-law, Cài Biàn 蔡卞, and *Píyǎ* by his student Lù Diàn 陸佃). Today they serve as one of the main sources for our understanding of Wáng’s no longer extant text, based on their extensive citations.

A wave of lexicographical production in the sixteenth and seventeenth centuries witnessed the production of several more *Ēryǎ*-style dictionaries. Two different compilations titled *Wǔyǎ* 五雅 (‘Five [*Ēr*]yǎ’) were produced during this period, arranging various *Ēryǎ*-style works through the Song dynasty. A notable text by the sixteenth-century philologist

²⁹ Mittag, ‘Becoming acquainted with nature from the *Odes*’, 326–8.

Zhāng Xuān 張萱, titled *Huǐyǎ* 彙雅 (‘Assembled [Ěr]yǎ’), similarly compiled various early works in the Ěryǎ tradition along with his own commentary and a supplement including *Píyǎ* and other Song commentaries. A sixteenth-century Míng imperial prince, Zhū Móuwěi 朱謀瑋, compiled a dictionary titled *Piányǎ* 駢雅 (‘Paired [Ěr]yǎ’), explaining the meaning of various literary disyllabic compounds, which closely followed the order of categories in the Ěryǎ. Fāng Yǐzhì 方以智, one of the great minds of the seventeenth century, composed *Tōngyǎ* 通雅 (‘Comprehensive [Ěr]yǎ’), the title of which similarly filiates it within the Ěryǎ tradition. Unlike the other works in this tradition, however, *Tōngyǎ* adopted very little in terms of structure or content from the Ěryǎ itself. Instead, it functioned as a massive, critical encyclopedia covering a vast range of topics, with particular attention to issues of phonology and script. In the work’s Editorial Principles (*fanli* 凡例), Fāng does suggest certain ways in which his text ‘imitates’ (倣) the structure of the Ěryǎ, but the end product is a work of considerably greater innovation and research than the preceding (and subsequent) models of Ěryǎ scholarship.³⁰

Another significant lexicographical work devoted to explicating the meaning of terms emerged from a very different intellectual context in the twelfth century. The Ěryǎ was associated with classicism and a tradition of textual exegesis that flourished in the Hàn dynasty. By contrast, Neo-Confucian practitioners of ‘the learning of the Way’ (*dào xué* 道學) of the eleventh and twelfth centuries opposed what they saw as the negative effects of such exegetical commentaries for a true understanding of the meaning of the classics and the Way of antiquity. The writings of the leading thinker and codifier of Neo-Confucian thought, Zhū Xī, would come to form the basis for civil service examinations from the fourteenth century through the end of the imperial period. One of Zhū Xī’s disciples, Chén Chún 陳淳, composed an influential dictionary. Despite the philological nature of the genre, Chén’s work differed strongly from any previous Chinese dictionary. Titled *Běixī zìyì* 北溪字義 (‘Meanings of characters by Beixi [i.e., Chén Chún]’), this text provided in-depth, philosophical definitions of key terms in Neo-Confucian metaphysics.³¹

The twenty-six entries of *Běixī zìyì* proceed sequentially, beginning with ‘[heaven’s] mandate’ (*mìng* 命) and concluding with the ‘false doctrines’ of Buddhism and Daoism. Each entry provides an extended essay on the headword and its philosophical implications. The text was apparently meant to be read from beginning to end, as later entries make frequent reference to

³⁰ Fang, *Tongya*, 1b. ³¹ Chén Chún, *Neo-Confucian Terms Explained*, is a translation.

previous entries. It is also unique in that the headwords are sometimes multi-character concepts, as opposed to single-character terms. For example, a five-character headword ‘humaneness, righteousness, propriety, wisdom, and faithfulness’ (仁義禮智信) discusses the relationship of these five cardinal virtues in Neo-Confucian thought.

Běixī zìyì is exceptional in the history of Chinese lexicography, which predominantly involved phonological and script analysis. It also differs from the *Ēryǎ* tradition in its emphasis on the meaning of concepts, rather than glossing terms from particular texts. *Běixī zìyì* underwent numerous later printings and enjoyed popularity in Japan and Korea. Its style of philosophical exegesis may have also been influential for later Neo-Confucian lexicographers, such as Wèi Jiào. In general, definitions in Chinese dictionaries comprised short glosses composed of a synonym or brief description. The bulk of entries in dictionaries concerned with usage involved citation of occurrences in the textual record, rather than description. *Běixī zìyì* provided an alternative model for discussing meaning removed from a context of literary usage.

Dialect Lexicography

One of the names most associated with philology in pre-modern China was Yáng Xióng 楊雄, whose *Fāngyán* 方言 (‘Regional language’: see Chapter 3) established a precedent for documenting regional varieties of Chinese speech. Despite the fact that many subsequent lexicographers claimed to be broadly following in Yáng’s footsteps, the *Fāngyán* itself received little in the way of later commentary, nor did it inspire a productive tradition of scholarship on topolects. There were, however, isolated texts from this period that documented colloquial terms or vocabulary from a particular topolect. For example, the sixteenth-century scholar Chén Shìyuán 陳士元 compiled several texts addressing colloquial terms, primarily in the textual record. His *Gǔ súzì lüè* 古俗字略 (‘Overview of ancient vernacular characters’) was a dictionary delineating what he identified as vernacular character forms in ancient texts. His *Lǐyán jiě* 俚言解 (‘Explanation of colloquial speech’) was similarly organized as an encyclopedia of vernacular usages in classical works, as well as more recent texts. Another dictionary of his, titled *Zhūshǐ yíyǔ jiěyì* 諸史夷語解義 (‘Explanation of the meaning of foreign language in the various histories’), glosses what he identified as foreign language terms in various historical works. Similar encyclopedic texts of colloquial terms, such as *Mùqián jí* 目前集 (‘Collection of [things] before one’s eyes’) and *Shìshì*

tōngkǎo 世事通考 (‘Comprehensive examination of worldly affairs’), were also produced during the sixteenth and seventeenth centuries.³²

One of the first dictionaries dedicated to a specific topolect was composed towards the end of this period. Lǐ Shí’s 李實 *Shǔyǔ* 蜀語 (‘Language of Shu’) was a seventeenth-century dictionary of the topolect of Sichuān province. The standard entry format in this work is ‘[short definition or standard term] is called [topolectal term]’ with occasional pronunciation notes. For instance, “‘unintelligent’ is called 昏憒 [*hūnchóu*], 憒 is pronounced 刀 [*dāo*].”³³ Máo Qíling 毛奇齡, an influential thinker of the period, authored a similarly structured work titled *Yuèyǔ kěngqīng lù* 越語肯綮錄 (‘A record of the essential points of Cantonese’), which analysed Cantonese terms and their pronunciation. A number of texts following this model continued to be produced in the following centuries.

Topolectal terminology is also documented in non-lexicographic sources, such as gazetteers and *biji* notebooks. There existed glossaries of slang used in the marketplace as well, such as *Shiyǔ shēngsòu* 市語聲嗽 (‘The lively sounds of market argot’). The vast majority of dictionaries in this period were focused on the literary language, explaining the pronunciation and meaning of single characters. Throughout this period, however, there was an awareness of important regional differences and colloquial usages, which were both documented by scholars and infused into the writings of literary figures.

Buddhism and Sanskrit Studies

The relationship of Chinese philology to Buddhist circles has been debated. It was once commonly held that the early history of phonology in China, and in particular the derivation of *fǎnqiè*, was linked to the phonological learning of Indian monks.³⁴ Later scholars predictably dismissed or highlighted the evidence for this link according to various ideological positions. Regardless of the veracity of this claim for the earlier period, Buddhist communities came to assume an important place in lexicographic production from the medieval period onwards. Much of the Buddhist lexicographic contribution centred around the issue of how to pronounce Sanskrit terms, which abounded in Chinese translations of Buddhist texts.

³² Present-day compilations which contain the texts cited in this section, as well as later examples, include *Zhongguo minsu fangyan yaoyan congan chubian*, and Nagasawa, *Min Shin zokugo jisho shūsei*.

³³ ‘不慧曰昏憒，憒音刀’, cited in Yang, ‘Li Shi “Shuyu” lunlue’, 85.

³⁴ Mair (trans.), ‘Preface to the Seven Sounds’.

For example, *Yīqièjīng yīnyì* 一切經音義 (‘Meaning and sounds of all sutras’), first compiled in the seventh century and expanded in the ninth, listed headwords consisting of Chinese transcriptions of Sanskrit terms, followed by definitions and phonological glosses. The tenth-century *Lóngkān shǒujiàn* 龍龕手鑑 (originally titled *Lóngkān shǒujīng* 龍龕手鏡 ‘Hand mirror for dragon abodes [i.e., Buddhist scriptures]’) was similarly compiled by a monk for the purposes of scriptural study, although it glossed single characters rather than entire terms. The majority of such works dedicated to the phonology of sutras was produced between the seventh and twelfth centuries, but there are later examples, such as the late sixteenth-century *Běijīng Wúdàbù zhíyīn huìyùn* 北京五大部直音會韻 (‘Beijing [edition of the] collected rhymes of the Five Great Sutras with phonetic glosses’).

The rhyme table tradition was also heavily associated with Buddhist scholars. Tenth-century monks employed a phonological terminology that would be adopted in most subsequent rhyme tables. A genre known as *ménfǎ* 門法, focusing on explicating the workings of earlier rhyme tables, developed in later centuries and included notable contributions from Buddhist monks, for example, the sixteenth-century *Yù yàoshi ménfǎ* 玉鑰匙門法 (‘Jade key *ménfǎ*’). Monks were forced to consider phonological issues as a result of the emphasis on chanting and accurate pronunciation in sutra recitation.³⁵

Chinese scholars considered these works to be part of the Chinese lexicographical tradition. Glossaries of non-Chinese languages were not usually viewed as a part of ‘philology’, which primarily concerned issues of classical exegesis and studies of Chinese script and phonology. However, works related to Sanskrit transcription in Buddhist sutras were often listed alongside dictionaries of classical usage as part of the same project. Because the entries often consisted of transcriptions of Sanskrit terms in widely read texts, they functioned as a kind of extended usage within Chinese. Scholars in the eighteenth century looked back at dictionaries such as *Lóngkān shǒujiàn* as important, if flawed, compilations of Chinese character variants and pronunciations.

Conclusion

Chinese dictionaries of the period 600–1700 were produced in great number and variety. Through the variation, some recurrent themes emerge. One is an attempt to uncover the nature of antiquity through language: dictionaries

³⁵ Copp, *Body Incantatory*, 4.

focused on delineating the pronunciations of antiquity as separate from the present, or sought to outline the principles governing the construction of written characters in ancient times. Another is the desire to set a standard for composition, again in terms of the proper script form and phonology. Finally, there emerged an increasing desire to describe linguistic characteristics abstractly, as evidenced by the proliferation of rhyme tables and theoretical commentaries in the *Shuōwén jiězì* tradition. While many of the individual dictionaries in this period were forgotten or actively attacked in later periods, these three lexicographic priorities have maintained relevance to the present day.

India and Tibet, c. 500–c. 1750

LATA MAHESH DEOKAR

(With a Section on Tamil by Jean-Luc Chevillard)

This chapter tells part of the story of lexicography in India and Tibet: that is, the geographical areas influenced by the Indian and Tibetan cultures. In other words, this is the story of the lexicography of the Sanskrit, Pali, Prakrit, and Tibetan languages. The time of the story to be told here is between the fifth and the seventeenth centuries. The story before the fifth century is narrated in Chapter 4, and the one after the seventeenth century in Chapters 18 and 30.

Although Sanskrit has been a predominant language in academic circles from ancient times to the present, Pali and various Prakrits also have an important place in Indian literature. These three languages are chiefly associated with three major Indian religions: Brahmanism, Buddhism, and Jainism. However, it must be remembered here that Buddhists and Jains have also contributed to the development of Sanskrit literature. Although we find numerous dictionaries in Sanskrit, this is not the case with the other two Indian languages. In Pali, we have only one large dictionary composed in the thirteenth century, apart from a glossary of monosyllabic words. In Prakrit, the situation is almost the same, with only two extant Prakrit dictionaries.

Tibet's contact with India and its subsequent adoption of an Indian Mahāyāna form of Buddhism as its state religion gave paramount importance to Sanskrit. In order to bring the Buddha's teachings to Tibet, Tibetan monk-scholars relied mainly upon Buddhist texts written in Sanskrit. A unique translation activity, which started in Tibet around the seventh century with the collaborative efforts of kings and Indian and Tibetan monk-scholars, necessitated the creation of a bilingual Sanskrit–Tibetan dictionary. After the initial phase of translating Buddhist works into Tibetan, secular Sanskrit literature attracted the attention of Tibetan literati. This gave rise to translations of secular works belonging to various genres such as grammar, lexicography, short poems, medicine, poetics, and metrics. In the field of lexicography, Tibetans chose three texts for translation. During the medieval period, a few bilingual dictionaries were composed in Tibet.

As regards classical Sanskrit lexicography, there are several types of Sanskrit dictionaries: dictionaries of synonyms, homonyms, monosyllabic words, indeclinables, words having several spellings, and series of things or concepts whose number of items is fixed; medico-botanical glossaries; dictionaries of astronomy and astrology, Buddhist nomenclature, verbs, and so on. We will discuss the first two types, synonymic and homonymic, since they are the most common.

As already noted in Chapter 4, synonymic dictionaries in the Sanskrit tradition deal chiefly, but not exclusively, with synonymous words. They also treat homonyms, indeclinables, rules on gender, adjectives, and miscellaneous to a small extent, synonyms and non-synonymic material such as homonyms being treated in separate sections. Homonymic dictionaries, on the other hand, deal only with homonymous words. Except for those dictionaries which deal with Buddhist vocabulary (*dharmasaṃgraha*), all the Indian dictionaries in the period under discussion were written in verse. This was done in order to facilitate their easy memorization.

Before we proceed with our story, some points should be borne in mind clearly. Although numerous synonymic dictionaries were compiled in Sanskrit, Sanskrit lexicographers never defined what they meant by a synonym (*ekārtha* or *samānārtha*). A. M. Ghatage discusses various aspects of synonymy as met with in the Sanskrit *kośa* 'dictionary' tradition. He concludes that by synonyms 'the Indian lexicographer means words which refer to the same individual or object and thus have the same denotation'.¹ But this is not quite as simple a definition as meets the eye. It can be explained by citing from Hemacandra's synonymic lexicon *Abhidhānacintāmaṇi*, a work of the twelfth century. According to Hemacandra, the word *nīlakaṇṭha* 'one having a blue neck' is synonymous with Śiva, the third god of the Hindu trinity. On account of consuming poison on a particular occasion, Śiva's neck became blue. Therefore, he came to be known as Nīlakaṇṭha. Thus, both Śiva and Nīlakaṇṭha refer to one and the same deity and are therefore, from the point of view of traditional Sanskrit lexicography, synonymous.

Sanskrit lexicographers did not, however, neglect different nuances of words altogether. For instance, Amarasiṃha, the author of the foundational dictionary *Amarakośa* (to which we shall return shortly) makes a distinction between 'a night', 'a dark night', 'a full-moon night', and 'a night with two days enclosing it'.² He records these nuances in the form of brief explanations, for example, 'jyautsnī is a night with moonlight'. At times, the

¹ Ghatage, 'Traditional lexica', 29. ² Amarasiṃha, *Amarakośa*, 1.4.3d–1.4.5d.

explanation consists of a single word or a compound. In such a situation, it becomes difficult to understand whether it is an explanation or a synonym. For instance, Amarasimha lists several words derived from the verbal root *lap* 'to chatter' with their explanations. He explains the word *pralāpa* 'talk' as *anarthakaṁ vacaḥ* 'speech without [any] meaning'. However, in the remaining cases, we come across pairs of words: *ālāpa* 'addressing' and *ābhāṣaṇa* 'speaking to'; *anulāpa* 'tautology' and *muhurbhāṣā* 'saying again and again'; *vilāpa* 'lamentation' and *paridevana* 'bewailing'.³ Whether to consider these pairs as synonyms or to consider one word as an explanation of the other is a question.

One more peculiarity of Sanskrit gave rise to a huge mass of words being treated as synonymous. In Sanskrit, relations which are regarded as leading to synonymy are of several types: for instance, different words which all refer to relationships of possession with a given thing, or to producers of a given thing, are seen as synonymous. For example, Hemacandra provides compounded expressions such as *bhūpāla*, *bhūdhana*, *bhūbhug*, *bhūnetā*, and *bhūpati*, all of which are formed by adding words – *pāla* 'the one who protects', *dhana* 'as wealth', *bhug* 'the one who enjoys', *netā* 'the leader', and *pati* 'chief' – to the word for the thing possessed, namely, *bhū* 'the earth'. Since the overall meaning of these words is similar, they are treated as synonyms. Inclusion of compound words is therefore a distinctive feature of synonymic lexica. By contrast, according to Ghatage,

in case of the homonymous Kośas it is mostly the simple words which are dealt with for the multiplicity of their meanings and only rarely do compound words play a part in them . . . there is no point in trying to show differences of meaning in compounds which can do so under the regular process of grammatical formations of which the lexicographer is not likely to take any special note.⁴

In case of the concept of homonymy as well, Sanskrit lexicographers did not clearly define what they meant by it, and did not seem to differentiate between homonymy and polysemy. The tenth-century lexicographer Halāyudha defines homonymy as 'occurrence of a one, single word in many meanings'.⁵ His twelfth-century successor Maṅkha distinguishes conventional and etymological meaning in his homonymic dictionary. In the case of words having only one conventional meaning, their etymological meaning is also provided. For example, the only conventional meaning of the

³ Amarasimha, *Amarakośa*, 1.6.25. ⁴ Ghatage, 'Traditional lexica', 27.

⁵ Halāyudha, *Abhidhānatnamāla*, 787.

word *gīti* is the name of a particular metre. In order to show why it belongs in a dictionary of homonyms, its etymological meaning, namely ‘song’, is also provided. On the other hand, if the word has more than one conventional sense, an etymological meaning may or may not be indicated. For instance, in the case of the word *tvaṣṭṛ*, which is conventionally used in two senses, namely, ‘the sun’ and ‘the architect of the gods’, its etymological meaning, namely ‘carpenter’, is also indicated. On the contrary, with regard to the word *priyaka*, which has a number of conventional meanings such as ‘a bee’, ‘saffron’, and ‘a certain medicinal plant’, its etymological meaning, namely ‘a lover’, is not listed.⁶

The *Amarakośa*

The story of Sanskrit, and, for that matter, of Indian, lexicography of this period starts with the most famous Sanskrit lexicon, titled *Nāmaṅgānuśāsana* ‘treatise [dealing with] nouns and [their] gender’, and popularly known as *Amarakośa*, from the first part of the name of its author, Amarasimha, plus *kośa* ‘dictionary’. This is the first complete Sanskrit lexicon available to us. As far as I am aware, no text-critical edition of this popular lexicon has ever been produced. The existence of manuscripts in a huge number and their availability in various regional and pan-Indian scripts could be the main reasons for this glaring lacuna.

Amarasimha’s lexicon was and is still considered as the most authentic classical lexicon of the Sanskrit language. Numerous commentators of classical literature from various parts of India have used it extensively to justify the usage or the attestation of a particular word. This certainly highlights the high esteem this lexicon enjoyed throughout the Indian subcontinent. Another testimony to its popularity is the fact that the lexicon has been commented upon no fewer than eighty times.⁷ Apart from this vast commentarial literature, at least two adaptations and six translations were made in neighbouring countries such as Tibet, Mongolia, and Sri Lanka, and it continued to play a significant part in Indian lexicography into the nineteenth century (see Chapter 18).

Amarasimha, like authors of most of the classical and medieval Indian texts, begins his dictionary with a benediction. Benediction, the Indian tradition believes, is a way to make sure that the text reaches its completion. It also

⁶ Mañkha, *Mañkhakośa*, 6; see Ghatage, ‘Traditional lexica’, 31.

⁷ Raghavan, *New Catalogus Catalogorum*, I.324–31.

reflects the author's faith, and may thus hint at the kind of religious vocabulary to be expected in a given dictionary. In this respect, *Amarakośa* stands apart. Its benedictory verse does not mention the name of a specific deity.⁸ This has given rise to a debate among commentators and modern scholars as to the religious affiliation of Amarasimha. The earliest commentators interpret the benedictory verse as homage to the Buddha, whereas the later South Indian commentators try to establish Amarasimha as a follower of the Brahmanical tradition. Early commentators, such as the twelfth-century Sarvānanda and the fifteenth-century Rāyamukūṭa, say that since Amarasimha wanted his lexicon to be acceptable to people of all faiths, he did not reveal his Buddhist identity in clear terms. Modern scholars generally interpret the verse in favour of the Buddha. According to them, Amarasimha's placement of epithets of the Buddha before important deities of the Brahmanical tradition and those of the Bo tree before other plants hints at his being a Buddhist. Although Amarasimha includes some vocabulary which is typically Buddhist, he did not emphasize Buddhist matters in the same way as his successor Puruṣottamadeva (to whose lexicon and its Buddhist content we shall return). Hence, the debate regarding his religious identity is still inconclusive.

If we glance through other prominent early lexica, we immediately notice that other lexicographers clearly revealed their faith. To take two examples five or more centuries apart, the early lexicographer Śāsvata pays homage to Viṣṇu, the second of the Hindu trinity, and the fourteenth- or fifteenth-century Vāmanabhaṭṭa Bāṇa to Śiva. Halāyudha, whose definition of homonymy I cited above, pays homage to 'the supreme spirit in the form of speech'. Even Yādavaprakāśa, an eleventh-century lexicographer who appears to begin his work without a benedictory verse, mentions the auspicious word *svarga* 'heaven' at the beginning of his dictionary, and this is considered benedictory. Three Buddhist lexicographers of the late eleventh or early twelfth centuries, Ajayapāla, Puruṣottamadeva, and Maheśvara, pay homage to the Buddha. Among the Jain lexicographers of Sanskrit, Hemacandra is the only one who begins his dictionary by paying homage to the Jain teacher Mahāvīra. In the first benedictory verse of his lexicon, Harṣakīrti, the second Jain lexicographer, pays homage to *paramātman* 'the supreme spirit' instead of Mahāvīra or any other Jain teacher. In the next three verses, he pays homage to the goddess of knowledge, Sarasvatī, and to Gaṇeśa, the remover of obstacles. Sometimes the homage with which

⁸ Amarasimha, *Amarakośa*, I.I.I.

a dictionary opens may help us to understand its maker: there is disagreement as to whether or not the twelfth- or thirteenth-century lexicographer Śrīdharasena was a Jain, and the fact that he pays homage to *Bhagavān* ‘lord’ and *dharma* ‘teachings’, which are the two most venerated objects in the Buddhist tradition, while he does not refer to Jain subjects, is suggestive.⁹

There was an uninterrupted tradition of dictionaries starting from the *Amarakośa* up to the eighteenth century. What could have motivated generations of scholars to write these dictionaries? If we consult the introductory portions of prominent lexica, their motives can be deduced as follows. The first was to provide a lexicon complete in all respects: this is already to be found in the *Amarakośa*.¹⁰ The second, expressed by lexicographers after Amarasimha, was to fill the lacunae of previous lexica.¹¹ The third was to help ignorant people understand Sanskrit literature.¹² The fourth and last was to help poets in their composition.¹³ This last motive does not mean that lexicographers coined an artificial new vocabulary to be used in literary compositions. As Katre has pointed out, ‘the artificial creations must be limited to cases where the lexicographer has sanskritized a vernacular expression current during his days’.¹⁴ Therefore, it seems reasonable to assume that lexicographers brought together words and usages current in their region and time at one place, which proved helpful to poets. Sometimes, due to the lack of a critical study of attestations of vocables in literature, common people and scholars tend to think that lexica sometimes contain artificial vocabulary. Such a contention is, however, proved wrong when new texts using such vocabulary come to light. For example, the word *punḍarīka* in the sense of ‘tiger’ was believed to exist only in lexica until the publication in 2008 of a work of the first half of the eleventh century, the Sanskrit text of Jñānaśrīmitra’s *Vṛttamālāstuti* (‘Garland of praise of Mañjuśrī’s conducts in various metres’), where the word occurs in this sense.¹⁵ There is, however, another dimension of such late attestations of rare words. We see a growing

⁹ Śrīdharasena, *Abhidhānaviśvalocana*, 1a; Vogel, *Indian Lexicography*, 75, sees him as a Jain, and Lozang Jamspal (in Śrīdharasena, *Abhidhānaviśvalocana*, introduction, vii–ix) as a Buddhist who was ‘born in the royal clan of the Sena kings [of] Bengal’.

¹⁰ Amarasimha, *Amarakośa*, 1.1.2.

¹¹ Puruṣottamadeva, *Trikāṇḍaśeṣa*, 1.1.2; Maheśvara, *Viśvaprakāśa*, 18–19.

¹² Śāśvata, *Anekārthasamuccaya*, 2–4.

¹³ Śāśvata, *Anekārthasamuccaya*, 5; Halāyudha, *Abhidhānaratnamālā*, 1.2; Dharaṇidāsa, *Anekārthasāra*, 5d.

¹⁴ Katre, ‘Amara’s contribution’, 202.

¹⁵ The lexica in which the word occurs in this sense are Amarasimha, *Amarakośa*, 3.3.11c; Śāśvata, *Anekārthasamuccaya*, 3 (verse 14); Halāyudha, *Abhidhānaratnamālā*, 2.71; Ajayapāla, *Nānārthasamgraha*, 57 (verse 11c); Puruṣottamadeva, *Trikāṇḍaśeṣa*, 102 (verses 32–3); and Medinikara, *Nānārthasābdakośa*, 16 (verse 200). See also Böhtlingk, *Sanskrit-*

tendency among the post-seventh-century Sanskrit writers to use rare words deliberately just to exhibit their scholarship, resulting in the revival of some obsolete or obsolescent words.

In order to achieve their aims, Sanskrit lexicographers culled their vocabulary from a wide selection of literary texts. This is evident from the large number of attestations which are provided by commentators. As these attestations show, the literary sources on which Sanskrit lexicographers drew belong to various genres of Sanskrit literature, such as *Vedas*, epics, law, short poems, plays, prose, great poems with their commentaries, hymns, didactic literature, grammatical literature, poetics, drama, metrics, lexicography, astronomy, medicine, veterinary science, philosophical literature, politics, and regional Sanskrit literature such as the text from Kashmir quoted by Mañkha in the auto-commentary to his homonymic lexicon. Commentators use citations to provide information regarding the exact usage, different meanings, variant spellings, gender, and number of a particular word and mythology. Commentaries also act as a supplement to the lexicon by supplying the vocabulary that was missing in it. So, for instance, the commentators Kṣīrasvāmin and Mallinātha often added words that were missing from the *Amarakośa* at the end of a particular section of their commentaries on it. Such additions later found their way into the main body of the *Amarakośa* as represented by some of its manuscripts.

In an introduction to a lexicon, next to the statement of motive, comes the description of its lexicographical sources. Instead of naming his particular sources, Amarasimha speaks of them in very general terms as 'other works'.¹⁶ Later lexicographers usually name their sources. This helps in identifying previous, long-forgotten lexicographers and also in assigning dates to them. The description of lexicographical sources is usually followed by rules of interpreting a lexicon and devices used for securing brevity.

After these introductory verses, the lexicon proper begins. As mentioned, earlier Sanskrit lexica had separate sections dealing with synonyms on the one hand, and homonyms and other material on the other. Amarasimha too followed this overall plan. Thus, he deals with synonyms in the first two sections (*kāṇḍa*) of his lexicon, which treat nouns pertaining to heaven and earth respectively, in ten chapters apiece. The third section deals with general vocabulary, in five chapters. The first section contains words related to heaven, sky, quarters of the sky, time, thought, sense-objects such as

Wörterbuch, IV.756. For the date of Jñānaśrīmitra, see M. Hahn in Jñānaśrīmitra, *Vṛttamālāstuti*, 6.

¹⁶ Amarasimha, *Amarakośa*, I.1.1.2.

sound, drama, the nether world, serpents, hells, and water. The second section deals with words related to earth, towns, mountains, medicinal herbs, animals and so on, human beings, the priestly class, the warrior class, the merchant class, and the servant class. The third section treats adjectives, miscellaneous words, homonyms, indeclinables, and rules on gender.

Commentators of the *Amarakośa* have subdivided its sections in different ways. For instance, some have divided the first section into two chapters rather than ten, the one dealing with nouns related to heaven and the other with nouns related to the nether world, the division between the two coming just after the section on words connected with drama. The Tibetan translators of the *Amarakośa* recognized only this broader division of the first section. It appears that commentators were not happy with the inclusion of words related to hell in the section related to heaven. They tried their best to justify Amarasimha by saying that hell is a concept opposite to heaven and therefore it is justifiable to include words belonging to these two concepts in one section.

However, later lexicographers, Indians as well as those non-Indian ones influenced by Amarasimha, did not follow him in this respect. They usually grouped words related to heaven up to drama in the first section, words belonging to the earth into the second one, and the vocabulary concerning the nether world into the third one. Instances of this type of division are found in the *Abhidhānaratnamālā* of Halāyudha and the *Trikāṇḍaśeṣa* of Puruṣottamadeva. In addition to the three sections found in the previous lexica, the eleventh-century Yādavaprakāśa adds a fourth dealing with words related to atmosphere and a fifth treating general vocabulary. Within each of these sections, there are subdivisions. For instance, the first section is further divided into first gods, regents of quarters of the world, and semi-divine beings; the second one into planets and stars, clouds, birds, sounds, and so on. In the twelfth century, the *Abhidhānacintāmaṇi* of Hemacandra has five sections dealing with the vocabulary related to chief gods or Jinas, the gods of the Brahmanical and the Buddhist traditions, mortals, animals, and hell-dwellers. The synonymic part of the *Abhidhānaviśvalocana* of Śrīdharasena, belonging to the twelfth or thirteenth century (and extant only in its Tibetan translation), falls into two sections.¹⁷ The first section deals with heaven and the nether world, and the second with the earth, towns, mountains, herbs, animals, man, the four castes, adjectives, and varia. Although separate

¹⁷ Vogel, *Indian Lexicography*, 76.

chapters on the sky, its quarters, time, thought, and other important topics are missing, most of them are dealt with in other suitable places. Later, in the fourteenth or fifteenth century, Vāmanabhaṭṭa Bāṇa arranged synonyms in twenty-two chapters in his *Śabdaratnākara*, including a chapter on sages, and dividing ‘many of the chapters of [*Amarakośa*] . . . under subtopics with different captions’.¹⁸ So, for instance, according to the editor of this lexicon, the chapter on plants and trees deals with the subject matter in great detail, including terms related to minor semantic categories such as cookery. Later still, Harṣakīrti treats his synonymic vocabulary in three sections: the first is further subdivided into three chapters, which deal with gods, sky, and the earth; the second treats words related to the body, social life, music, and scholars; the third deals with words related to priests, kings, peasants, servants, and miscellanea. Towards the end of the period under consideration here, in the seventeenth century, Viśvanātha’s *Kośakalpataru* is divided into five sections. The first three sections deal with vocabulary related to heaven, earth, and the subterranean world (the subdivision of chapters within each section is similar to the one found in the *Amarakośa*); the fourth treats adjectives, miscellanea, verbs, genders, and indeclinables; the fifth deals with homonymous words. Here, words are arranged according to their final consonant following the Nāgarī alphabet. These are ‘further grouped by the alphabetical order of the individual consonants and by the increasing number of syllables’.¹⁹

As well as adding sections or chapters and rearranging the sections, lexicographers rearranged the sets of vocabulary as well. For instance, after listing the epithets of Brahmā, the creator of the universe, Halāyudha brings in the synonyms of speech, the sacred literature of Brahmanism, and vocabulary related to the same.²⁰ This reshuffle is justified since, according to Hindu mythology, speech is related to Brahmā. Amarasimha listed these words in the independent chapter of the first section in the *Amarakośa* which deals with sounds.²¹ Yādavaprakāśa follows Halāyudha and lists synonyms of speech after the epithets of Brahmā.²² One more case of reshuffling is based on philosophical grounds. Yādavaprakāśa was a great exponent of the Viśiṣṭādvaita, a school of philosophy in which sound is considered as a quality of the sky. In his *Vaijayantī*, he therefore includes words related to sound in the section related to the sky.²³

¹⁸ B. R. Sharma, in Vāmanabhaṭṭa Bāṇa, *Śabdaratnākara*, foreword, 5.

¹⁹ Vogel, *Indian Lexicography*, 93. ²⁰ Halāyudha, *Abhidhānatnamāla*, 7cd, 8–10.

²¹ Amarasimha, *Amarakośa*, 1.6.11cd–1.6.12ab.

²² Yādavaprakāśa, *Vaijayantī*, 6cd–9ab (Brahmā), 9cd (speech).

²³ Yādavaprakāśa, *Vaijayantī*, 2.3.

This kind of reshuffle is vital evidence for the importance attached by an individual lexicographer to a particular concept or personality or deity. Amarasimha, for example, placed epithets of the Buddha before those of important Brahmanical deities. All later lexicographers, except Halāyudha, arranged the section of epithets of the Buddha alongside epithets related to Viṣṇu, the second of the Hindu trinity. This corresponds to the status of the Buddha in the medieval period, where he was looked upon as an incarnation of Viṣṇu. This in turn also explains the increase in the number of epithets of the Buddha in later lexica. Halāyudha is the only lexicographer who arranged the epithets of the Buddha alongside epithets for evil fortune. This reflects his antagonistic approach towards the Buddha. Among the Jain lexicographers, Hemacandra included the Buddha among the rest of the Vedic deities, whereas Harṣakīrti placed him before the principal Vedic deities but after the Jina Mahāvīra.²⁴ Śrīdharasena did not deal with Jain matters and was inclined towards Buddhism. This analysis highlights the importance of Sanskrit lexica from the sociocultural point of view.

It will be clear from the preceding discussion that, in synonymic lexica, there was not much scope for improvement except rearranging the sections, the chapters, or the semantic categories. On the other hand, homonymic lexica, starting from the very early *Anekārthasamuccaya* of Śāśvata, came to be arranged more systematically and precisely than the homonymic section of the *Amarakośa*, in which homonyms are arranged simply on the basis of their final consonants, with no further structure within each of these divisions. (The same was true of the other non-synonymic sections of synonymic dictionaries.) Śāśvata arranged homonyms according to the regressing length of a verse required to describe them: for example, words requiring a complete verse for their description are listed first, whereas words described in a single foot occur last. The fifth section of the *Abhidhānaratnamālā* of Halāyudha, which may be contemporary with or later than the work of Śāśvata, deals with homonyms, but presents them at random, not even maintaining the structure introduced by Amarasimha. Thereafter, the homonymic lexica tend to be more tightly structured. In the eleventh or twelfth century, Ajayapāla arranged the entry words according to their initial letters. In the early twelfth century, in the chapter of the *Trikāṇḍaśeṣa* that deals with homonyms, Puruṣottamadeva introduced a far better arrangement, by the initial and then the last letter of the entry words – so, for instance, words beginning with *a*, the first vowel of the Sanskrit language and ending with *ka*, the first

²⁴ Hemacandra, *Abhidhānacinātmanī*, 1.2.146–150c; Harṣakīrti, *Śāradīyākhyānāmamālā*, 6–7.

consonant of the language, are listed first, followed by words beginning with the vowel *ā* and ending with *ka* – but without further alphabetization.²⁵ In the same century, Dharaṇidāsa listed the entry words according to the alphabetical order of their final consonants and then according to the length of a verse required to list the meanings of a particular word (first those requiring one foot, then two feet, and then a verse), and then according to their number of syllables. In the twelfth or thirteenth century, Śrīdharaśena followed the practice of alphabetization laid down by Puruṣottamadeva, adding alphabetization of the second syllable of the word, and in the thirteenth century Medinīkara divided his lexicon on similar principles, but with further classification according to the number of syllables.²⁶ Yādavaprakāśa had already developed a more elaborate structure, classifying the lemmata first according to the number of syllables, then according to their genders, and within each of these subclasses, according to the alphabetical order of their first letters; due to this improvement, the homonymic part of his lexicon is quoted profusely in the Sanskrit commentarial literature. Later lexicographers more or less followed one of these patterns. To conclude, it is clear that, although we certainly see development in the arrangement of homonymic vocabulary, it was not adopted uniformly: lexicographers rather seem to have worked with their individual preferences.

So far we have seen the arrangement of various Sanskrit lexica and its development over the centuries. Now let us see how Sanskrit lexicographers dealt with meanings current in their own times and regions.

Lexicographers incorporated words into their own lexica that had been omitted from previous lexica or that had come into use in their own times. Notably, Puruṣottamadeva compiled a completely new lexicon called *Trikāṇḍaśeṣa* with the sole aim of supplementing the *Amarakośa*. Since the latter has three sections (*tri-kāṇḍa*), the title *Trikāṇḍaśeṣa* ('Remainder to the *Trikāṇḍa*') is appropriate. The nature of this supplement was twofold. First, it included additional words or synonyms which were missing in the *Amarakośa*. For instance, *Amarasimha* records seven synonyms of Gautama, the Buddha, and Puruṣottamadeva supplies three more.²⁷ Secondly, it supplied semantic categories that were missing in the *Amarakośa*. For instance, Puruṣottamadeva includes the epithets of son of the Buddha and those of a *Pratyekabuddha*, that is, a Buddha who lives in seclusion and obtains emancipation for himself only.²⁸ Similar instances are

²⁵ Puruṣottamadeva, *Trikāṇḍaśeṣa*, 3.5.1. ²⁶ Medinīkara, *Nānārthaśabdakośa*, verse 13.

²⁷ *Amarasimha*, *Amarakośa*, 1.1.14d–15; Puruṣottamadeva, *Trikāṇḍaśeṣa*, 1.1.11cd–12a.

²⁸ Puruṣottamadeva, *Trikāṇḍaśeṣa*, 1.1.12, 13.

not infrequent. For example, Amarasiṃha lists the word *udarka* in the sense ‘future result of an action’. Śāśvata includes one more meaning, namely, ‘future time’. Later lexicographers such as Ajayapāla, Puruṣottamadeva, Yādavaprakāśa, and Maṅkha follow Śāśvata.²⁹ Mallinātha, a commentator of the *Amarakośa*, must have taken some hint from these lexica, for he remarks that ‘*Udarka* means the future result of an action. *Udarka* also means future time.’³⁰

Thus, while Sanskrit lexicographers followed the overall plan laid down by their predecessors as far as possible, they also incorporated new vocabulary and kept their works up to date, thereby attracting generations of scholars and poets.

Prakrit

This phase of the story of Prakrit lexicography comprises two dictionaries, both of which call themselves dictionaries of provincial words. Hemacandra, the second lexicographer of this period, cited works of his predecessors, who had presumably composed lexica of provincial words. Their names are Abhimānaciḥna, Gopāla, Devarāja, Droṇa, Dhanapāla, Pādaliptācārya, Rāhulaka, and Śīlāṅka; out of these eight, only Dhanapāla’s *Pāiyalacchī* has come down to us. So, in this section, we will look at Dhanapāla’s work first, and then at Hemacandra’s.

Dhanapāla’s *Pāiyalacchī* or *Prākṛṭalakṣmī* ‘the wealth [or the glory] of the Prakrit language’ is a synonymic dictionary.³¹ Although Dhanapāla calls his treatise ‘a garland of nouns’, he also includes other parts of speech such as adverbs, verbal forms, particles, and affixes. According to him, his lexicon contains provincial or native words (*deśī*).³² Hemacandra defines such words as those used by Prakrit authors, which are not derived from Sanskrit by the rules of Prakrit grammar and have a different meaning from that of the corresponding word in Sanskrit.³³ Thus, an author’s knowledge of Sanskrit literature and lexica, and to some extent the character of his Prakrit grammar, would decide which words he treated as purely provincial and which ones he regarded as loanwords from Sanskrit. Since Dhanapāla compiled his work ‘as

²⁹ Amarasiṃha, *Amarakośa*, 2.8.29d; Śāśvata, *Anekārthasamuccaya*, 347cd; Ajayapāla, *Nānārthasamgraha*, 14 (verse 2ab); Puruṣottamadeva, *Trikaṇḍaśeṣa*, 93; Yādavaprakāśa, *Vaijayantī*, 240 (12); Maṅkha, *Maṅkhakośa*, 40ab.

³⁰ Mallinātha, in Amarasiṃha, *Amarakośa* (1971), 1.501.

³¹ G. Bühler, in Dhanapāla, *Pāiyalacchī*, 70. ³² Dhanapāla, *Pāiyalacchī*, 1d; 278d.

³³ P. K. Ramanujaswami, in Hemacandra, *Deśināmamālā* (1938), 3.

an introduction into the language of the Prakrit poets', he probably included whatever he remembered.³⁴ As a result of this policy, provincial words 'form not more than one fourth of the total of the words' given in the *Pāiyalacchī*.³⁵ The rest are either loanwords from Sanskrit or words that are related to Sanskrit but have undergone Middle Indic phonological change. Georg Bühler, the nineteenth-century editor of the *Pāiyalacchī*, is therefore justified in criticizing Dhanapāla by saying that many of the words included 'are so well known that not even the merest beginner in Sanskrit and Prakrit could mistake them for Deśī-terms and ... are by no means required for the purpose of explaining the real Deśīs'.³⁶ The *Pāiyalacchī* begins with homage to God who is *paramapurisa* ('the Supreme Spirit') and *purisuttama* ('the highest of men'). It is divided into four sections: synonyms requiring a verse, or a hemistich, or a single metrical foot to explain them, and single words explained by one synonym. These principles are laid down at the beginning of each section. The first three sections include names of gods, saints, and sacred things; in them, as Bühler points out, 'nouns denoting the most various things, adjectives, adverbs, inflected verbal forms, particles, and even affixes are mixed up without any order or principle of arrangement'.³⁷

Dhanapāla also, by the way, compiled a Sanskrit lexicon, now lost. Subhūticandra, a commentator of the *Amarakośa*, quoted two Sanskrit citations from it, calling it *Dhanapālābhīdhāna*. Hemacandra in his auto-commentary to the *Abhidhānacintāmaṇi*, and Bhānuji Dīkṣita, another commentator of the *Amarakośa*, also cited it.³⁸

The second Prakrit dictionary of this period is Hemacandra's *Deśināmamālā*. Hemacandra himself calls the work *Ratnāvalī* ('A string of jewels').³⁹ The title *Deśināmamālā* ('A garland of provincial words') occurs in the best manuscripts of the text. Yet another title, *Deśīśabdasaṃgraha* ('A collection of provincial nouns'), is found in the work.⁴⁰ Hemacandra's lexicon begins by paying homage to the Jina. It is divided in eight chapters arranged according to the initial letter of the words following the Nāgarī script. In each chapter, synonyms are given first and then homonyms. They are further subdivided in alphabetical order by their first vowels and then in order of the increasing number of syllables. Hemacandra, who was a great scholar of Sanskrit and Prakrit, sought to exclude all words derived from Sanskrit.

³⁴ G. Bühler, in Dhanapāla, *Pāiyalacchī*, 78–9. ³⁵ G. Bühler, in Dhanapāla, *Pāiyalacchī*, 76.

³⁶ G. Bühler, in Dhanapāla, *Pāiyalacchī*, 78. ³⁷ G. Bühler, in Dhanapāla, *Pāiyalacchī*, 76.

³⁸ Subhūticandra, *Kavikāmadhenu on Amarakośa*, 211; Hemacandra, *Abhidhānacintāmaṇī*, 76, 79; Bhānuji Dīkṣita, *Vyākhyāyāsudhā*, 33.

³⁹ Hemacandra, *Deśināmamālā*, 8.77. ⁴⁰ Hemacandra, *Deśināmamālā*, 1.2 and 8.77.

However, at times, even he included words which were evidently derived from Sanskrit. The *Deśināmamālā* also contains some words which are of Dravidian origin and, as noted by its editor P. K. Ramanujaswami, there are some Persian and Arabic words as well.⁴¹

Pali

From the late tenth century onwards we see a revival of Pali literature in Sri Lanka and Myanmar, with the composition of many new Pali texts. Although several of these were original in nature, many were modelled on Sanskrit texts.

For our present story, it is interesting to note that Cūla Moggallāna, a Sri Lankan scholar of the late thirteenth century, compiled his Pali lexicon *Abhidhānappadīpikā* by modelling it after the *Amarakośa*. Being a Buddhist, Cūla Moggallāna starts his lexicon by paying homage to the Buddhist trinity: the Buddha, his teachings, and the order of monks. He writes that he composed this lexicon to help people in attaining proficiency in understanding nouns and their gender for comprehending the Words of the Buddha.⁴² He follows the same principles and methodology for his lexicon as laid down by Amarasimha.⁴³ Although Cūla Moggallāna does not mention his sources by name, it is clear from his vocabulary that he used *Amarakośa* along with materials culled from the Pali canonical and commentarial literature.

The main divisions of the *Abhidhānappadīpikā* are the same as those of the *Amarakośa*. However, it does not adopt the *Amarakośa*'s further subdivisions into chapters. Cūla Moggallāna follows the commentarial tradition of the *Amarakośa* and therefore includes the contents of only the first seven chapters of the *Amarakośa*, up to drama, in his first section. The second section of the *Abhidhānappadīpikā* includes the last four chapters of the first section and the entire second section of the *Amarakośa*. The third section of the *Abhidhānappadīpikā* is similar to that of the *Amarakośa* with the single exception that it omits the chapter dealing with gender. Within these sections, we observe some changes in their internal arrangement, compared to that of the *Amarakośa*. For instance, synonyms of directions and subdirections, and the names of the guardian elephants of eight directions and those of the regents of four directions, are listed in the first section in the context of the guardian deity of the east (*Sakka*). In the *Amarakośa*, these appear in the third chapter,

⁴¹ P. K. Ramanujaswami, in Hemacandra, *Deśināmamālā* (1938), 8–9.

⁴² Cūla Moggallāna, *Abhidhānappadīpikā*, 4–5.

⁴³ Cūla Moggallāna, *Abhidhānappadīpikā*, 6–9; Amarasimha, *Amarakośa*, I.1.6–8.

which deals with directions. It seems that Cūla Moggallāna must have felt it fit to list these nouns in the context of *Sakka* as he is considered to be the lord of the eastern direction. Moreover, by shifting the nouns related to the directions to the first section, Cūla Moggallāna could achieve better consistency in the text since now the synonyms of sky are immediately followed by those of clouds, lightning, and the like.

After the introductory verses the lexicon begins with the epithets of the Buddha. These are followed by synonyms of *nibbāna* 'liberation'. After this, although, as I have just observed, Cūla Moggallāna follows the same sequence as that of the *Amarakośa*, he includes only selected vocabulary from it. So, he tends to omit epithets of deities belonging to the Brahmanical tradition, since they do not occur in the Buddhist literature. In contrast, he brings in vocabulary that is typically Buddhist. For instance, after listing the synonyms of Mount Meru, he lists seven mountains that are found in the Pali literature. He likewise incorporates a good many terms of Buddhist flavour in general and of Theravāda in particular, such as *samatha* 'tranquillity', *dhīti* 'steadfastness', and *virīya* 'right effort'. However, he also incorporates many words which are not attested in the Pali literature, merely converting Sanskrit words into Pali. K. R. Norman is therefore justified in saying that, 'A proportion of the vocabulary in the *Abhidhanappadīpikā* is therefore artificial, in the sense that it had no existence in Pali until it had been specially coined for inclusion in the dictionary'.⁴⁴ When later poets culled such words from the *Abhidhānappadīpikā* and used them in their compositions, these artificially coined words came into existence in their own right.

The twelfth-century Pali grammatical work *Saddanīti* of the Burmese scholar Aggavaṃsa contains discussions regarding synonyms and their functions. For instance, while explaining the meaning of the verbal roots *kakhi*, *vakhi*, and *makhi* 'to doubt' in the *dhātumālā* ('garland of verbal roots'), Aggavaṃsa cites a list of synonyms of doubt from the canonical text *Dhammasaṅgaṇi*, and points out those synonyms which are more popular.⁴⁵ In the *padamālā* ('garland of words'), another part of his work, he cites a list of synonyms of 'a being' from the *Mahāniddeśa*.⁴⁶ This list is followed by three verses which give synonyms of 'a being'. From the structure of these verses, they appear to be from some Pali lexicon not known to us today. There are similar verses in the *Abhidhānappadīpikā*, but Aggavaṃsa does not mention

⁴⁴ Norman, *Pāli Literature*, 167. ⁴⁵ Aggavaṃsa, *Saddanīti*, 330, citing *Dhammasaṅgaṇi*, 85.

⁴⁶ Aggavaṃsa, *Saddanīti*, 64, citing *Mahāniddeśa*, 68.

the name of this text or of its author, and so it is not possible to say conclusively whether he knew this lexicon or not.

Before we conclude our story of Indian lexicography, it is necessary to understand the important place held by the commentarial literature of Indian lexica. As we have seen, *Amarakośa* has been commented upon at least eighty times, and other lexica such as *Abhidhānaratnamālā* and *Viśvaprakāśa* have been commented upon a few times. Hemacandra and Mañkha wrote auto-commentaries to their respective lexica, and the *Abhidhānappādīpikā* was commented upon several times in Sri Lanka and Myanmar. This art of writing a commentary on a lexicon seems to have attracted scholars even in the modern age. For instance, a Sri Lankan monk-scholar, Seelakkhandha Mahathera, wrote a commentary on Puruṣottamadeva's *Trikāṇḍaśeṣa* in the early twentieth century.⁴⁷ These commentaries provide information regarding the etymologies, grammatical status, spellings, and genders of words. Many of them also provide mythological information and citations from the literature. The study of the commentarial literature also throws light on the Sanskrit literature that is lost to us irretrievably. For instance, Subhūticandra's *Kavikāmadhenu* commentary on the *Amarakośa* cites from many texts which are otherwise unknown to us. Thus, on the one hand, the commentarial literature carries out some functions of modern dictionaries and encyclopedias and, on the other, it points to the literature that has passed into oblivion.

Indian tradition believes that grammar and lexica are two main pillars of learning. It regards Pāṇini's grammar titled *Aṣṭādhyāyī* as the world's mother and *Amarakośa* its father. As a result, both these texts formed an important part of the curriculum of traditional Sanskrit schools, and students were expected to learn these texts by heart. According to Indian tradition, etymology is like a female bird. Lexica and grammar are her two wings. With their help, one should dive into the infinite sky of learning.

Tibetan

In the Indo-Tibetan Buddhist thought and civilization, the science of lexicography is regarded as one of the minor sciences ancillary to the major science of grammar. All the secular as well as religio-philosophical sciences were considered 'to make up the complete cultural equipment of the Bodhisattva and of the educated person modeling his career on that of the Bodhisattva'.⁴⁸

⁴⁷ Seelakkhandha Mahathera, *Sārārtha Candrikā*.

⁴⁸ Ruegg, 'Sanskrit–Tibetan and Tibetan–Sanskrit dictionaries', 132.

The sixteenth-century scholar sGom sde Nam mkha' rgyal mtshan says that 'since Buddhas in all three times give teachings in Sanskrit, there exists the motive that we should implant [in ourselves] a predisposition (*vāsanā*) in respect of this language'.⁴⁹ According to the eighteenth-century polymath Si tu paṇ chen, after one fixes the use of lexemes and grammatical gender, one acquires unwavering insightful intelligence in carrying out the study and understanding of the words of the Buddha.⁵⁰ Hence, lexicography constitutes an auxiliary required by those seeking liberation. Bearing this motive in mind, let us now turn to the history of Tibetan lexicography.

Mahāvyutpatti (*Bye brag tu rtogs par byed pa*), the first Sanskrit–Tibetan dictionary, is a well-known text. It was compiled with the objective of standardizing the Tibetan vocabulary of the translated Tibetan literature and achieving uniformity in terminology and translation techniques. The focus of this dictionary was obviously the Sanskrit Buddhist literature which was being translated at that time in Tibet: it was compiled as 'a response to the need for a suitable tool for translating Indian Sūtras and Śāstras into Tibetan'.⁵¹ Not a single classical Sanskrit dictionary which we know of today pays special attention to Buddhist matters. Moreover, the *Mahāvyutpatti*, unlike the classical Sanskrit dictionaries, was written in prose. Whereas Sanskrit synonymic lexicographs are arranged in sections related to heaven, the earth, and miscellanea, the *Mahāvyutpatti* is divided into several chapters that deal with various semantic categories, especially Buddhist ones. More than half of its chapters are related to Buddhist terminology. The remaining chapters include non-Buddhist philosophical terms belonging to six Indian schools of philosophy, names of important non-Buddhist teachers, and secular vocabulary that occurs in Buddhist texts. These facts mean that the *Mahāvyutpatti* differs substantially from classical Sanskrit lexicographs, and a number of the words which it records, although attested in the pre-Amarasimha Buddhist as well as classical literature, are not recorded in the *Amarakośa* or any known Sanskrit lexicon. From this point of view, the contribution of the *Mahāvyutpatti* still calls for detailed study.

The autobiography of Khro phu Lo tsā ba Byams pa'i dpal, written in 1235, mentions that Zhang Lo tsā ba Dge ba instructed him in the *Amarakośa* and other linguistic sciences.⁵² This might be seen as the beginnings of the study of the *Amarakośa*, its commentarial literature, and other Sanskrit lexicographs in

⁴⁹ Ruegg, 'Sanskrit–Tibetan and Tibetan–Sanskrit dictionaries', 116 n. 1.

⁵⁰ Si tu paṇ chen, *Collected Works*, vol. V, fos. 210b–211a.

⁵¹ Ruegg, 'Sanskrit–Tibetan and Tibetan–Sanskrit dictionaries', 131.

⁵² Van der Kuijp, 'On the vicissitudes', 6.

Tibet. A few decades later, a biographer of the Sa skya scholar Kun dga' rgyal mtshan writes that the latter had studied the *Amarakośa* and Maheśvara's *Viśvaprakāśa* in the first decade of the thirteenth century. His study of both these texts resulted in the compilation of the *Tshig gi gter* ('Word treasury'), which is an abridged metrical translation of select passages from the *Amarakośa*, made no later than c. 1280.⁵³ Unlike the later Tibetan translations of the *Amarakośa*, this lexicon begins with the epithets of the Buddha. The sequence of other semantic categories remains the same as that of the *Amarakośa*. Just like the *Amarakośa*, the text of the *Tshig gi gter* is also divided into three main sections: epithets of the Buddha, words related to heaven, and those related to the subterranean regions.⁵⁴ In the first section, the author has added a few more epithets of the Buddha. The second section of the *Tshig gi gter*, following the commentarial tradition of the *Amarakośa*, ends with the words related to drama. Kun dga' rgyal mtshan deals with the contents of the second section of the *Amarakośa* rather briefly.⁵⁵ The third section of the *Tshig gi gter* deals with the world of serpents, the subterranean region, beings residing in hell, sufferings, and water.⁵⁶

The *Tshig gi gter* was followed by the first Tibetan translations of the *Amarakośa* and Subhūticandra's *Kavikāmadhenu* commentary thereupon, made no later than the mid fourteenth century, by Grags pa rgyal mtshan and Kīrticandra. Since both the translations were faulty, they were revised before the mid sixteenth century by Chos skyong bzang po, who was also responsible for the translation of Śrīdharaśena's *Abhidhānaviśvalocana*. The translation of the *Amarakośa* was further revised in the seventeenth or early eighteenth century. The last and best revision of the translations of the *Amarakośa* and the *Kavikāmadhenu* was carried out by Si tu paṇ chen in the eighteenth century. Si tu is also credited with a Tibetan commentary on the *Amarakośa*. Here, he has listed the Sanskrit terms occurring in the *Amarakośa* in their original sequence, followed by their Tibetan equivalents – which are not always those which occur in the versified translation, where metrical constraints affected the word choice – and occasional comments. These comments are mainly about variant readings, gender, number, or declension of particular words, and the technical instructions given in the *Amarakośa*; they do not occur in the versified translation.

⁵³ Van der Kuijp, 'On the vicissitudes', 22. ⁵⁴ Kun dga' rgyal mtshan, *Tshig gi gter*, 4.

⁵⁵ Kun dga' rgyal mtshan, *Tshig gi gter*, 201 (words related to drama), 201–6 (material corresponding to second section of *Amarakośa*).

⁵⁶ Kun dga' rgyal mtshan, *Tshig gi gter*, 206–50, corresponding to *Amarakośa* 1.8–11. Cf. van der Kuijp, 'On the vicissitudes', 14.

The sixteenth century witnessed the compilation of a Tibetan dictionary titled *Mkhas pa'i ma rgyan*, compiled by Ngag dbang 'jig rten dbang phyug grags pa'i rdo rje. A main feature of this dictionary is that it is compiled in verse. The dictionary is divided into four sections. The first three sections deal with heaven, the nether world, and the earth respectively; in each, the words are arranged in semantic categories, but within each category, they are placed randomly. The fourth section, which is further subdivided into two chapters, deals with miscellanea and homonyms. The chapter dealing with homonyms is bilingual: that is, it includes the Sanskrit word with its Tibetan equivalent. Homonyms are arranged according to their final syllable and then by the number of syllables. Although the author does not strictly follow the division of the *Amarakośa*, he does try to include many of its semantic categories. As well as the *Amarakośa*, he identifies the *Kavikāmadhenu* of Subhūticandra, the *Abhidhānaviśvalocana* of Śrīdharaśena, and the *Tshig gi gter* and a commentary on it as sources of his lexicon. The influence of Sanskrit lexica, and especially of those compiled by Buddhist authors, on Indo-Tibetan lexica was, in conclusion, considerable.

The Development of the Poetical/Classical Tamil Lexicographical Tradition

JEAN-LUC CHEVILLARD

The Classical Tamil lexicographical tradition became an independent field of knowledge with the composition of autonomous treatises.⁵⁷ The most frequent generic designation for these is nowadays *nikaṇṭu*, although their most ancient designation is *uric col* (or *uric col paṇuval*), after the title (*Uri-yiyal*) of the seventeenth chapter in the *Tolkāppiyam*, the most ancient Tamil text with lexicographical content (see the section on Tamil in Chapter 4).⁵⁸ Like the *Tolkāppiyam*, these treatises were in verse, and they were probably intended to be recited from memory.

Currently available to us are the texts of 18 *nikaṇṭus*, which occupy 1,700 pages in A4 size, with 50 lines per double-column page, in a recent popular

⁵⁷ An indispensable reference on Tamil lexicography in general is Gregory James' well-known book *Col-porūl*; see also Zvelebil, *Lexicon of Tamil Literature*.

⁵⁸ The well-known thirteenth-century grammar called *Nannūl* explicitly refers in the third line of its 460th *cūttiram* to the *Piṅkalam* as being one of those *uric col* treatises. Some *Nannūl* commentators use the longer expression *uric col paṇuval*, where *paṇuval* means poetical composition.

collected edition.⁵⁹ The two most ancient of these are the *Tivākaram* (containing 2,518 *cūttirams*, or versified rules, totalling 4,365 metrical lines) and the *Piṅkalam* (4,121 *cūttirams*, totalling 6,782 metrical lines).⁶⁰ These two works may date back to the eighth or ninth century, and to the tenth, respectively, although providing precise dating is a very difficult challenge, for many reasons – the first of which is that palm-leaf manuscripts cannot survive for a long time in the Indian climate and have to be periodically copied, again and again, so that there are no very ancient manuscripts in Tamil Nadu, such as are found in Europe. A comparison of the size of the *Uriyiyal* and of the *Piṅkalam* (which would take more than eight hours for a full recitation) shows how much development took place between the age of the *Tolkāppiyam* and the age of the *nikaṇṭus*. It should not be thought, however, that progress in lexicographical knowledge was completely linear, because one can find a few words which are in the *Uriyiyal* and not in the *Tivākaram*, and one can also find words which are in the *Tivākaram*, but not in the *Piṅkalam*.⁶¹ These two most ancient *nikaṇṭus* are followed by the very concise *Uric col nikaṇṭu*, composed by Kaṅkēyar, possibly in the fourteenth century.⁶² A long tradition ensued: the *Pārati tīpam* in the fifteenth century; the *Kayātaram*, composed by Kayātarar, in the fifteenth or sixteenth; the *Cūṭamaṇi Nikaṇṭu*, the *Akarāti* (which is the first among the *nikaṇṭus* to make use of this title, used in Modern Tamil to mean ‘dictionary’), and the *Kailācam* in the sixteenth; two more in the seventeenth century, two more in the eighteenth, and six more in the nineteenth.⁶³

Inside this long series, however, a natural milestone must however be placed after the seventeenth-century *nikaṇṭus*. That milestone, which is a monolingual work called *Caturakarāti* (‘Quadruple dictionary’), was composed by an Italian Jesuit missionary called Constanzo Beschi (for whom see below, and also Chapter 30), who was the first Westerner to really leave his mark in the literary history of Tamil, becoming himself a Tamil *pulavar*, still remembered today

⁵⁹ *Tamiḷ Nikaṇṭukaḷ* (2008).

⁶⁰ *Tivākaram* (1990–3) and *Piṅkalantai eṇṇum Piṅkala Nikaṇṭu* (1968) are recent editions.

⁶¹ For instance, the word *puṇiṇu*, which is attested many times in ancient Tamil literature, as a specifier to the name of female animals (who have just delivered, at the end of pregnancy), is explained in the *Uriyiyal* (TC375c) as meaning *īṇṛ-aṇimai* ‘recency of delivery’ and in the *Tivākaram* (Ti-1502), with a slightly different gloss, but seems to have been forgotten by the *Piṅkalam*.

⁶² Kaṅkēyar, *Uric col nikaṇṭu* (2003), is an edition. For the date, see Aruṇācalam, *Tamiḷ Ilakkiya Varalāru*. Among the objective elements for dating the *Uric col nikaṇṭu*, we have the fact that it contains the expression *calām ceytal* ‘to do *salām*’ (borrowed from Arabic), in its 176th verse, which enumerates words used for a salutation (or expression of respect).

⁶³ *Kayātaram* (1939) is an edition of the second of these.

under his adopted name of Vīramāmuṇivar.⁶⁴ After the age of Beschi, Classical Tamil lexicography progressively becomes part of a wider domain, which we can tentatively call ‘Global Tamil lexicography’, for reasons which should become clear in due course. Before that, however, we must return to our examination of the lexicography of *Cen-Tamiḷ* ‘Refined Tamil’, in the age which starts with the *Tivākaram*.

The two most important visible parameters which must be discussed when giving a material account of the Tamil *nikaṇṭus* are the general scheme which is followed in the division of the subject matter into sections, and the metre used, with its practical consequences for ease of memorization (for those trained to memorize learned texts) and difficulty of use (for those not so trained). Regarding the former, it can be said that the majority of Tamil *nikaṇṭus* follow the structural model of the *Tivākaram*, which was divided into twelve chapters, and that the *Piṅkalam* is the most visible exception to that scheme, being divided into ten chapters, the content of which cannot be very easily mapped in a linear manner with the content of the *Tivākaram*, because it is rearranged (and vastly augmented). At a deeper level, however, the *Piṅkalam* is not very different from the *Tivākaram*, and both *nikaṇṭus* can be roughly divided into three sections, treated here in the order of the *Tivākaram* (and of its followers).

First is a thematic (encyclopedic) section in which the main concern is to enumerate many quasi-synonyms for a given notion, expressed by a main term, and in which we can also find a few definitions for some difficult words. This section typically constitutes more than 70 per cent of the whole work, and is divided into ten chapters (or ‘collections’), of which the first enumerates the names – as well as other attributes – of the gods; the second to the seventh enumerate the names of human beings, animals, trees and plants, places, ‘many [natural] substances’, and ‘fabricated forms’; and the last three cover ‘names based on qualities’, ‘names based on actions’, and ‘names based on sound’.

Next is a dictionary section (in the Western sense), in which the task at hand is to deal with polysemy by enumerating all the possible meanings of many polysemic words, and which occupies roughly 20 per cent of the global structure. In the *Tivākaram* and its followers, this section is the eleventh chapter, whereas in the *Piṅkalam* it is the tenth. Some of the

⁶⁴ Beschi, who has a statue on the Marina Beach in Chennai, is also known for writing four grammars of Tamil (three in Latin and one in Tamil), and for composing Christian Tamil poetry (and notably the *Tēmpāvaṇi*). For him and his dictionary, see Muttusami Pillai, *Brief Sketch*, and Inṇāci, *Caturakarāṇi Ārāycci*.

polysemic words have, indeed, very many meanings, like the word *ari*, which has fifteen meanings enumerated in the *cūttiram* 2,266 of the *Tivākaram* and forty-one distinct meanings enumerated in the *cūttiram* 3,085 of the *Piṅkalam*. Staying away from such an extreme case, the average number of meanings for the 1,091 polysemic words enumerated by the *Piṅkalam* is 3.4, an increase (on both counts) compared with the 381 polysemic words enumerated in the *Tivākaram*, with an average of 3.07 meanings.

The twelfth chapter in the *Tivākaram* and its followers, but the third chapter in the *Piṅkalam*, is a ‘traditional groups’ section, which occupies a little more than 5 per cent of the total structure, and contains elements of information concerning groups (of items) which are traditionally associated with numbers, and which can sometimes be used as substitutes for those numbers (as in the Sanskrit *bhūta saṃkhyā* method).⁶⁵ The groups are enumerated in ascending order, starting with one; moving to two, exemplified for instance by ‘the two luminaries’ (sun and moon) and ‘the two parenthoods’ (maternal and paternal); then to three, as in ‘the three fires’ and ‘the three faults’, and so on – although the first group named in the *Piṅkalam* is ‘the seven *rishis*’ – and frequently culminating in eighteen, as in ‘the eighteen *purāṇas*’.

Although this is not a matter which can be completely understood without a lot of technical explanations (and without long practice), some information about the metrical form of the *nikaṇṭus* should be provided. The metrical variety of these works parallels the variety seen in other types of scholarly productions, such as grammars (in a wide sense of ‘grammar’ which includes poetics). One introductory point should be made: the normal type of rhyme in poetical Tamil is not final, but relies on alliteration. The most prominent rhyming technique, called *etukai*, is based on a repetition (between two or more lines) of the consonant in the second syllable. Thus, for example, in *cūttiram* 1,881 in the thematic section of the *Piṅkalam*, which contains ten metrical lines and enumerates forty-three quasi-synonyms for the word *tukkam* ‘sadness’, (metrical) line 1 starts with *allal* and alliterates with line 2 which starts with *cellal*. Similarly *evvam* (l. 3) alliterates with *kavvai* (l. 4), *paṭar* (l. 5) with *iṭar* (l. 6), *kalakkam* (l. 7) with *alakkaṇ* (l. 8), and *alamaral* (l. 9) with *iḷcam* (l. 10). The second type of alliteration is called *mōṇai* and relies on a phonetic similarity between the first syllables of two (or more) words (generally taken from the same metrical line). In the same stanza, we can

⁶⁵ For the *bhūta saṃkhyā* method, see Plofker, *Mathematics in India*, 47. A similar practice is seen in some Tamil technical texts, where big numbers have to be written as a rebus.

say that, in line 1, there is *mōṇai* between the first word (*allal*) and the third (*arantai*); in line 2, there is *mōṇai* between the first word (*cellal*) and the third (*ciṛumai*); in line 3, there is *mōṇai* between the first word (*kavvai*) and the fourth (*kavalai*), and so on.

By way of an overview, we can say that the two earliest *nikaṇṭus* (like the earliest grammars) make use of a metre called *ācīriyam*, in which there is no constraint on the number of metrical lines found in a stanza, but in which each metrical line contains four metrical feet, the use of rhyme being optional. The following *nikaṇṭu*, which is the *Uric col nikaṇṭu*, makes use of a different metre, called *nēricai veṇṇpā*, in which there are always four lines per stanza (of four feet each, excepting the fourth line, which always has three feet), and in which two *etukai*-rhyming groups are found in each stanza. The practical consequence is that, where the *Tivākaram* and *Piṅkalam* respectively enumerate sixty-four and ninety-five names for Civaṇ (Śiva) in single verses of twenty-five and thirty lines, the *Uric col nikaṇṭu* needs seven individual stanzas to enumerate seventy-nine names of Civaṇ. Another consequence is that where the *Tivākaram* and the *Piṅkalam* may make use of many short individual *cūttirams*, the *Uric col nikaṇṭu* will pack several entries (up to six if they are very brief) inside one verse, which can create ambiguities. Among the other *nikaṇṭus*, we see still other metres, but these cannot be discussed without long technical explanations which would take us away from our main narrative.

We now turn to a dimension which is probably easier for a modern reader to appreciate, because it is nowadays so ubiquitous, and therefore usually taken for granted, namely alphabetical order. Examining the *nikaṇṭus* which were composed up to the seventeenth century, the main observations we can make are as follows. The *Tivākaram* does not make any use of alphabetical order (and the *Uriyiyal* of the *Tolkāppiyam* did not either), but we have occasional sporadic groupings of elements based on alliteration. The *Piṅkalam* makes use of alphabetical order in its tenth chapter, which deals with 1,091 polysemic words and which is divided into 10 subsections: the first one contains the 227 polysemic words that start with a vowel, the second one contains the 204 words starting with [k], and so on. The sixteenth-century *Cūṭamaṇi Nikaṇṭu* also makes use of a special ordering for its eleventh chapter, dealing with polysemic words, but the order used is not alphabetization but the *etukai* method of alliteration. There being eighteen consonants in the Tamil alphabet, the chapter is divided into eighteen subsections. The first words in each of the four lines of the first verse of the first subsection are (respectively) *pakavē*, *pakalē*, *makarē*, (*y*)*akamaṇam*, with *etukai* alliteration on

[k], and this is also found in the following verses: the first words in each of the four lines of the first verse of the second subsection are (respectively) *puṇkamē*, *tiṅkaḷam*, *kaṇkenpa*, *toṅkalē*, with *etukai* alliteration on [ṇ]; and so on.

The *Akarāti Nikaṇṭu*, also of the sixteenth century, which contains only a polysemic section, dealing with 3,340 items, is an example of still another scheme. The words are grouped by initial letter, in ten large sections (corresponding to the sections of the tenth chapter of the *Piṅkaḷam*); each section is further subdivided into subsections, where all the words have the same initial (C)V syllable; those subsections are divided into sub-subsections, based on the number of meanings. As an illustration, we can say that section 2, where all words start with *k*, is divided into twelve subsections (containing words starting respectively with *ka*, *kā*, *ki*, *kī*, *ku*, *kū*, *ke*, *kē*, *kai*, *ko*, *kō* and *kau*), and that these subsections are themselves divided into sub-subsections: words starting with *ka* and having one meaning, words starting with *ka* and having two meanings, and so on, up to words starting with *ka* and having twenty-two meanings. This last feature does not make the *Akarāti Nikaṇṭu* an efficient reference tool, because the user cannot be sure how many meanings are ascribed to a given word: so, for example, the user looking for the word *puṇṇiru* must examine one by one the seventy-eight items in the relevant subsection, before concluding that it is not there.

It is now time to return to the examination of some events which took place at approximately the same period as the composition of the *Cūṭāmaṇi Nikaṇṭu* and the *Akarāti Nikaṇṭu*, the consequence of which was that Classical Tamil lexicography progressively became a component in Global Tamil lexicography. I have already mentioned Beschi and his *Caturakarāti*, but he was not the first Westerner to try to master the very complex linguistic reality of Tamil Nadu. The interactions between Western missionaries and native Tamil scholars had their starting point in the sixteenth century, which is the time of the arrival of the Portuguese in South India (for the missionaries' study of Tamil in the context of their other activities in South Asia, see Chapter 30). The history of the external linguistic description of Tamil Nadu starts with the sixteenth-century *Arte em Malabar*, by Henrique Henriques, which is a brief grammatical description of a dialect spoken by Tamil fishermen (on the Pearl Fishery Coast) and was published for the first time in book form in 1982.⁶⁶ On the lexicographical side, the earliest remaining trace which we have of missionary efforts is the *Vocabulario Tamulico com a significacão Portuguesa*, by Antão de Proença, which was printed posthumously in

⁶⁶ For it, see Zwartjes, *Portuguese Missionary Grammars*, 28–44.

Ambalacatta in 1679, and which contains 16,208 (non-lemmatized) bilingual entries on 508 pages.⁶⁷ Concerning the traditional *nikanṭu*s, the *Vocabulario* contained a few entries (such as *peṇṭir*) belonging to the Tamil poetical vocabulary.⁶⁸ However, its preface clearly stated that

I shall leave out also poetic words (of which the vocabulary of Fr Ignaccio Bruno is full) because they are useless for practical purposes or for prose, and the Tamil poets have their own vocabularies for them, in which those who are interested and might wish to compose verses may find them.⁶⁹

Another parameter in the linguistic situation of Tamil Nadu, which I have not yet emphasized, is the fact that several among its authors must have been bilingual scholars, equally at ease in Classical Tamil and in Sanskrit, although having Colloquial Tamil as a mother tongue.⁷⁰ That bilingual competence of many Tamil scholars is one of the reasons why we see the vocabulary of Tamil progressively absorbing many Sanskrit words, and the importance of that vocabulary was noted by Proença, immediately after the passage which I have just quoted:

But I include many Sanskritic words, both for the sake of those who read the books of Fr Robert [Nobili] in which they are numerous, and also because they commonly occur in the ordinary conversation of the Brahmins, whose language is more elevated, and whom Tamilians who consider themselves learned and wish to speak seriously and with care try to imitate.⁷¹

This background being now given, we are now in a position to understand what Beschi did when he composed his *Caturakarāṭi* or ‘Quadruple

⁶⁷ See Chevillard, ‘How Tamil was described once again’, and Chevillard, ‘Challenge of bi-directional translation’.

⁶⁸ Proença, *Vocabulario Tamulico* (1966), 217 col. 2, *peṇṭir*, glossed and labelled ‘Molher, honorifice, hê palaura de uerso’ (‘woman: honorific; poetic word’); further examples in Chevillard, ‘Challenge of bi-directional translation’, 213.

⁶⁹ Proença, *Vocabulario Tamulico* (1966), 13 (translation by E. C. Knowlton and X. S. Thani Nayagam); the original is fo. [v]v, ‘Deixarey tambem as palauras poeticas (de que o vocabulario do P. Ignacio Bruno está cheo) porque não seruem nada, pera â pratica, ov prosa. e dellas tem os poetas Tamuis seus vocabularios, â onde os curiosos, que quizerem compor versos, as podem ver.’

⁷⁰ For instance, Kayātarar has embedded in the text of his *nikanṭu* (at the end of each section) a number of details about himself, such as his being a Brahmin and being competent in the ‘Northern language’ (*vaṭamoḷi*, verse 147) – Sanskrit – and in ‘fertile Tamil’ (*celun-tamiḷ*, verse 187), or ‘Southern arts Tamil’ (*ten-kalai-tamiḷ*, verse 223).

⁷¹ Proença, *Vocabulario Tamulico* (1966) 13; the original is fös. [v]v–[vi]r, ‘Ponho porê muitas palauras puramente Grandonicas, assy pera quem ler os liuros do P[adre] Roberto, aonde estão muitas, como por q[ue] fã ordinarias no ordinario modo de fallar dos Bramanes. Cujo fallar hê mais subido, aquem os Tamuis, q[ue] se presaõ, de doutos, & querem fallar graues, & selecto, querem imitar.’

dictionary'. If the second component of the name, *Akarāti*, is taken from the *Akarāti Nikaṇṭu*, Beschi went much further than this dictionary in the efficient use of alphabetical order. Additionally, he took over (and re-ordered) the three-part structure which I have described earlier when explaining the general organization of *nikaṇṭu*s, but added to it a fourth category, inspired by the success of the grouping by *etukai*, which I have described above when presenting the division of the eleventh chapter of the *Cūṭāmaṇi Nikaṇṭu* into eighteen subsections. The general structure of Beschi's *Caturakarāti* is therefore quadripartite, as suggested by his title. A first section emulates the dictionary section of earlier *nikaṇṭu*s: it is called *Peyar akarāti* 'dictionary of nouns', and occupies 44.5 per cent of the global structure. A second emulates the thematic (encyclopedic) section of earlier *nikaṇṭu*s: it is called *Poruḷ akarāti* 'dictionary of things', and occupies 19 per cent of the global structure. A third emulates the 'traditional groups' section of earlier *nikaṇṭu*s: it is called *Tokai akarāti* 'dictionary of collections', and occupies 8.5 per cent of the global structure. A fourth emulates the organizational principle of the eleventh chapter of the *Cūṭāmaṇi Nikaṇṭu*: it is called *Toṭai akarāti* 'dictionary of [alliterative] rhymes', and occupies 28 per cent of the global structure.⁷²

As we can see, by creating such a structure, Beschi positioned himself as a champion of *Cen-Tamiḷ* 'Refined Tamil', taking a stand completely the opposite of the utilitarian position illustrated by the first quotation above from the preface of the *Vocabulario*. Moreover, in his *Grammatica latino-tamulica*, Beschi explicitly criticizes the choices made in the *Vocabulario* to print (vulgar) spoken forms such as *kaṇṇu* 'calf', instead of using the corresponding high form, which in this case would be *kaṇru*: he justifies this particular objection on the grounds that *kaṇ* 'eye' and *kaṇru* 'calf' had both become *kaṇṇu* in spoken Tamil, so that the high forms avoided a confusion present in the spoken language.⁷³ Although that grammar is supposed to be a grammar of Koṭuṇ-Tamiḷ (and should therefore be a description of spontaneously spoken forms), we are clearly in a situation where this is not the case, because the forms which native speakers use spontaneously are deprecated by Beschi, in conformity with a diglossic view of the world, which is still present nowadays, where what is ordinary language is considered as vulgar.

⁷² According to Inṇāci, in Beschi, *Vīramāmuṇivar aruḷiya Caturakarāti* (1989), v–vi, the *Caturakarāti* was printed ten times in the nineteenth century, but the first edition, published in 1819, contained only the *Poruḷ akarāti*. The following edition, which came out in 1824, added the text of the remaining three parts.

⁷³ Beschi, *Grammatica Latino-Tamulica*, 19.

I must now fast-forward to the nineteenth century, by which time the colonial British domination over Tamil Nadu was well established.⁷⁴ This was the time when several of the works which I have described were printed in book form for the first time, after being transmitted for centuries partly orally and partly as (fragile) palm-leaf manuscripts. As far as *nikaṇṭus* are concerned, the first to be printed may have been the very popular *Cūṭāmaṇi Nikaṇṭu*, of which chapters 1 to 10 were printed in 1834–5 and chapter 11 was printed in 1836.⁷⁵ As for the *Tivākaram*, its first ten chapters were jointly edited and published in 1839 by Tāṇṭavarāyamutaliyār (chapters 1 to 8) and by Nayanappamutaliyār (chapters 9 and 10): that is to say, they not only collated the text from manuscripts but also produced the equivalent of a minimal commentary by presenting the text of the *cūttirams* in parallel with a *peyarppirivu* ‘division into nouns’ in which they undid the sandhi, occasionally reordered the items, removed decorative words, and decided which was the main item, using it as a title (the titles, but not the remainder of the *peyarppirivu*, were retained in the first complete edition of the *Tivākaram*, published in the following year). The edition of 1839 was in other respects not a completely faithful reproduction of the *Tivākaram*, because the editors made some modifications in the beginning of the text, placing first the *cūttiram* enumerating sixty-four names of the god Civaṇ (Śiva), whereas the original contained in that place (because of its being rooted in the Jain religion) the *cūttiram* enumerating forty-three names of Arukaṇ (Arhat). They had also, in a spirit of improvement, interpolated a significant number of *cūttirams*, dutifully indicated by using stars. However, the following editors were not as scrupulous, and did away with the stars, and it was not until 1990–3 that a critical edition of the *Tivākaram* appeared, published by the University of Madras, in which the original text was restored, based on the examination of many manuscripts. The nineteenth century also saw the printing of the *Piṅkalam* (in 1890). In that edition, we have the illustration of another task which *nikaṇṭu* editors considered as useful, namely providing the modern Tamil equivalents for many of the words in the original text (those lexical substitutions are still to be found in more recent editions of the *Piṅkalam*).⁷⁶ It should also be added that some of the scholars who edited Tamil *nikaṇṭus*, such as Tāṇṭavarāyamutaliyār, had had their initial training

⁷⁴ For the study of Tamil in this period, see Venkatachalapathy, ‘Grammar, the frame of language’.

⁷⁵ Gros, ‘Review’, 348.

⁷⁶ See, for the *Piṅkalam*, Chevillard, ‘On a 1968 incarnation of the *Piṅkalam*’, 25.

(in the art of making books using a printing press) by being charged with producing an edition of Beschi's *Caturakarāṭi*.⁷⁷

In Guise of a Conclusion (as an Esquisse)

The history of Tamil lexicography in the nineteenth and the twentieth centuries (and also in the twenty-first century) is the history of the search for an elusive or slowly moving '*point d'équilibre*' between three driving components of the linguistic reality of Tamil Nadu. The first of these is Classical Tamil, which is the language, with many historical layers, of the body of literature which has been (rediscovered and) edited from the nineteenth century onwards, during what has been described as the 'Tamil renaissance'. The part of it which is the most alive among native speakers of Tamil is probably the vast devotional (or *bhakti*) corpus. The second is Formal Tamil, which is the H component of Tamil diglossia and the oralized and written component of intellectual exchanges between Tamil-speakers, used in schools, in books and journals, and also in exchanges on the internet. The third, Spoken Colloquial Tamil, which is the L component of Tamil diglossia, has a wide range of dialectal variation, and is strongly deprecated in written communication, although every native speaker of Tamil uses it in every ordinary circumstance. It is also the language used in modern cinema.

The monumental *Tamil Lexicon* sponsored by the University of Madras contains more than 117,500 entries (over 4,400 pages) and was compiled and printed during the first half of the twentieth century, under the guidance of its general editor, S. Vaiyapuri Pillai, whose prefatory materials include the first history of Tamil lexicography.⁷⁸ The *Tamil Lexicon* could be described as 'a-chronous', because it intends to be of use for both Classical and Formal Tamil, which is of course an impossible task. Attempts have been made, of course, to remedy this situation, and the best-known of the dictionaries which are dedicated to Formal Tamil might be the Tamil–Tamil–English *Kriyāvin Tarḱālat Tamil akarāṭi: Tamil–Tamil–Āṅkilam*, or 'Cre-A dictionary', now in its second edition. As far as Classical Tamil is concerned, the most useful available tool at the moment is the five-volume *Varalārru Murai Tamil Ilakkiyap Pērakarāṭi* ('Tamil dictionary/Glossary on historical principles project'), totalling 2,361 pages for the entries (not counting the front material), which was prepared over many decades by a group of traditional Tamil

⁷⁷ Tāṇṭavarāyamutaliyār is named in the undated edition of 1824 or later preserved as Pondicherry, EFEO Library, shelfmark EO DIC TA 5.

⁷⁸ Vaiyapuri Pillai, 'History of Tamil lexicography'.

scholars and published in 2001 by the Cānti Cātaṇā trust in Chennai. Attempts to describe Spoken Colloquial Tamil, on the grammatical and on the lexicographical side, continue to meet with strong resistance, due to the collective diglossic psyche, of which I have provided an example when mentioning Beschi's criticism of Proença's dictionary.

Arabic to c. 1800

RAMZI BAALBAKI

Interest in philology by the Arabs is reported in the sources to have begun only a few decades after the Qurʾān had been revealed to Prophet Muḥammad. The *Hegira* of Muḥammad from Mecca to Medina in AD 622 marks the beginning of the Islamic calendar, and thus the first century after the *Hegira* (AH) corresponds to the seventh century AD. But it is not until the second half of the second/eighth century that the earliest works in both grammar and lexicography were authored. Sībawayhi's *al-Kitāb* laid the foundations of the grammatical tradition, and most later authors were faithful to its analytical methods, arguments, terminology, and scope. Sībawayhi was deeply influenced by his teacher al-Ḥalīl b. Aḥmad, himself the author of the first (non-thematic) Arabic lexicon, *Kitāb al-ʿAyn*. But although grammar and lexicography became two distinct disciplines at a very early stage, they shared much of their material and were both closely related to other linguistically oriented sciences, such as *qirāʾāt* (Qurʾānic readings), *Ḥadīth* (Prophetic tradition), *fiqh* (jurisprudence), and *tafsīr* (exegesis). In fact, for centuries to come, many scholars straddled both disciplines, as did, for example, the renowned fifth-/eleventh-century grammarian al-Zamahṣārī, who also authored a unique dictionary titled *Asās al-balāḡa*.

The influence of the Qurʾān on early philological activity cannot be exaggerated. Not all the words in the Qurʾān were readily understood by some speakers of Arabic and thus the need arose to explain these words and demonstrate the semantic change undergone by others in a Qurʾānic context. Furthermore, the spread of *lahn* (solecism; see below), especially in reading the Qurʾān, is considered in the biographical sources of the third/ninth century onwards to have contributed significantly to the awareness of the need for determining proper usage in syntax and morphology, and for elucidating the meanings of certain vocables and

identifying Qur'ānic words that are not of Arabic origin.¹ In both grammar and lexicography, scholars of Arabic – mostly in the second half of the second/eighth century and the first half of the third/ninth – competed in collecting linguistic data from the *A'rāb* or Bedouin informants, whose speech was considered to be the 'purest' form of Arabic and thus superior to other varieties, in particular the speech of city-dwellers.²

The collection of data from the *A'rāb* certainly preserved much of the Bedouin cultural life that would have otherwise been lost to us. Linguistically, it preserved a large number of lexical items that later found their way to the lexica and provided us with valuable knowledge about a host of Arabic dialects. Yet, in transmitting usage by the *A'rāb*, the lexicographers were more interested in rare, even obsolete, words than in more commonly used words in spoken or literary language. The bulk of the rare and strange words is referred to as *ġarīb* material and, as we shall see, several scholars authored dictionaries devoted exclusively to these words. Reliable Bedouin informants, however, became a rare commodity by the end of the fourth/tenth century, after which the philologists stipulated that the corpus should no longer be kept open. Accordingly, neologisms and the semantic development of lexical items were rarely recorded in subsequent lexica, and most authors became primarily interested in expanding their material by incorporating earlier works into their lexica and perfecting the arrangement of the roots and, to a lesser extent, of the material within the lemmata.

In addition to the Qur'ān and Bedouin speech, pre-Islamic and early Islamic poetry was the third major source of data for the lexicographers. Being the genre which contains the vast majority of *ġarīb* material, poetry features prominently in grammatical and lexicographical works alike. Most dictionaries abound with attested data (*šawāhid*) derived from poetry and often include lexical items attested only in poetry. Of the sixteen metres of Arabic poetry, the *raġaz* metre is the one most closely identified with *ġarīb* usage, most probably due to its close affinity to desert life. The lexicographers considered it the embodiment of the Bedouins' linguistic purity and usually competed in reciting it and explaining the strange usage that it contains. For example, the lines of *raġaz* in Abū Zayd al-Anṣārī's book on rare usage, titled *al-Nawādir fī l-luġa*, amount to two-thirds of the number of lines of all other fifteen metres put together. The philologists also availed themselves of two

¹ Cf. Ibn Sallām, *Ṭabaqāt*, 1.12; Zubaydī, *Ṭabaqāt*, 12. See also Baalbaki, *Legacy*, 2–5.

² The process of data collection is discussed in Baalbaki, *Arabic Lexicographical Tradition*, 16–29.

other relatively minor sources of linguistic data, namely the Prophetic tradition and proverbs.

With Arabic as a spoken language expanding in vast areas of the Islamic world and acquiring a prestigious stature as the language of administration following the Arabization of the *dawāwīn* (official registers) during the reign of 'Abdalmalik b. Marwān (r. 65/685–86/705), early scholars felt the need to identify criteria for 'correct' usage. Accordingly, the earliest grammarians focused on what they considered to be 'good' Arabic syntactically and morphologically, often contrasting it with less acceptable variants or with anomalous usage. On the other hand, the early lexicographers focused on two types of semantic study, namely, the classification of the huge number of lexical items which they amassed into various topics (such as the human body, animals, insects, rain, swords, and wells) or linguistic phenomena (such as *ḡarīb* usage, Arabized words, homonyms, and synonyms), and the inclusion of all Arabic roots and their derivatives in one lexicon in which they are arranged according to a certain principle that users should familiarize themselves with. During this early period, the less frequently attested lexical items had to be authenticated or else were preserved but described as doubtful or, less frequently, incorrect.

The two types of semantic study conducted by the lexicographers obviously gave rise to two types of lexica which dominated the tradition for several centuries. The fifth-/eleventh-century Andalusian philologist Ibn Sīda clearly distinguishes between these two types, which he calls *mubawwab* and *muḡannas*.³ The first term refers to the onomasiological lexica or thesauri in which meaning leads to sign. Hence a book that includes terminology related to horses, such as al-Aṣma'ī's *al-Ḥayl*, is an onomasiological lexicon which specializes in a specific theme or semantic field. It lists lexical items in no systematic order, although some items are grouped together based on the topic under discussion, for instance desirable traits in horses or their manner of movement. The other term, *muḡannas*, refers to the semasiological lexica in which sign leads to meaning. In this type, lexical items are uniformly arranged based on their constituent letters and not according to their semantic fields. Contrary to *mubawwab* lexica that specialize in one or more topics, lexica of this type normally aspire to treat the whole corpus of Arabic exhaustively. Exceptionally, however, the author might limit the scope to a specific type of usage, as did al-Šaybānī, whose *Kitāb al-Ġīm* is, as we shall see, the first *muḡannas* lexicon to be alphabetically arranged, but whose

³ Ibn Sīda, *Muḡaṣṣaṣ*, 1.10, 12.

material is almost totally restricted to *ġarīb* usage. It should also be noted that the boundaries between *mubawwab* and *muğannas* lexica are not always clear. For instance, some *mubawwab* dictionaries are alphabetically arranged, as in certain works that deal exclusively with proverbs or with verbal and nominal patterns. On the other hand, Ibn Durayd, author of the *muğannas* lexicon titled *Ġamharat al-luġa*, curiously includes in its final part several miscellaneous topics that obviously belong to the *mubawwab* type, such as certain morphological patterns, alliteration (*itbāʿ*), metathesis (*qalb*), Arabized words (*muʿarrab*), and lists of words that pertain to a specific semantic field (food, bows and arrows, blades, and footwear).

The interrelatedness between *mubawwab* and *muğannas* lexica is further demonstrated by the fact that authors of each type incorporated into their corpus lexical items derived from the other. In the introduction to his *muğannas* lexicon titled *Tahdīb al-luġa*, the fourth-/tenth-century lexicographer al-Azhari mentions among his sources several works of the *mubawwab* type, such as works on *ġarīb*.⁴ Conversely, Ibn Sida incorporates into his *mubawwab* lexicon, *al-Muḥaṣṣaṣ*, a vast body of material derived from the *muğannas* lexica *Kitāb al-ʿAyn* and *Ġamhart al-luġa*. It is remarkable that *mubawwab* and *muğannas* dictionaries coexisted throughout the tradition. This is obviously due to the fact that each type serves a purpose not shared by the other. Ibn Sida, who authored a major *muğannas* lexicon, *al-Muḥkam*, as well as the voluminous *mubawwab* lexicon *al-Muḥaṣṣaṣ*, explicitly says in the introduction of the latter that because his other lexicon is intended to guide the searcher to the exact place in which any lexical item is recorded, he wished to match that with a *mubawwab* lexicon. He describes the latter as more appropriate for the orator (*ḥaṭīb*) and poet (*šāʿir*) since it provides them with the possibility of choosing from among the various synonyms and adjectives listed the *mot juste* that exactly suits their need in poetry and rhymed prose.⁵

In the rest of this chapter, we will survey the most important topics of the *mubawwab* lexica and trace their historical development, and then turn our attention to the *muğannas* lexica.⁶ The number of the *mubawwab* lexica in the tradition exceeds by far the number of the *muğannas* ones, but whereas the latter are mostly exhaustive of Arabic roots and are thus multivolumed, the former mainly consist of short treatises. Accordingly, many scholars authored

⁴ Azhari, *Tahdīb*, I.13. ⁵ Ibn Sida, *Muḥaṣṣaṣ*, I.10.

⁶ For a full discussion of both types, see Baalbaki, *Arabic Lexicographical Tradition*, chapters 2 and 3. See also Haywood, *Arabic Lexicography*, and Naṣṣār, *al-Muğam al-ʿArabi*, for a survey of the tradition.

several *mubawwab* works, but only few authors wrote more than one *muğannas* lexicon.

Mubawwab (Onomasiological) Lexica

As we shall see, lexica of this type, in which meaning leads to sign, mostly deal with a single theme, although there are also a few that are multi-thematic. In the early period of lexicographical activity, short treatises that specialize in one topic are clearly the result of the process of data collection referred to above. Some scholars have argued that the beginnings of the lexicographical tradition witnessed three stages. In the first stage, *mubawwab* lexica totally lacked organization, and were lists of semantically unconnected words that had been collected from the speech of Aʿrāb. In the second stage, words were classified under single topics (e.g. plants, animals, rain, etc.) and then in multi-thematic works, but only in the third stage did authors arrange lexical items on the basis of form (*lafẓ*); hence the emergence of the *muğannas* lexica.⁷ Although this proposed sequence might sound logical, the chronology inferred from the sources does not support it in the least. In fact, the process of data collection, as previously mentioned, continued in the first half of the third/second half of the ninth century with scholars such as al-Māzinī and Abū Ḥātim al-Siğistānī. More importantly, both al-Ḥalīl's *Kitāb al-ʿAyn* and al-Šaybānī's *Kitāb al-Ğīm* are of the *muğannas* type and were authored well before the data collection was complete. Furthermore, Abū ʿUbayd's multi-thematic dictionary *al-Ğarīb al-muṣannaf* appeared before the completion of data collection. Hence, single-topic wordlists, multi-thematic thesauri, and *muğannas* lexica represent concurrent types of lexical writing, and not consecutive stages in the tradition.

Mubawwab lexica can be divided into nine genres: strange usage (*al-ğarīb*); proverbs (*al-amṭāl*); specialized semantic fields; Arabized words (*al-muʿarrab*); solecisms (*laḥn al-ʾamma*); semantic phenomena; particles/letters (*al-ḥurūf/al-aṣwāt*); morphological patterns (*al-abniya*); and multi-thematic works.

Strange Usage (al-Ğarīb)

Interest in lexical items considered to be *ğarīb* – also referred to as *nādir* (rare), *ḥuṣṣī* (uncouth), and *šādd* (anomalous) – characterized the era of data collection and continued throughout the tradition. Among the earliest works in

⁷ Amīn, *Ḍuḥā l-Islām*, 2.263ff. Cf. Marçais, *Articles*, 149–51; Ğurāb, *al-Maʿāğim al-ʿArabiyya*, 48–50.

this genre are *al-Nawādir fī l-luġa* by Abū Zayd al-Anṣārī and *al-Nawādir* by Abū Miṣḥal al-Aʿrābī, both of the early third/mid ninth century. Both works largely lack organization and consist of a vast amount of reported data, mainly anomalous forms and expressions, dialectical variations, synonyms, and homonyms. Several developments took place in the genre, most notably in Ibn Ḥālawayhi's book *Laysa fī kalām al-ʿArab* (literally 'There is not in the speech of the Arabs [except] . . .'). It consists of 188 short chapters, each of which is devoted to a morphological pattern or a specific phenomenon which is so rare that it is attested only in a very small number of words.

There are also two subgenres of *ġarīb* lexica, the first of which deals with *ġarīb* words in the Qurʾān. A number of words in the Qurʾān were not readily understood by some Arabs, and scholars often differed in explaining them or attributed them to foreign origin. Yet, in addition to such words, books on the Qurʾānic *ġarīb* cite words that hardly qualify to be in this category. Examples from the earliest extant source, *Ġarīb al-Qurʾān wa-tafsīruhu* by Abū Miṣḥal al-Aʿrābī's contemporary al-Yazīdī, include fairly common words such as *raʿūf* (merciful), *ḥasra* (sorrow), *anṣaʿnā* (we started), *baġtatan* (suddenly), and *matīn* (tough).⁸ The reason for their inclusion is probably to specify their exact meaning within a Qurʾānic context. Early works of this subgenre were arranged according to the order of Qurʾānic chapters and verses. Full alphabetical order was later introduced by al-Harawī, whose book *Kitāb al-Ġarībayn* combines the *ġarīb* of the Qurʾān and of the *Ḥadīth* (Prophetic tradition). The second subgenre deals with the *Ḥadīth*, and the earliest book authored on the subject, Abū ʿUbayd's *Ġarīb al-Ḥadīth*, had a huge impact on later authors. Interestingly, the late third-/ninth-century lexicographer al-Ḥarbī, in his voluminous book *Ġarīb al-Ḥadīth*, probably influenced by al-Ḥalīl's *Kitāb al-Ayn*, adopts the system of permutations, in which the various combinations of the radicals of a root are listed together (see below).

Proverbs (al-Amṭāl)

This genre of writing was among the earliest to appear in the tradition. Some sources even attribute books on *amṭāl* to scholars of the first/seventh century, although it is impossible to verify the authenticity of such attribution. The unique structure and syntactic idiosyncrasies of *amṭāl*, in addition to their abundance in pre-Islamic fables and narratives – e.g. *Ayyām al-ʿArab* ('Battle-days of the Arabs') – made them an appealing subject to the philologists. The earliest extant works of the genre precede the third/ninth century.

⁸ Yazīdī, *Ġarīb*, 83, 86, 134, 136, 154.

Two examples of these very early works are al-Mufaḍḍal al-Ḍabbī's *Amtāl al-ʿArab* and Muʿarriḡ al-Sadūsī's *al-Amtāl*. Shortly after that, Abū ʿUbayd authored *al-Amtāl*, which includes 1,386 proverbs arranged thematically. Alphabetical ordering, however, was introduced relatively late, in the third/ninth century, by Ḥamza al-Iṣfahānī in *al-Durra al-fāḥira*. But the most famous work was compiled a century and a half later: al-Maydānī's *Mağmaʿ al-amtāl*, which includes 4,765 proverbs. There are also more specialized works, such as ones including proverbs that begin with the comparative adjective of the pattern *afʿal* ('more than . . .'); proverbs derived from the Qurʾān, the *Ḥadīṭ*, and poetry; and proverbs that are classified as *muwallad* (post-classical) or *ʿammī* (vernacular).⁹

Specialized Semantic Fields

There are numerous wordlists that contain vocabulary specific to a narrow semantic field. Many of these were authored in the third/ninth century by philologists who collected linguistic material from the Bedouin, and although scholars continued gathering wordlists of this type as late as the eleventh/seventeenth century, they could add only little material to that of their predecessors. The early wordlists are particularly important because their material was incorporated not only in multi-thematic works or thesauri, but also in the *muğannas* lexica.

The semantic fields in these wordlists include plants (*nabāt*) in general, trees (*šağar*), kinds of herbage (*kalaʿ*), palm trees (*naḥl*), vineyards (*karm*), animals (*ḥayawān*) in general, camels (*ibil*), sheep (*šāʿ*), wild animals (*wuḥūš*), horses (*ḥayl*), lions (*asad*), wolves (*diʿb*), bees (*naḥl*), insects (*ḥašarāt*), the human body (*ḥalq al-insān*), rain (*maṭar*), wind (*rīḥ*), clouds (*saḥāb*), weapons (*silāḥ*), wells (*biʿr*), calamities (*dawāhī*), gambling (*maysir*), wine (*ḥamr*), and so on. The wealth of vocabulary in these wordlists is matched by the wealth of attested usage (*šawāhid*), mainly from pre-Islamic and early Islamic poetry, which is quoted to confirm the authenticity of the words under discussion.

Arabized Words (al-Muʿarrab)

Early grammarians and lexicographers identified many words as loaned from other languages. Several works, such as Sībawayhi's grammar book *al-Kitāb* and Ibn Durayd's semasiological lexicon *Ġamharat al-luġa*, include small chapters on Arabized words.¹⁰ Writing in the sixth/twelfth century, al-

⁹ See Sellheim, 'Mathal', for further detail on works devoted to proverbs.

¹⁰ Sībawayhi, *Kitāb*, 4.303–7; Ibn Durayd, *Gamhara*, 3.1322–7, 1329.

Ġawālīqī was probably the first author to incorporate earlier data into a book which exclusively deals with Arabized words. His *al-Mu‘arrab min al-kalām al-a‘ġamī* embraces 743 words alphabetically arranged on each word’s first letter only, and was hardly surpassed by later authors. Words are typically traced to one language (for instance Persian, Syriac, or Greek), although at times more than one source language is proposed, or the loanword is non-committally described as *a‘ġamī* (foreign).

Solecisms (Laħn al-‘Āmma)

Cited in the tradition as having triggered an awareness of the need for establishing criteria for correct usage, *laħn al-‘āmma* (literally ‘errors of the commoners’) were the subject of numerous monographs authored as early as al-Kisā’ī’s late second/late eighth or early ninth-century *Mā talħan fīhi l-‘āmma*. The use of the term *laħn al-‘āmma*, however, is often misleading, since numerous works not only cite linguistic errors that allegedly occur in the everyday speech of the generality of people, but also errors committed by men of letters and eminent scribes. Al-Ĥarīrī’s *Durrat al-ġawwās fī awhām al-ħawāšš* is an example of a work that is devoted to the errors of the *ħawāšš*, in other words the elite or people of distinction. Two types of arrangement were followed in this genre: the alphabetical (which appeared relatively late in the sixth/twelfth century) and the thematic. The latter arrangement was more frequently used perhaps because most authors wanted to highlight the types of changes involved when a supposedly correct usage is altered – an objective that an alphabetical arrangement cannot achieve. Most of the errors cited are due to changes in vowels and consonants and in morphological patterns, as well as semantic changes that are deemed unacceptable. Syntactic errors are less frequent, given that they were thoroughly discussed in grammatical works.

Semantic Phenomena

In the introductory part (*risāla*) of *al-Kitāb*, Sībawayhi briefly mentions a number of semantic relationships, including homonymy and synonymy.¹¹ But these do not feature elsewhere in the book nor are they used as analytical tools by Sībawayhi or elsewhere in the grammatical tradition, which focused almost exclusively on syntax, morphology and morphophonology, and to a lesser extent phonology. Semantic issues were, however, discussed by lexicographers as early as the beginning of the third/ninth century and

¹¹ Sībawayhi, *Kitāb*, I.24.

there emerged three distinct themes, namely, *al-aḍḍād* (words with two contradictory meanings), *al-muštarak* (homonyms), and *al-mutarādif* (synonyms).

For several centuries, scholars were in disagreement as to whether one word can express two contradictory meanings or not. The eminent lexicographer Ibn Fāris asserts that this phenomenon is one of the norms (*sunan*) of the speech of the Arabs.¹² Many of the authors who argued against the existence of the *aḍḍād*, on the other hand, probably did so in reaction to the claims of the *Šuʿūbiyyūn* (i.e., those who exalt mainly Persian culture over Arabic culture) that the existence of the *aḍḍād* is indicative of the Arabs' lack of wisdom and abundance of ambiguity in their dialogue. One such author is Abū Bakr ibn al-Anbārī, whose book *al-Aḍḍād* includes 357 items, the largest collection in the genre. His main argument is that the context would eliminate any ambiguity in words claimed to have two contradictory meanings.¹³ Yet the refusal to admit the existence of the *aḍḍād* was not restricted to Arab authors. For instance, Ibn Durustawayhi, who was of Iranian descent, mentions in one of his works that he had authored a book titled *Ibtāl al-aḍḍād* (*The Annulment of the aḍḍād*).¹⁴

Homonyms which do not indicate contradictory meanings (*al-muštarak*) were the subject of several short wordlists in the third/ninth century. However, Kurāʿ al-Naml's book, *al-Munaḡḡad fī l-luḡa*, is considerably larger than earlier works. It includes a total of 884 words arranged partly thematically and partly alphabetically on the basis of the first two letters of the word, and not its root.¹⁵ By the sixth/twelfth century, the size of the material grew appreciably, and Ibn al-Šaḡarī's *Mā ttafaqa lafẓuhu wa-ḥtalafa ma'nāhu* ('What agrees in form and differs in meaning') embraces 1,670 words of the *muštarak* type, arranged alphabetically based on the first letter only.

Wordlists of the *mutarādif* genre resemble semasiological lexica in giving synonyms of lexemes, but whereas the former typically cite several synonyms for each word, the latter normally cite only one and often give an explanation for the word rather than a synonym. Most works of this genre follow a thematic, not an alphabetical, arrangement. The third-/ninth-century Kufan lexicographer Ibn al-Sikkīt's *al-Alfāz*, as preserved in a fifth-/eleventh-century commentary, al-Tibrīzī's *Kanz al-ḥuffāz fī Kitāb Tahdīb al-Alfāz*, is probably the most prominent work of the genre. It includes 148

¹² Ibn Fāris, *Šāḥibī*, 97. For the various views of Arab scholars on *al-aḍḍād*, see Suyūṭī, *Muzhir*, 1.387–402.

¹³ Ibn al-Anbārī, *Addād*, 1. ¹⁴ Ibn Durustawayhi, *Taṣṣīḥ*, 1.359.

¹⁵ See Omer, 'Early Arabic lexicons', 5–6.

mostly short chapters in which are listed words and expressions pertaining to a specific field, such as wealth, poverty, battalions, stinginess, illness, wine, colour, old women, death, love, the sun, the moon, and so on. In the chapter on cowardliness, for example, several adjectives and phrases denoting the cowardly man are listed: *ḡabān*, *yarā'a*, *manḥūb*, *maf'ūd*, *wahil*, *iḡfīl*, *hawā'*, *ra'īb*, *bayyin al-ri'dīda*, *aḡbanu min ṣāfir*, and others.¹⁶ Accordingly, Ibn al-Sikkīt's book – which has a strong component of *mutarādīf* words – can also be classified under multi-thematic works (see below).

As with *al-aḏḏād*, several authors denied the very existence of *al-muṣṭarak* and *al-mutarādīf*.¹⁷ The principal argument they adduced was that the admission of the existence of these two types was incompatible with the principle of *ḥikma* (wisdom) that was usually ascribed to the inventor of the language (*wāḏī' al-luḡa*). In *al-mutarādīf*, this is clear in the first chapter of al-ʿAskarī's *al-Furūq al-luḡawiyya* ('Linguistic/semantic differences'), one of a number of books that strive to explain the nuances of words that are claimed to be synonymous, thus denying their synonymy.¹⁸ One example is the differentiation between *qirā'a* and *tīlāwa*, for whereas the former can apply to reading a single word, the latter applies only to two or more words given that its root, *talā* (to follow), implies more than one element.¹⁹

Particles/Letters (al-Ḥurūf/al-Aṣwāt)

One of the major themes in this genre is the study of particles or *ḥurūf al-ma'ānī* (literally 'letters with meanings'), a term that refers mainly to particles that grammatically govern the verb or the noun, in contrast to the term *ḥurūf al-ḥiḡā'* or letters of the alphabet, which obviously have no grammatical significance. The material in wordlists of this type is the same as in grammatical works, but is normally arranged in the form of a lexicon. In some works, the particles are alphabetically arranged regardless of the number of their letters – for instance al-Mālaqī's *Raṣf al-mabānī fī ṣarḥ ḥurūf al-ma'ānī* – or according to that number, which ranges from one to five – for instance al-Murādī's *al-Ḡanā l-dānī fī ḥurūf al-ma'ānī*.

There are also wordlists that deal with specific letters of the alphabet. Due to orthographic and dialectal factors, words with a glottal stop (*hamza*) were collected or classified in several works. The letter *lām* (l) was also a subject of interest, mostly at the syntactical level, as in al-Zaḡḡāḡī's monograph *al-Lāmāt*, in which thirty-one types of *lām* are identified. But it is with the two

¹⁶ Ibn al-Sikkīt, *Alfāz*, in Tibrīzī, *Kanz*, 176–83.

¹⁷ See Suyūṭī, *Muzhir*, 1.369–86, 402–13, for the various views on both types.

¹⁸ ʿAskarī, *Furūq*, 11. ¹⁹ ʿAskarī, *Furūq*, 48.

letters *ḍād* (ḍ) and *zā'* (z) that most wordlists in this genre deal. These letters were the source of confusion both in pronunciation and dictation, and wordlists typically point out pairs of words distinguished only by the alternation of ḍ and z, such as *'ida* 'a great tree with thorns' and *'iza* 'exhortation'.²⁰ Interestingly, one of the most extensive works on *ḍād* and *zā'*, al-Ġarbādaqānī's *al-Rawḥa*, follows a phonetic arrangement of roots – a rare occurrence in onomasiological lexica.

A number of works deal with other characteristics of letters, such as assimilation (*idḡām*), consonantal substitution (*ibdāl*), and metathesis (*qalb*). One example is Abū l-Ṭayyib al-Luḡawī's *al-Ibdāl*, whose extensive material is arranged into chapters that follow strict alphabetical order.

Morphological Patterns (al-Abniya)

Closely related to grammar, of which morphology constitutes a major part, morphological themes that were deemed worthy of study in independent wordlists include *al-iṣṭiqāq* (derivation of proper nouns or attributes of God), *al-muḍakkār wa-l-mu'annat* (masculine and feminine), and *al-maqṣūr wa-l-mamdūd* (abbreviated and prolonged patterns, ending with *-ā* or *-ā'* respectively). Wordlists on these subjects go back to the early third/ninth century and vary in the arrangement of their material, but some are alphabetically arranged. For instance, the first alphabetically arranged work, albeit on the basis of only the first letter, in the realm of the *maqṣūr* and *mamdūd*, is Ibn Wallād's *al-Maqṣūr wa-l-mamdūd*. Another type of wordlist deals with *muṭallatāt* (triplets). These are groups of three words whose roots and patterns are identical, but one of whose radicals has an *-a-*, *-i-* or *-u-* vowel (*faṭḥa*, *kasra* and *ḍamma* respectively), with a change in meaning in most cases, as in *ṭalā* (fawn), *ṭilā* (thick beverage), and *ṭulā* (necks).²¹ The first extant work on triplets was authored by Sībawayhi's student Quṭrub, and includes thirty-two triplets listed in no order. Later authors considerably expanded the material and arranged it alphabetically, as did the sixth-/twelfth-century al-Baṭalyawsī, whose *al-Muṭallat*, according to his own enumeration, embraces 833 triplets.²²

More extensive works were authored in which nominal and/or verbal patterns were listed and explained. Many of these would have qualified to be semasiological lexica had it not been that, by focusing on morphological patterns, they were not intended to include all Arabic roots. The first author to combine nominal and verbal patterns in one book is the fourth-/tenth-century al-Fārābī, whose *Dīwān al-adab* is meticulously divided into sections

²⁰ Ibn Mālik, *I'timād*, 47. ²¹ Quṭrub, *Muṭallatāt*, 41–2. ²² Baṭalyawsī, *Muṭallat*, 1.298–9.

and subsections based on morphological considerations. It follows the rhyme system in the arrangement of roots (see below). Two of the most important lexica devoted solely to verbal patterns were authored by the Andalusian scholars al-Saraqustī and Ibn al-Qaṭṭā' and titled *al-Af'āl*. Their arrangement on morphological grounds, however, makes them quite difficult to use.

Multi-thematic Works

The first extant multi-thematic work belongs to the third/ninth century: Abū 'Ubayd's *al-Ġarīb al-muṣunnaf*. Its early date, which precedes the completion of data collection, shows that this genre began soon after the single-topic wordlists started to appear. It may even be argued that, based on their titles, several works that are lost to us and are attributed to authors in the second half of the second/eighth century may have been multi-thematic in nature. In fact, throughout much of the tradition, authors of multi-thematic works often drew on material in single-topic wordlists, and vice versa. Furthermore, the classification of some lexica as single-topic or multi-thematic works is problematic. One such case is *al-Munağğād* (see above). Although its partial thematic division into such categories as the human body, animals, weapons, and the sky is reminiscent of multi-thematic thesauri, the primary occupation of its author with homonyms (*al-muṣṭarak*) firmly places it within that genre.

Al-Ġarīb al-muṣunnaf contains twenty-six sections, elaborately divided into about 900 (mostly short) chapters. Yet its material is but a fraction of the most extensive opus in the genre, compiled two centuries later, Ibn Sīdā's *al-Muḥaṣṣaṣ*. This contains two main parts: the general themes and the philological issues. Al-Ṭālibī identifies four major areas into which the themes may be classified, namely, human beings and their environment, animals, nature and plants, and man in society.²³ The philological part embraces a host of issues, such as duals and plurals, particles, verbal and nominal patterns, masculine and feminine, *al-aḍḍād*, etc. With *al-Muḥaṣṣaṣ*, the genre reached its peak and, given that later authors had precious little to add to its vast material, authoring thesauri of this magnitude practically came to an end.

Muğannas (Semasiological) Lexica

Unlike the specialized lexica and the multi-thematic thesauri of the *mubawwab* type, in which meaning leads to sign, *muğannas* lexica are typically

²³ Al-Ṭālibī, *al-Muḥaṣṣaṣ*, 70–2.

unspecialized and arranged based on formal criteria, whereby sign leads to meaning. Another major characteristic of *muğannas* lexica is that they usually try to be exhaustive of all the roots of the language and, in some of the more extensive works, the largest possible number of their derivatives. Based on their formal criteria of arrangement, these lexica are normally classified into three types: the phonetic-permutative, the alphabetical, and the rhyme-based. (S. I. Sara's attempt at identifying subcategories within this classification and his inclusion in it of dictionaries of the modern era is both confusing and anachronistic.)²⁴ In this part, we shall examine the emergence and nature of each of the three types of the *muğannas* lexica and then survey the main works, arranged purely chronologically, irrespective of the type to which they belong.²⁵ The survey is limited to the philologically oriented lexica, to the exclusion of lexica that deal with other subjects, such as toponyms, philosophical or medical terms, names of drugs, or biographies of various types.

The originator of the phonetic-permutative system is almost certainly al-Ḥalīl b. Aḥmad, who died in the late second/eighth century. His introduction to *Kitāb al-'Ayn*, the first lexicon of this type, lays the phonetic and morphological foundations of its arrangement. Based on their points of articulation, which al-Ḥalīl examined through *ḍawāq* (sampling of letters), the letters of the alphabet were arranged in the following order:

‘, ḥ, h, ḥ, ġ, q, k, ġ, š, ḍ, š, s, z, t, d, t, z,
ḍ, t, r, l, n, f, b, m, w, alif, y, hamza

Two other arrangements were available to al-Ḥalīl but, unlike the phonetic arrangement he adopted, both lacked any linguistic justification. The first of these is the ‘, b, ġ, d, h, w, z, etc. order, which is of foreign origin and used in calculation (*ḥisāb al-ğummal*). The other is the ‘, b, t, t, ġ, h, ḥ, etc. arrangement, which is derived from the previous arrangement primarily on the basis of the form of the Arabic letters.

Given that the letters of the alphabet form a closed set of twenty-nine elements, and that the number of radicals in Arabic words ranges from two to five, al-Ḥalīl established that the roots are finite in number. To these two finite sets he then applied the principle of permutations, known in later sources as *taqālīb*. The number of possible permutations is 2 in biliterals (for instance ‘K and K’), 6 in trilaterals (*KSD*, *KDS*, *DSK*, *DKS*, *SDK*, *SKD*), and

²⁴ Sara, ‘Classical Arabic lexicographical tradition’.

²⁵ For a historical arrangement of these works within each of the three types, see Baalbaki, *Arabic Lexicographical Tradition*, 279–401.

theoretically 24 and 120 in quadrilaterals and quinqueliterals respectively. Only the combinations which actually occur (*musta'mal*), of course, are explained in the lexicon (for instance the first five permutations of *KSD*); they are cited together in the lemma of the root that precedes the other roots of its group in the arrangement. Accordingly, no Arabic root, whether used or not, could escape al-Ḥalīl's count. To exhaust the derivatives of these roots, however, was more challenging due to the vast material involved, and later authors continually expanded their material from this perspective.

In the alphabetical system, the letters are arranged as follows:

ʾ, b, t, ṭ, ġ, ḥ, ḏ, d, r, z, s, š, ṣ, ḏ, t, z,
, ġ, f, q, k, l, m, n, h, w, y (at times w, h, y)

Authors of *mubawwab* and *muğannas* lexica often justify their choice of this arrangement with reference to its easiness and familiarity to the generality of people.²⁶ Among extant *muğannas* lexica, the early fourth-/tenth-century Ibn Durayd's *Ġamharat al-luġa* was the first in which the alphabetical system was applied to all radicals of the root. In *mubawwab* lexica, partial application of the system (by the first radical only or by the first two radicals) continued through much of the tradition, side by side with its full application.

Kitāb al-Ġīm by Abū 'Amr al-Šaybānī, who lived a century before Ibn Durayd, is generally recognized as the first *muğannas* lexicon to adopt the alphabetical arrangement, albeit on first-letter basis only. This lexicon is not exhaustive of all Arabic roots since it is primarily concerned with lexical items of the *ġarīb* type. Yet its inclusion with *muğannas* lexica – which normally, though not always, include all roots – is essential for demonstrating that alphabetical ordering appeared contemporaneously with the phonetic-permutative system. The death dates of al-Ḥalīl (175/791) and Abū 'Amr (206/821) suggest that the former's *al-'Ayn* preceded the latter's *al-Ġīm*. However, these dates may be misleading in determining which lexicon was authored first since, according to some reports, Abū 'Amr lived to the age of 90 (or even 119 in other reports).²⁷ If one accepts the death date of 206/821, which is the one most commonly given in the sources, then, given his long life span, there is a very good chance that he authored his lexicon simultaneously with *al-'Ayn* or even earlier.

The third type of classification is the rhyme system. This is an alphabetical system but is set apart from the previous one because it starts with the last

²⁶ See, for instance, Ibn Durayd, *Ġamhara*, 1.40, and Ibn Wallād, *Maqṣūr*, 2–3.

²⁷ Cf. Ibn al-Anbārī, *Nuzha*, 80.

radical of the root. The other radicals are then arranged from first to intermediate. Accordingly, instead of 1+2+3 and 1+2+3+4 in trilaterals and quadrilaterals respectively, the order becomes 3+1+2 and 4+1+2+3. Al-Ġawharī perfected the rhyme system in *al-Šiḥāḥ*, and it was adopted by the majority of the extensive *muğannas* lexica authored after him.

Yet al-Ġawharī is not the originator of this system. His main contribution is that he followed a single criterion in the arrangement of roots, namely, the order of the radicals, without any consideration for nominal or verbal patterns. This is a huge step forward compared to *Dīwān al-adab* – authored by his maternal uncle al-Fārābī – in which the roots are indeed arranged based on the order of all their radicals starting with the last, but are listed in numerous chapters, each of which is devoted to a specific morphological pattern, an arrangement which would be followed by Maḥmūd al-Kāšġarī in his *Dīwān Luġāt at-Turk* (see Chapter 11). Several decades earlier, al-Bandanīġī had arranged the lexical items (not the roots) in *al-Taqqiya fī l-luġa* based on their final radicals. Accordingly, he is credited with inventing the rhyme system. Like *al-Ġīm*, which is the first lexicon to follow the alphabetical system, *al-Taqqiya* is neither intended to be exhaustive nor takes more than one radical (in this case, the last) into consideration in the arrangement.

Baalbaki has recently demonstrated that the use of the rhyme system precedes *al-Taqqiya*.²⁸ Although part of a famous work titled *Islāḥ al-manṭiq* by Ibn al-Sikkīt (who died in 244/858, four decades before al-Bandanīġī), the lengthy chapter on the two patterns *faʿl* and *faʿal*,²⁹ which is arranged based on the final letters of the words listed, has curiously escaped the notice of modern scholarship. The order of the letters in this chapter is probably haphazard, but words that end with the same letter are exclusively placed together (for instance *nazḥ*, *ṭarḥ*, *faḥ*, *ṭallḥ*, *ṣabḥ*, *ṣarḥ*, *nadḥ*, *qarḥ*, and *qarn*, *ġabn*, *ḥazn*, *ʿaġn*, *fann*, *sann*, *saḥn*). It is very likely that al-Bandanīġī devised *al-Taqqiya* based on Ibn al-Sikkīt's chapter by incorporating and expanding its material, largely preserving the order of its lexical items, but arranging the chapters according to the ʿ, b, t, ṭ, etc. order. Our knowledge of the early works that adopted the rhyme system, however, is far from complete. What can be established with confidence is that this system was more widely used as of the third/ninth century than had been hitherto acknowledged. Strongly pointing in that direction is that al-Bandanīġī's close contemporary al-Dīnawarī, in his lexicon on plants, *Kitāb al-Nabāt*, part of which is

²⁸ Baalbaki, *Arabic Lexicographical Tradition*, 364–6; Baalbaki, 'Precursor', 5–14.

²⁹ Ibn al-Sikkīt, *Islāḥ*, 37–84.

alphabetically arranged, expresses his preference for the alphabetical arrangement over the arrangement by final letters.³⁰ This striking statement implies that the latter type of arrangement was sufficiently common to rival the alphabetical arrangement by first letters as a model for lexicographers.

The question arises as to why lexicographers would even consider starting with the final radical in arranging roots. The most satisfactory explanation is that this facilitates poets' search for suitable rhymes.³¹ A direct link between the rhyme system and poetry is made by the unidentified disciple of al-Bandanīgī who wrote *al-Taḡfiya*'s introduction. He explains that the lexicon is titled *al-Taḡfiya* (literally 'Rhyming') because it is based on *qawāfiʿ*, plural of *qāfiya*, which he defines as a line of poetry.³² The link with poetry is further confirmed by the fact that al-Bandanīgī cites only lexical items, but not roots (which would have been useless for poets) and that he does not distinguish within the same group of words between geminated (e.g. *ṣadd*, *kadd*, *madd*) and ungeminated ones (e.g. *ṣald*, *mahd*, *ṣahd*), since both types can occur as last words in the same poem.³³

The following purely chronological arrangement of the most important extant *muḡannas* works demonstrates how the three above-mentioned types of arrangement coexisted for several centuries, following which the rhyme system became the system of arrangement *par excellence* in the major lexica. Due to space constraints, the commentaries, abridgements, and expansions to which most of these works gave rise will not be discussed.

1. *Kitāb al-ʿAyn* by al-Ḥalīl bin Aḥmad (d. 175/791). In many respects, this lexicon laid the foundations of the lexicographical tradition as a whole. Its phonetic arrangement and method of citing permutations were entirely adopted by several authors as late as Ibn Sīda (see item 10). Even authors who gave up on al-Ḥalīl's phonetic arrangement in favour of the alphabetical one still preserved his system of permutations (e.g. Ibn Durayd; see item 4) or of dividing the material into chapters based on the number of radicals in words (e.g. Ibn Fāris; see item 8).

Doubts were raised as of the third/ninth century about the identity of *al-ʿAyn*'s author, primarily due to alleged mistakes whose ascription to al-Ḥalīl was deemed to be incompatible with his scholarly status. But even if al-Ḥalīl were not the true author of *al-ʿAyn* in full or in

³⁰ Dīnawarī, *Nabāt*, 397.

³¹ For other suggestions, see 'Umar's introduction to Fārābī, *Dīwān al-adab*, 1.18–19, and Carter, 'Arabic lexicography', 110–11. Both are refuted by Baalbaki, 'Influence', 23–4.

³² Bandanīgī, *Taḡfiya*, 36. ³³ Bandanīgī, *Taḡfiya*, 306–7.

part – although there is strong evidence to the contrary, such as its grammatical material, which is consistent with what Sibawayhi attributes to al-Ḥalīl in *al-Kitāb* – most sources agree that the lexicon must at least have been authored shortly after his death by a disciple of his, such as al-Layṭ bin al-Muẓaffar.³⁴ *Al-ʿAyn*’s lengthy introduction, however, obviously bears al-Ḥalīl’s mark. Written by an unidentified disciple of his, this valuable document sets forth al-Ḥalīl’s phonetic and morphological views which are reflected in the lexicon’s divisions and arrangement of roots. The lemmata are replete with attested data (*ṣawāhid*), but the order in which the root’s derivatives are mentioned does not follow a rigorous system.

2. *Kitāb al-Ġīm* by al-Šaybānī (d. 206/821). Previously mentioned as the first *muğannas* lexicon to adopt alphabetical order, albeit on the basis of the first letter only, its author was primarily interested in recording dialectal material and poetry of several tribes – hence its strong *ġarīb* component. Although titled *al-Ġīm*, the lexicon does not start with the letter ġ (*ġīm*), unlike *al-ʿAyn* which indeed begins with that letter. The reason for its name remains a mystery.
3. *Al-Taḳfiya fī l-luġa* by al-Bandanīġī (d. 284/897). The arrangement and scope of this lexicon are discussed above.
4. *Ġamharat al-luġa* by Ibn Durayd (d. 321/933). This lexicon is unique in that it adopts an alphabetical arrangement based on all the letters of the root and, at the same time, retains al-Ḥalīl’s permutations. Its major divisions are a function of the number of radicals of roots (biliteral, trilateral, etc.), and each division is then alphabetically arranged independently of the others. As Ibn Durayd explains in the introduction, the title of his lexicon reflects his interest in recording the *ġumhūr* (the main body, the bulk) of the speech of the Arabs to the exclusion of what is *waḥṣī* (uncouth, barbarous) and *mustankar* (objectionable).³⁵ Yet Ibn Durayd’s inclusion of several non-standard dialects, such as the Azdī and Šaʿamī dialects, gave his opponents reason to accuse him of including items of questionable authenticity.
5. *Al-Bāriʿ fī l-luġa* by al-Qālī (d. 356/967). Although this lexicon – the first *muğannas* lexicon in Arab Spain – is heavily influenced by *al-ʿAyn*, it has certain characteristics of its own. The phonetic ordering of the letters differs from al-Ḥalīl’s in some respects, and several modifications were introduced into the arrangement of the material based mostly on

³⁴ For the evidence for his authorship, see Talmon, *Arabic Grammar*, 256–9.

³⁵ Ibn Durayd, *Ġamhara*, 1.41.

morphological grounds. Judging by its extant parts, the lexicon's material is quite extensive and it is not surprising that lexical items are regularly ascribed to the early philologists who reported them given that the period of data collection had practically come to an end by al-Qālī's time.

6. *Tahdīb al-luġa* by al-Azharī (d. 370/980). The author follows al-'Ayn's system very closely but his material, spread over fifteen volumes, dwarfs that of al-'Ayn. Al-Azharī is critical of several of his predecessors in his introduction, and to avoid their pitfalls he was extremely meticulous in checking the data available in the sources. Obviously, he was keen to highlight his critical approach by his choice of the word *Tahdīb* for the title (literally 'Correction'). To his credit is that, during several years in captivity, he heard the 'pure' speech of the eloquent Bedouins, who by his time were difficult to come by. He thus boasts the fact that he admits only material that is undoubtedly authentic because it derives from reliable sources.
7. *Al-Muḥīṭ fī l-luġa* by al-Šāḥib bin 'Abbād (d. 385/995). Like *Tahdīb al-luġa*, this lexicon faithfully follows al-Ḥalīl's system of arrangement but considerably expands its material. Al-Šāḥib frequently indicates that a certain root was considered by al-Ḥalīl to be unused (*muḥmal*) but adduces usage that proves the contrary. It is remarkable that al-Šāḥib opted to emulate al-'Ayn's phonetic-permutative system although he was a student of Ibn Fāris, author of two alphabetically arranged lexica (see item 8).
8. *Maqāyīs al-luġa* by Ibn Fāris (d. 395/1004). This lexicon is unique in two major aspects: its arrangement and its identification of *uṣūl*. The author follows al-Ḥalīl in considering root length a criterion in the arrangement of material but opts for an alphabetical arrangement of roots based on all their radicals. The unique feature in this arrangement is that it envisages the letters of the alphabet as successive points on a circle. Hence, after the first letter of the root, the letter that immediately follows it in the alphabet is considered, and then the next, until the last letter of the alphabet. Only then are the letters that precede that first letter in alphabetical order considered. The other unique feature is the application of the notion of *uṣūl* (pl. of *aṣl*) or basic meanings associated with the roots. This is uniformly applied in biliterals and trilaterals. For example, Ibn Fāris identifies four *uṣūl* for the root ŠKR, each of which is embodied in one or more words derived from it. These *uṣūl* indicate thanking, fullness, a type of plant, and taking a spouse.³⁶ Ibn Fāris' other lexicon, *Muġmal al-luġa*, has the same arrangement as the *Maqāyīs*, but the

³⁶ Ibn Fāris, *Maqāyīs*, 3.207–8.

lemmata are not internally divided on the basis of *uṣūl*. The popularity of this lexicon stems from its conciseness and brevity of expression, as its title (literally ‘The Synopsized’) suggests.

9. *Al-Ṣiḥāḥ* by al-Ġawharī (d. c. 400/1010). Al-Ġawharī not only perfected the rhyme system (see above), making his lexicon in this respect a model that was emulated by lexicographers for several centuries, but was also renowned for restricting his material to what he deemed – albeit without clear criteria – to be correct or authentic, hence the lexicon’s title *al-Ṣiḥāḥ/al-Ṣaḥāḥ* (literally ‘Correct usage’). The lemmata reveal a somewhat more systematic organization of the root’s derivatives than in previous works (for instance, the perfect is normally followed by the imperfect, the verbal noun, and the assimilate adjective). As an aid to the reader, al-Ġawharī frequently specifies the correct vowel (for instance of the imperfect) or gives the word’s pattern to ensure that it is not misread.
10. *Al-Muḥkam* by Ibn Sīda (d. 458/1066). This is the last major *muḡannas* lexicon of the phonetic-permutative type, and Ibn Sīda proudly states that it is the most comprehensive lexicon ever authored.³⁷ Yet the real value of the work lies in its author’s unprecedented criticism of the way the definitions of his predecessors are formulated and in setting criteria for how the data should be presented and what the lemmata should include or exclude. Although Ibn Sīda did not rigorously apply these criteria in authoring *al-Muḥkam*, it is certainly one of the most organized and well-planned lexica in the entire tradition.
11. *Asās al-balāḡa* by al-Zamaḥṣarī (d. 538/1144). This is the first lexicon in which roots are arranged in full alphabetical order, but without the morphological divisions and permutations used in earlier alphabetically arranged lexica. It is also the first lexicon, and most probably the only one, that systematically distinguishes between literal or veridical (*ḥaqīqī*) and tropical (*maḡẓāzī*) usage. Interested primarily in the latter, al-Zamaḥṣarī focuses on context and normally cites words, not as isolated units, but as components of a structure that requires explanation.
12. *Al-‘Ubāb al-zāḥir* by al-Ṣaḡḡanī (d. 650/1252). Authored towards the end of a rich scholarly life and after several *muḡannas* and *mubawwab* lexica of his, *al-‘Ubāb* – which adopts the rhyme system – is among the most comprehensive Arabic lexica in spite of being largely free from digression and from reporting the conflicting views of earlier lexicographers. It is arguably the most organized lexicon as far as the internal arrangement of

³⁷ Ibn Sīda, *Muḥkam*, 1.48.

the lemmata is concerned. In a step that considerably facilitates the search for the required word, the ground form of the verb is followed by its nominal derivatives and each of the augmented verbs is then listed, followed by its nominal derivatives as well.

13. *Lisān al-ʿArab* by Ibn Manẓūr (d. 711/1311×1312). Embracing 9,273 roots (hence lemmata) arranged by last letter, *al-Lisān* is second only to *Tāǧ al-ʿarūs* (see item 15), which contains 11,978 roots.³⁸ Ibn Manẓūr incorporated into it five earlier lexica, including *Tahdīb al-luġa*, *al-Šihāḥ*, and *al-Muḥkam*. In the introduction, he defines two basic criteria for evaluating lexical works, namely, scope and easiness of use.³⁹ He rhymingly names these two criteria as *ġamʿ* and *waḍʿ* respectively and judges earlier works accordingly. But although he ascribed no credit to himself other than merging his five sources into one lexicon, he made a great effort to organize the content of the material by identifying semantically related lexical items and separating them from other groups of words within the same lemma.
14. *Al-Qāmūs al-muḥīṭ* by al-Firūzābādī (d. 817/1415).⁴⁰ So popular was this work, which also follows the rhyme system, that the word *qāmūs* (a loan from Greek *ōkeanós* used by the Arabs as a word for ‘sea’) has largely replaced *muḡam* to denote ‘lexicon’. Al-Firūzābādī notes that al-Ġawharī excluded from *al-Šihāḥ* half the language or more, and thus he wanted to make up for this deficiency by restoring the roots and words which he considered worthy of inclusion.⁴¹ To keep the work within reasonable limits, he adopted a template according to which the material is presented in a condensed and systematic manner. He also consistently applied certain abbreviations, such as *d* for *balad* (town), *ġ* for *ġamʿ* (plural), and *m* for *maʿrūf* (well-known).
15. *Tāǧ al-ʿarūs* by al-Zabīdī (d. 1205/1791). Intended to supplement al-Firūzābādī’s *al-Qāmūs* by adding material from 116 sources listed in the introduction, this rhyme-based lexicon crowns the ‘classical’ period of Arabic lexicography as the most comprehensive lexicon in the whole tradition. By restoring to the text of *al-Qāmūs* the names of the philologists to whom the data are ascribed, the *šawāhid* (attested material), and the missing roots, words and meanings, al-Zabīdī authored a truly encyclopedic work that contains not only linguistic data but also information derived from works which specialized in biographies, place-names, animals, proverbs, medicinal plants, stones, metals, and the like. Some

³⁸ Mūsā and Šahīn, *Dirāsa ihṣāʾiyya*, 9; Mūsā, ‘Dirāsa tiqniyya’, 149.

³⁹ Ibn Manẓūr, *Lisān*, Introduction, 1.8.

⁴⁰ For Firūzābādī, see Strotmann, *Majd al-Dīn al-Firūzābādī*.

⁴¹ Firūzābādī, *Qāmūs*, Introduction, 1.4.

Egyptian vernacular usage is also recorded (for instance in the entries *RW*, *ṢNT*, *LT*). It is also worth mentioning that Lane's voluminous English–Arabic lexicon (see Chapter 19) is primarily based on *al-Tāğ*.

Concluding Remarks

Arabic lexicography is best understood within the wider context of Islamic scholarship. In particular, there is close affinity between the beginning of grammatical and lexicographical writing, and both disciplines share a common background and are closely related to the study of the Qur'ān. Both *al-Kitāb* and *al-'Ayn* were authored in the same period, and each left its mark on its own discipline by laying its foundations at a very early stage of linguistic enquiry. With its roots firmly based in the Arabic-Islamic culture, Arabic lexicography is almost certainly the fruit of indigenous scholarly activity and not the result of direct borrowing from an available system that merely had to be Arabized. It is clear in *al-'Ayn*'s introduction, in which al-Ḥalīl's phonetic and morphological views are expounded, that he followed an intuitive approach in discovering the phonetic traits of Arabic and that the terms he used to describe Arabic sounds represent a fresh and experimental approach and are not translations of foreign terms.⁴² Parallels have indeed been drawn between the Arabic and Indian phonetic traditions in order to establish the influence of the latter on the former. But although both traditions adopt an arrangement that has an articulatory basis and proceeds at the source of the airstream and works outward, it has been convincingly demonstrated that this resemblance is merely superficial and that the fundamental concepts underlying the two approaches are substantially different.⁴³ Furthermore, it was suggested that the rhyme system of arrangement may have been influenced by a Hebrew dictionary known as the *Agron*, whose second part is arranged by the last letter.⁴⁴ But this is disproven by the fact that its author, Saadia Gaon, also known as Sa'īd b. Yūsuf al-Fayyūmī (for whom see Chapter 9), was born in 279/892, thirty-five years after Ibn al-Sikkīt's death in 244/858, and was a boy of five when al-Bandanīğī died in 284/897.

Throughout its long history, the Arabic lexicographical tradition has manifested an amazing variety in content and form. As far as the *mubawwab*

⁴² Baalbaki, *Arabic Lexicographical Tradition*, 58–9.

⁴³ Law, 'Indian influence'; Baalbaki, 'Introduction', xxiii.

⁴⁴ Haywood, *Arabic Lexicography*, 68–9, 120–1; Āl Yāsīn, *al-Dirāsāt al-luğawiyya*, 90; Fenton, 'Sa'adyā', 661.

lexica are concerned, their topics cover a wide spectrum, as demonstrated above. Furthermore, the *mubawwab* lexica that are arranged according to form vary considerably in their arrangement within each genre and subgenre. In this regard, they resemble their *muğannas* counterparts. In both types, the variety pertains not only to the use of the phonetic-permutative, the alphabetical, or the rhyme arrangements, but also to arrangement by first letter, by first two letters, or by all letters. Even the basic principle set by al-Ḥalīl of arranging entries by roots was sometimes violated by lexicographers who took augments into account in the arrangement of lexical items, such as Kurāʿ al-Naml in *al-Munagğad*. Also noteworthy is that several authors used different systems of arrangement in different works, as did al-Zamaḥṣārī, who arranged *Asās al-balāga* alphabetically but *Muqaddimat al-adab* by last letter.

Given that the philologists were little interested in a comparative approach to the study of Arabic, it is not surprising that polyglot dictionaries are quite rare in the tradition, and these appeared relatively late. One example is al-Zamaḥṣārī's *Muqaddimat al-adab*, an Arabic–Persian lexicon that belongs to the genre of morphological writing (see above and Chapter 11). Its material includes nouns, verbs, and particles, but only in the section that deals with verbs does al-Zamaḥṣārī adopt a formal arrangement of the words, which he divides based on their patterns and in accordance with the rhyme system. There are also a few Arabic–Persian and Arabic–Mongolian glossaries, in addition to a list of Arabic words with Turkish, Persian, Mongolian, Greek, and Armenian equivalents, now known as the Rasulid Hexaglot, which was compiled in the eighth/fourteenth century, by or for the sixth Rasulid king of Yemen, al-Malik al-Afḍal.⁴⁵ Only in the modern era do polyglot dictionaries constitute a sizeable part of lexicographical writing in the Arab world.

Ibn Manẓūr proposed the criteria of *ğamʿ* and *waḍʿ* (see above) to evaluate the works of his predecessors. Indeed, lexical works, particularly of the *muğannas* type, vary tremendously as to these two criteria. There are two aspects to *ğamʿ*, in both of which exhaustiveness is the ultimate aim. From the onset of the tradition, al-Ḥalīl had resolved the first aspect by devising a method which ensures that no root would be excluded, at least theoretically, from his lexicon. In spite of the abandonment by many lexicographers of al-Ḥalīl's phonetic arrangement, morphological divisions, and root permutations, all subsequent authors were indebted to his scheme, which achieves exhaustiveness of Arabic roots and represents the discovery stage

⁴⁵ Carter, 'Arabic lexicography', 108; Naṣṣār, *al-Mu ʿğam al-ʿArabī*, 1.91–6; Golden, 'Rasulid Hexaglot'.

in the tradition. As for the other aspect of *ġam'*, namely the inclusion of the attested lexical items derived from the various roots, the above survey demonstrates how lexicographers sharply differed in their approach to this matter. But given the vast wealth of Arabic vocabulary, it is not unexpected that certain words are not found even in the most comprehensive lexica. For example, about fifty words that feature as early as the third/ninth century in the works of one of the most renowned Arab men of letters, al-Ġāḥiẓ, are not noted in any lexicon.⁴⁶ Similarly, the temporal limits imposed by most lexicographers on acceptable usage – generally not exceeding the end of the fourth/tenth century – resulted in the exclusion of neologisms and new meanings of old words. Also, contemporary usage was only sporadically included, e.g. the medical terms recorded by al-Firūzābādī and words from the Egyptian vernacular noted by al-Zabidī.

The criterion of *waḍ'* primarily involves the issue of word definition and that of the internal arrangement of the lemmata. Clear and succinct definitions do feature in many lexica, but it is not uncommon for definitions to be obscure or imprecise, as when a word is defined as being the contrary of its antonym, or by a synonym whose meaning may not be readily accessible, or simply by describing it as *ma'rūf* (well known), in which case no definition is provided. Lexica also vary in the arrangement of the material within the lemmata. Most *muġannas* lexica lack a template that determines the order in which verbal and nominal patterns and other elements are listed, and the reader might have to read much of the lemma, or all of it, to find the required word. There were, however, some relatively late attempts to amend this situation, as in Ibn Manẓūr's habit of listing semantically related items together and, more significantly, in the introduction by Ibn Sida in *al-Muḥkam* and al-Ṣaġānī in *al-'Uḇāb* of a specific order for citing the derivatives within the lemma.

The 'classical' Arabic lexicographical tradition extended for almost a millennium, during which the lexicographers painstakingly gathered the corpus, adopted various ways in arranging the material, and hardly left a subject of philological interest for which they did not devote a *mubawwab* lexicon. Their discussion of questions of authenticity and correctness of the data and how inclusive a lexicon should be attests to their awareness of major methodological issues that any lexicographer has to handle. One can only admire their originality, their passion for Arabic, and their role in preserving the vast philological and cultural heritage of the Arabs.

⁴⁶ Al-Sāmarrā'ī, '*al-Ma'āġim al-'Arabiyya*', 196–203.

Hebrew to c. 1650

AHARON MAMAN

The first section of this chapter introduces the changing fortunes of the Hebrew language, and the ways in which it came to be an object of study, before the emergence of formal Hebrew dictionaries. The first of these was written in 902 by Rav Saadia Gaon, and stands at the head of a tradition of medieval and Renaissance Jewish lexicography, in which both Biblical and post-Biblical Hebrew were addressed. Having described the principal dictionaries in this tradition, this chapter will comment on some of the ideas about the Hebrew language, and about lexicography, which they share.

The Hebrew Language and Hebrew Lexicography

Hebrew lexicography is inextricably tied to the history of Hebrew itself. As is well known, Hebrew and its speakers suffered severe traumas throughout their history, traumas that did not allow transmission and continuous natural speech throughout the ages. The exile of the ten tribes to Assyria in 722 BC led to the loss of the Hebrew dialects of the northern Land of Israel, and the exile of the Kingdom of Judah to Babylon from 597 to 586 BC caused the loss of the natural living speech of the Judaic Hebrew state, the one represented basically in the books of the Bible as transmitted to us.¹ The destruction of the Second Temple in AD 70 and the failure of the Bar-Kochba revolt in 135 resulted in the disappearance of Hebrew speech, and consequently in the loss of transmission of part of the knowledge of Hebrew, so that some modern scholars declared that Hebrew was dead for 1,700 years until it was revived in natural daily speech towards the end of the nineteenth century.

Furthermore, these traumas changed even the nature of formal Hebrew. When the Jews returned to Israel from the Babylonian exile, they came across an entirely different dialect, more popular, that became the standard and the

¹ Exile of the ten tribes: 2 Kings 17:1–6; exile of the kingdom of Judah: 2 Kings 25.

norm in the Land of Israel. That popular Hebrew dialect, initially used orally, had some effect on the biblical books composed or edited during the Second Temple period. Moreover, in post-biblical literature it became a literary language, later nicknamed 'the language of the Sages' (*leshon Ḥakhamim*), to the point that the first canonical composition penned after the Bible, the Mishnah, as well as later rabbinic works, were written in it. But even that dialect was pushed back after the loss of sovereignty and by the spread and the takeover of Aramaic. Beginning in the fifth century BC, Aramaic became the *lingua franca* and the diplomatic language of the Persian empire, from Mesopotamia to the Land of Israel and to Elephantine in Egypt.² It is no wonder that some biblical chapters in the books of Daniel and Ezra were composed in Aramaic.³ In the course of the five following centuries, Aramaic was upgraded from the rank of a diplomatic language to that of the language of everyday speech, the most common language in the entire area.

This obviously affected the degree of mastery of Hebrew. Some fascinating anecdotes recorded in the Babylonian Talmud show a reduced knowledge of Hebrew even among scholars. In some cases, the Sages had to learn the meaning of a biblical or even Mishnaic word from servants or older people (*Sages* denotes the rabbis who codified the Mishnah from its beginning up to AD 220; later it refers to the authors of the Talmud and other rabbinic texts). In a particular case, the Sages had forgotten the meaning of the word *mat'ate* (Isaiah 14:23), until they heard it used in current speech with the meaning of 'broom' by a maidservant, who was one generation older and perhaps also from a rural society that was still familiar with natural Hebrew usage. They also forgot the meaning of the post-biblical word *serugin*, until the maidservant asked them one day, 'Why are you coming in *serugin* [alternating, one by one] and not at once?' (Talmud: Megillah, 18a). Such Talmudic anecdotes as these testify to an inadequate knowledge of Hebrew, even among scholars who frequented the most important Jewish academy of its time.

Moreover, during the time of the Mishnah, Hebrew was soaked with Greek and Latin words, following the conquests of Alexander the Great, and government by the Romans from the Hasmonean period onwards. Some of these words have reached spoken Hebrew today, such as *ambatia* (< Greek

² Cowley, *Aramaic Papyri*, xv; Kutscher, *History of the Hebrew Language*, 71; Botta, *Aramaic and Egyptian Legal Traditions*, 15.

³ Daniel 2:4b–7:28; Ezra 4:8–6:18 and 7:12–26; also Aramaic are the single expression *Jegar-Sahadutha* in Genesis 31:47 and the single verse denouncing idolatry in Jeremiah 10:11.

embatē 'bath') and *itstadion* (< Greek *stadion* 'race-course, amphitheatre').⁴ The Babylonian *amoraim* (composers of the Talmud), who lost live and natural contact with Greek and Latin, were unable to decipher the Greek and Latin words in the Mishnah, and explained them as if they were acronymic amalgams of Hebrew and Aramaic. For example, the term *prosbul*, from *pros boule* 'before the council, assembly', initially designating a document issued by the Jewish court called *beth-din* and meant to protect the interest of lenders after the Sabbatical year, without which the debt would be cancelled, was reinterpreted by the *amoraim* as an acronym of *pros* 'an enactment for' *buli* 'the rich' and *buti* 'the poor'.⁵

The Mishnah defines many words for *halakhic*-legal purposes, such as two in successive verses of Leviticus: 'What are considered *gleanings*? Whatever drops down at the moment of reaping' and 'What is considered a *grape*? Whatever drops at the moment of vintage.'⁶ Gaonic *responsa* to questions from learners throughout the Jewish world (*Gaon* was the title of the leader of the highest Jewish academy who also served as spiritual leader of Jews from the seventh through the eleventh centuries) are also full of definitions of biblical and Talmudic words.⁷ This is also true for sporadic interpretations and later systematic commentaries. That one can single out the lexicographic material embedded in commentaries can be seen in the example of the modern scholar Yitshaq Avineri, who accumulated all such lexicographical material from the commentaries of Rashi (the acronym by which the eleventh-century Rabbi Shelomo Itzhaki is known) and put it together into his *Heikhal Rashi*, as if it were Rashi's dictionary on biblical and Talmudic literature. Avineri's work may be criticized, yet it proves that scattered lexicographic material in commentaries or other texts may be accumulated and presented as substantial lexica.

Glossaries

Rav Saadia Gaon, a tenth-century scholar and teacher, who translated part of the Bible from Hebrew into Arabic, with a commentary, also composed *Alfāz*

⁴ Kutscher, *History of the Hebrew Language*, 137; Lieberman, *Yəḡanit və-yaynut be-Erets-Yiśra'el*, 5–21; Krauss, *Griechische und lateinische Lehnwörter*, II.59, II.119.

⁵ For *prosbul*, see Krauss, *Griechische und lateinische Lehnwörter*, II.482; for the reinterpretation, see Talmud: Gittin 36b.

⁶ Talmud: Peah 4:10 (on *leqet* 'gleanings' in Leviticus 19:9), 'ezehu leqet? ha-nošer bi-š'at ha-qəšira', and Peah 7:3 (on *peret* 'grape' in Leviticus 19:10), 'ezehu peret? ha-nošer bi-š'at ha-bəšira'.

⁷ Brody, *Geonim of Babylonia*, 191.

alMishna ('The words of the Mishnah'), a series of glossaries to certain tractates, explained in Arabic.⁸ Many followed in his footsteps.⁹ The glossaries were usually bilingual, the language of the definitions depending on the spoken language of the author and his audience. In France, Hebrew–French glossaries were compiled and used by Rashi (for whose contributions to the lexicography of Old French, see Chapter 13).¹⁰ In Italy, the Hebrew–Italian glossary to the Bible composed by Yedidia of Rimini was copied many times.¹¹ There were Hebrew–Greek glossaries of biblical and Mishnaic Hebrew.¹² In Spain, Hebrew–Spanish glossaries were composed; in Persia, there were Hebrew–Persian glossaries; and in the Arabic-speaking countries, the lemmata were defined in a Judeo-Arabic *koine* or in local dialects.¹³ In some cases they add an etymological proof, in the form of another occurrence of the same lemma in the literature.

The Dictionary as Commentary

Glossaries and dictionaries were reference books for the study of the Bible before the exegesis genre was developed. Even when Bible commentaries were composed in the Middle Ages, they were usually written on only one book or a set of books from the Bible. Before Rashi, hardly anyone had written a comprehensive interpretation of the Bible and Talmud.¹⁴ In many places the dictionary fulfilled the role of a commentary, as it included the interpretation of the entire vocabulary of the Bible. Moreover, as these ancient dictionaries came to meet a social-cultural need, they often went beyond the pure definition of the headword and provided the interpretation of its biblical context, as if it were an exegesis by genre. The study of the Bible and especially the teaching of the Talmud were the foundations of traditional Jewish education from early childhood to old age, and the glossaries and dictionaries were the constant companions of the students.

Furthermore, these dictionaries, much like Saadia's commentary on and translation of the Bible, manifested a grammatical-philological approach,

⁸ Saadia Gaon, *Alfāz alMishna* (1985), is an edition; see also Brody, *Geonim of Babylonia*, 267–9.

⁹ On the beginnings of the study of words, see Dotan, *Nitsanim rishonim ba-hokhmat ha-millim*; on the differences between dictionaries and glossaries, see Maman, *Otsrot Lashon*, xxvi–xxx.

¹⁰ *Le glossaire de Bâle* (1972) is an edition. ¹¹ Maman, *Otsrot Lashon*, 26–9, 37.

¹² De Lange, *Greek Jewish Texts*, chapters 10 (Biblical) and 16 (Mishnaic); Schnoll, 'Mishnaic Hebrew glossary'.

¹³ Maman, *Otsrot Lashon*, 'Introduction', §8. ¹⁴ Grossman, *Haxme Tsarfut ha-rishonim*, 438.

according to which the Scriptures were interpreted in the literal *peshat* way, contrary to the impressionistic and associative interpretation that had prevailed for centuries in the *Midrash*. The effort was now, from the tenth century on, to establish the interpretation of the Bible on the scientific linguistic foundations that flourished then. The meanings of many biblical words were lost or drowned into the sea of midrashic commentary. Thus the *peshat* commentators took great pains to reconstruct the lost original meanings of those words, using all sorts of sophisticated philological methods.

Rav Saadia Gaon and the First Systematic Dictionaries

Saadia's glossaries have just been mentioned; he also composed two dictionaries. Perhaps it is no accident that this activity was carried out in what is now Iraq (though Saadia himself initiated it in Egypt), where the Jews lived in relative prosperity and personal and social security over several centuries. The economic, social, and political conditions allowed them to establish academies for advanced studies to which many students flocked from across the Jewish world. Throughout the Mishnaic, the Talmudic, and most of the Gaonite periods, in other words before the tenth century, no systematic dictionaries or grammars were composed as the conditions were not yet ripe. Apparently, as long as the Masoretes' work was not completed – as long as the vocalization and cantillation of the Bible were not fully set, and as long as a cultural decision was not made to finally prefer the Tiberian biblical tradition over the competing Palestinian and Babylonian traditions – it was impossible to formulate rules of grammar and to compose dictionaries for the vocabulary of the Bible.¹⁵ It seems that those conditions were met close to Saadia's time. It was thus he who composed the first Biblical Hebrew grammar and Hebrew dictionary.¹⁶ No doubt this was also related to his personality. He was a leader of great stature who paved the way in many areas, interpretive and philosophical. His philological work is also inextricably linked to his activity in poetry. Not only did he write many liturgical poems, but he also founded a school of poetry in which he preached a return to the Bible and the cultivation of its language and style.¹⁷

The first Hebrew dictionary, Saadia's *Egron* ('Thesaurus'), was exactly designed for this purpose: to convince poets to use biblical words and

¹⁵ Téné and Barr, 'Linguistic literature', col. 1352. ¹⁶ Brody, *Geonim of Babylonia*, 316–23.

¹⁷ Zulay, *ha-ʾaskola ha-payṭanit šel Rav Saʿadya Gaʿon*, 19; Tobī, 'Piyuṭe Rav Seʿadya Gaʿon'.

forms, with an emphasis on the rarer ones, along with post-biblical poetic innovations. After several centuries of Hebrew poetry in Palestine and Babylonia, Saadia sought to return biblical studies to the heart of Hebrew Jewish culture. In the introduction to his *Egron* he laments the state of knowledge of Hebrew at the time.¹⁸ His dictionary presents a selection of about a thousand lofty words from the Hebrew classic sources, as well as his own neologisms, which he recommends that the poets incorporate in their liturgical poems. Each word was translated into Arabic. The *Egron* entries were arranged in two alphabetical sequences, one according to the first letter of the lemma and one according to its last letter, the former to help poets in selecting words fitting their acrostics and the latter to help with rhymes. It seems that this was a kind of imitation of the poetic achievements of contemporary Arabic scholars (for whom see Chapter 8), who focused on the study of the language of the Qurʾān and poetry on the one hand and on the study of the pure Arabic language preserved by the Bedouins on the other. In general, indeed, the initiation and setting of systematic Hebrew grammar and lexicography were inspired by the Arab philologists of the time. In some cases one can even point to the specific Arab texts that were used as models for Hebrew philology.¹⁹

Saadia's second lexicographic work, *Kitāb al-sabʿīn lafza al-mufrada* 'The Book of the seventy *hapax legomena*', a list of biblical words with their Mishnaic counterparts, without any specific order, was meant to support a polemical point.²⁰ He sought to show that certain biblical words are intelligible only via the traditional rabbinic interpretation, against the Karaites who denied the Oral Law (post-biblical rabbinic literature) and only adhered to the Written Law (the Bible), as they believed that 'the teaching of God is comprehensive' (Psalm 19:8) and needed no external interpretation.²¹ Saadia shows in his *sabʿīn lafza* that this view was impossible and impractical, presenting as a proof a selection of about ninety (*sic*) biblical words that can only be understood by the use of their counterparts in rabbinic literature.

The First Successors of Saadia: Menaḥem and Alfāsi

Saadia's lexicographical works were limited in scope: his *Egron* did not focus on biblical vocabulary, and his *sabʿīn lafza* was a short list. His contemporary

¹⁸ Saadia Gaon, *Haʿegron* (1969), 150. ¹⁹ See Becker, *Māqorot ʿarviyim*.

²⁰ See his *Kitāb al-sabʿīn lafza*. ²¹ Wieder, *Judean Scrolls and Karaism*, 57.

Yehuda ibn Quraysh composed three interesting small dictionaries (see below), but these were likewise not biblical. Some years passed before the challenge imposed by Saadia – the need to study the language of the Bible – found a response. Towards the middle of the tenth century, the composition of formal, complete, and well-edited dictionaries began with *Maḥberet* by Menaḥem ben Saruq and *Jāmiʿ al-ʿalfāz* by David ben Abraham Alfāsi.²² The first of these has entries for about 2,500 roots and a total of 8,000 lexemes.

These dictionaries were confined to the vocabulary of the Bible alone (disregarding post-Biblical Hebrew), which was an appropriate response to Saadia's poetical ideal. Moreover, Alfāsi was a Karaite, and so it was natural that, although he refers to the rabbinic literature dozens of times, his work only deals with the vocabulary of the Bible, including its Aramaic parts.²³ As for Menaḥem, he was the court poet of Hasdai ibn Shaprut, who was governor of the Jewish community, personal physician of the Umayyad caliph of Córdoba, and finance minister in his government, and who set up a court in Córdoba, parallel to the royal court, which cultivated poetry, culture, and science in open cultural competition with the court in Baghdad.²⁴ Menaḥem wrote his dictionary at the invitation of his patron, designing it to meet the social and cultural needs which Hasdai perceived.

Though these dictionaries were real pioneers, they do not look as if they were initial experiments. To be sure, there are some fundamental differences between the two. For instance, *Jāmiʿ* is a bilingual Hebrew–Aramaic dictionary, and so extensive that the author himself had to make an abbreviated version of it; two later scholars composed even more concise editions.²⁵ By contrast, *Maḥberet* is essentially a compact monolingual Hebrew dictionary. Again, *Jāmiʿ* freely used etymological and semantic comparisons of Hebrew with Aramaic and Arabic, including cases where metathesis or consonant shift were required to show an affinity, whereas Menaḥem compared only Hebrew and Aramaic, refusing in principle to compare the Holy Tongue with Arabic, and even his comparisons with Aramaic are limited to a minimum of words where neither metathesis nor consonant shift were required.²⁶ Yet there are some common features (it is still unknown whether either Menaḥem or Alfāsi knew of his counterpart's work). Both dictionaries,

²² Menaḥem, *Maḥberet* (1986), and Alfāsi, *Jāmiʿ al-ʿalfāz* (1938), are editions; for a detailed description of the former, see Mirsky, *Torat ha-lašon šel Menaḥem ben Saruq*.

²³ Netzer, 'Lašon', 84–124; Maman, *Comparative Semitic Philology*, §9.12.1.1.

²⁴ Ashtor, *Qorot haYhudim bi-Sfarad haMuslamit*, I.111ff.

²⁵ Alfāsi, *Jāmiʿ* (1938), Introduction, chapter 4.

²⁶ Maman, *Comparative Semitic Philology*, §10.1.

for instance, combine the Hebrew and Aramaic entries in a single wordlist as if they were not two separate languages but one and the same language. The entries in both are arranged alphabetically according to their roots, assuming monoliteral and biliteral verbal roots.

The theory of monoliteral and biliteral verbal roots was, however, about to become obsolete, as a result of the most outstanding contribution of the late tenth-century scholar Yehuda Ḥayyūj to Hebrew philology in general and to lexicography in particular: the discovery of universal triliterality of the Hebrew verb and the majority of nouns.²⁷ The root theory before Ḥayyūj artificially brought together entry words that were far apart and created forced semantic relationships. For instance, the first section of root *b-ḥ* in *Maḥberet* presents verses which contain the forms *mi-bəḥi*, *nivḥe*, *ha-baḥa*, *navoḥu*, and *wa-yit'abbəḥu*, all of which Menaḥem regarded as derived from the same root *b-ḥ* and therefore forced to fit into one and the same meaning, 'movement'.²⁸ But from Ḥayyūj onwards they have been regarded as three separate roots, *b-w-ḥ*, *b-ḥ-y/ʔ*, and *ʔ-b-ḥ*, each with its own particular meaning. For example, *ʔ-b-ḥ* is explained by Yona ibn Janāḥ in his *Kitāb al-ʔUṣūl* (see below) as 'pick up, go up, or evaporate'.²⁹

Once the linguistic theory which gave *Maḥberet* and *Jāmi*^c their structure was superseded, one might have expected them to go out of the scene. Yet *Maḥberet* has survived to our time because, in addition to being compact, it was a monolingual Hebrew dictionary, which was welcome in Europe and outside the Arabic-speaking area. Had the Jewish intellectuals in Germany, France, Provence, and Italy been able to read the new linguistic works written in Arabic by the Spaniards Ḥayyūj and Ibn Janāḥ, they might have seen *Maḥberet* as totally outdated. But they did not master Arabic, stuck to *Maḥberet*, copied it with its old-fashioned teachings, and passed it down from generation to generation. Especially instrumental in its canonization was Rashi, who long made use of it and embedded many quotations from it in his commentaries on the Bible and Talmud.³⁰ Alfāsi's dictionary too survived to this day despite its outdated linguistic theory, because the Karaites took care

²⁷ Ḥayyūj, *Two Treatises* (1870), is an English translation, with an edition of an early Hebrew translation; Ḥayyūj, *Weak and Geminate Verbs* (1897), is an edition of the Arabic original; for a discussion, see Téné and Barr, 'Linguistic literature', §2.7.2.

²⁸ Menaḥem, *Maḥberet* (1986), *83: the verses are 'mi-bəḥi nəharot ḥubbas' (Job 28:11), 'nivḥe yam' (Job 38:16), 'cemeq ha-baḥa' (Ps 84:7), 'navoḥu 'edrei baqar' (Joel 1:18), and 'wa-yit'abbəḥu ge'ut 'ašan' (Isaiah 9:17).

²⁹ Maman, 'Flourishing era', 265.

³⁰ Maman, 'Lə-middat hissardutam šel ḥibburim bə-ḥoxmat ha-lašon', 227–8.

to preserve the works of their great scholars and transfer them to later generations.

Ibn Janāḥ

Rabbi Yona ibn Janāḥ wrote the first grammar and dictionary to be based on Ḥayyūj's root theory. He saw his work as simply a complement to Ḥayyūj's. Initially, around the year 1012, he composed a work called *Kitāb al-Mustalḥaq*, to supplement whatever Ḥayyūj had missed in his books; then he wrote short treatises to elaborate on particular biblical words or forms or argue with his great opponent Samuel Hanaguid (see below); and finally he compiled his systematic dictionary *Kitāb al-ʿUṣūl* ('The book of roots'), completed around 1050, where he virtually corrects the errors of all his predecessors.³¹ As he saw his initial writings along with Ḥayyūj's works as a single oeuvre, he did not bother to repeat material from the shorter works in *Kitāb al-ʿUṣūl*, but rather referred to them. This was of course inconvenient for the user, for which reason later lexicographers saw it fit to compose full and self-sufficient dictionaries.

Yet, Ibn Janāḥ's work had a great impact on Hebrew lexicography. He brought the use of comparative philology in dictionary-making to heights unknown before. Through the use of comparisons with Arabic and sometimes sophisticated Hebrew, Aramaic, and Arabic chains, he succeeded in discovering hidden meanings. For example, the form *meltaḥa* is a *hapax* in the Bible (2 Kings 10:22) and is unexplained in the Mishnah. In the absence of any tradition of interpretation or use, medieval philologists were uncertain about its meaning. Ibn Janāḥ found a decisive solution: 'It is possible that this term designated a closet for clothes storage, as it is translated in the Aramaic Targum *qəmatraya*. Now *qamtar* in Arabic is an ark, a closet.'³² We have here a sophisticated triangular pattern, combining a comparison of a Biblical Hebrew entry with entries from the Aramaic Targum and spoken Arabic. The Biblical Hebrew *meltaḥa* was semantically opaque; but the traditional Aramaic local translation *qəmatraya* could be used as an intermediate tool, although Aramaic had also ceased to be spoken among most Jews, so this term in Aramaic had also become obsolete; but it can be explained by its Arabic cognate. This in its turn illuminated the Hebrew.

³¹ Editions are Ibn Janāḥ, *Kitāb al-Mustalḥaq* (2006), Ibn Janāḥ, *Opusculs* (1880), and Ibn Janāḥ, *Kitāb al-ʿUṣūl* (1875); an edition of the translation of the latter into Hebrew (discussed below) is Ibn Janāḥ, *Sefer ha-šorašim* (1896).

³² Ibn Janāḥ, *Sefer ha-šorašim* (1896), 250.

Naturally, Ibn Janāḥ also relied on rabbinic traditional interpretation, whenever he saw fit, and on comparisons of Biblical to Rabbinic Hebrew.³³ Ibn Janāḥ's predecessors, such as Saadia, Ibn Quraysh, and Alfāsi also used this technique, and occasionally he adopted their view. But Ibn Janāḥ turned the method of comparison into a well-founded and systematic tool in biblical exegesis and compared 303 entries.³⁴ Most of his comparisons were original and all were instructive, as he never resorted to trivial comparisons, just as in his dictionary he never defined familiar words. Here is one example: King David did something to the Philistine idols after the war in Ba'al-Peratsim, the verb being *va-yissa'em* (2 Samuel 5:21), and Ibn Janāḥ explained this as 'he burned them', based on the comparison with *massi'in massu'th* 'lighting bonfires' (Talmud: Rosh Hashanah 2:3).³⁵ With this comparison he removed any concern that the reader would understand *va-yissa'em* as referring to King David and his men *taking* the idols, as the root *n-s-ʔ* is regularly understood, thus assigning to David the sin of practising idolatry. With this explanation Ibn Janāḥ also resolved the textual difference between 2 Samuel and its parallel in 1 Chronicles 14:12, *va-yomer David va-yisrafum ba-ʔesh* ('David commanded and they burned them').

Ibn Janāḥ's achievements were so outstanding that Joshua Blau says that one of the greatest achievements of Wilhelm Gesenius' dictionary (1810–12) is that he made proper use of the dictionary of Ibn Janāḥ.³⁶ *Kitāb al-ʔUṣūl* is rightly considered the crown of scientific achievements in the realm of medieval Hebrew lexicography. Ibn Janāḥ wrote his definitions in Arabic, the *koine* of the majority of world Jewry in Arab lands at that time. A century later, Yehuda ibn Tibbon translated the dictionary into Hebrew under the title *Sefer ha-šorašim* ('The book of roots') for the sake of Provençal scholars and other Jews who lived outside Arab lands; he also translated Ibn Janāḥ's Hebrew grammar, *Kitāb al-Lumaʕ*, as *Sefer ha-Riqmah*.³⁷ Bible dictionaries and other works written in future generations all relied either directly or indirectly on Ibn Janāḥ's book.

Later Dictionaries of Biblical Hebrew

Parallel to Ibn Janāḥ's dictionary, the poet and statesman Samuel Hanaguid composed *Kitāb al-Istighnāʔ* ('Book of amplitude'), a very detailed

³³ Neubauer, 'Notice sur la lexicographie hébraïque', 186ff.; Bacher, *Hebräisch-neuhebräische und hebräisch-aramäische Sprachvergleichung*; Netzer, 'Lāšon', 176–304.

³⁴ Netzer, 'Lāšon', 299. ³⁵ Netzer, 'Lāšon', 253. ³⁶ Blau, 'Millonim Miqra'iyim', 906.

³⁷ Ibn Janāḥ, *Sefer ha-šorašim* (1896), is an edition of ibn Tibbon's rendering of the dictionary.

Hebrew–Arabic biblical dictionary with the scope of a concordance. Yet, except for a few entries preserved in a *geniza* (a repository for unusable documents containing the name of God and, later, any document written in Hebrew letters, which could not be thrown away as ordinary waste), it did not survive. It appears that the book's size was too great for it to be transmitted.³⁸ From the vestiges published in 1916 it appears that every entry was rich in lexical, semantic, and grammatical data.³⁹ Some of the lost entries have been reconstructed through late citations from the dictionary, and from Hanaguid's usage in his Hebrew poetry.⁴⁰ Some scholars regard *al-Istighnā'* as the zenith of medieval Hebrew lexicography, but it is hard to support this view, due to the loss of the work.

One might have thought that Ibn Janāḥ's dictionary would have been sufficient for readers and scholars throughout the Jewish world for many generations to come, particularly once it was translated by Yehuda ibn Tibbon. Nevertheless, more dictionaries were composed after his, for several reasons. An early example is Shelomo ibn Parḥon's *Maḥberet he-ʿArukh*, completed in Salerno in 1161, a monolingual Hebrew dictionary preceded by a grammar section, which is an adaptation of Ibn Janāḥ's *Sefer ha-šorašim*, *Kitāb al-Mustalḥaq*, and *Sefer ha-Riqmah* with additions from the works of Ḥayyūj.⁴¹ By the turn of the thirteenth century, the Spaniard Yaʿakov ben Elʿazar composed a book titled *Kitāb al-Kāmil* ('The comprehensive book'), including a grammar and a dictionary. The book disappeared in the fourteenth century, and was considered lost until recently. In fact, some of its vestiges were discovered by 1880 in the Geniza collections in St Petersburg, and some specimens were published in 1916, but were misidentified; they have now been shown to be vestiges of *Kitāb al-Kāmil* and are in preparation for publication.⁴² Yaʿakov ben Elʿazar followed in the footsteps of his predecessors but added his own input as well. For instance, he made every effort to explain the exceptions as if they were normal. For instance, the form *li-mnuḥayxi* (Ps. 116:7; Return, my soul, to your rest) was taken by Ibn Janāḥ as an exception (to the rule that infinitives should not regularly be inflected in dual or plural), thus the (singular) form *li-mnuḥex* was expected. Yet Ben Elʿazar sees it as normal, a noun in the plural. It appears that his work was popular among Arabic-speaking scholars, as it was constantly quoted by two

³⁸ Maman, 'Lə-middat hissardutam šel ḥibburim bə-ḥoxmat ha-lašon'.

³⁹ See Kokovzov, *Novye materialy*. ⁴⁰ Perez, 'Muvaʿot'; Uqashi, 'Millon Šomuʿel ha-Nagid'.

⁴¹ Ibn Parḥon, *Maḥberet* (1844), is an edition; see also Téné and Barr, 'Linguistic literature'.

⁴² The publication of 1916 was in Kokovzov, *Novye materialy*; the modern identification is in Maman, 'Sefer diqduq'.

popular lexicographers, one in the east and one in the west. *Kitāb al-Taysīr* ('The book of facilitation'), a Hebrew–Arabic dictionary by the Karaite Shelomo ben Meborakh, is mainly based on *Kitāb al-Kāmil*, which it mentions explicitly seventy-two times (no doubt it incorporates much more of *Kāmil*'s material without mentioning it by name), though it also draws on sources such as Alfāsi, Ḥayyūj, and Ibn Janāḥ.⁴³ Likewise, Radak, to whom we now turn, quoted *Kitāb al-Kāmil* explicitly more than thirty times and probably – the point needs further investigation – implicitly incorporated more material.

R. David Kimchi, known by acronym as Radak, a descendant of a Spanish family that had moved to Provence, took the trouble to compose a new monolingual Hebrew dictionary, probably understanding that Ibn Janāḥ's dictionary was too scientific for the readers of Radak's time and place and incomplete in that it referred the readers to his earlier works and to those of Ḥayyūj, which made it less practical for the lay reader. Thus Radak composed a practical dictionary, in which he basically followed in Ibn Janāḥ's footsteps, but also incorporated significant midrashic and even Halakhic material, probably as he felt that his audience was not subtle enough to differentiate between *peshat*, which as we have seen is the basic meaning of a word, and *derash*, the sort of non-basic meaning found in midrashic and Halakhic exegesis.⁴⁴ Radak may be primarily considered as the agent of the linguistic achievements of Spanish scholars outside the boundaries of Arabic-speaking areas, such as Provence and Italy, and his dictionary was so popular in the centuries to come that it is still used by certain circles in our own time. There have been printed editions from the incunabular period onwards, and a new critical edition is being prepared.⁴⁵

Though Radak's *Sefer ha-šorašim* was widely circulated, more Hebrew dictionaries were composed. There is hardly a country where no Hebrew dictionary or glossary was composed, and there is not enough space here to list them all.⁴⁶ An interesting example is the dictionary *Maqre Dardege*, composed by Perez Trabot (a descendant of a French, and thence a Catalan, family, as he calls himself *Ẓarfati* or *Catalani*), which was written as a Hebrew–Italian–Arabic dictionary in Italy (Naples, 1488), and was then transformed into other versions for different countries, for example

⁴³ Shelomo, *Kitāb al-Taysīr* (2010), is an edition. ⁴⁴ Maman, 'Peshat and derash', 343.

⁴⁵ Kimchi, *Radicum liber* (1847), has been the basis of later editions; for recent critical work, see Petrover, 'Darko šel rabbi David Qimḥi', and Grunhaus, *Challege of Received Tradition*.

⁴⁶ For some unpublished examples, see Maman, *Otsrot Lashon*.

Hebrew–Spanish–Arabic in Spain.⁴⁷ It looks very much like a bilingual or trilingual dictionary of today.

Towards the end of the fifteenth century, the Spaniard Se'adiah ibn Danan composed *Sefer ha-šorašim*, a Hebrew–Arabic bilingual dictionary.⁴⁸ The definitions are concise and worded in western Judeo-Arabic. The entries are arranged in alphabetical order of their roots, following the classical method of Hebrew lexicography. The four-letter and five-letter roots come in separate groups, after the trilaterals, also according to Ibn Janāḥ's and Radak's method. In many cases, Ibn Danan follows Radak's opinion rather than Ibn Janāḥ's, as is the case in the entry *ḥbḥ*, which shows a scientific regression, as does his basing of etymologies on similarities of sound. The issue of the exchange of letters, which was a serious one in the eyes of his predecessors, seems to him minor. Thus, he seeks etymological affinity between remote roots. For example, *va-ye'abeq* (Genesis 32:25) 'fought' is for him nothing but *va-yehabeq* 'embraced'. In his eyes, it is easier to assume a sound shift *ḥ/h* than to accept Ibn Janāḥ's or Radak's opinions, or to determine homonymy between the roots *ḥbq*¹ and *ḥq*².

Josef Kaspi of Argentièrre in Provence composed a still unpublished monolingual dictionary of Biblical Hebrew.⁴⁹ As Kaspi was a logician and philosopher, he applied logic to Hebrew grammar and lexicography. This is reflected very well by the definitions in his dictionary, where he tends to find the closest possible links between words from the same root. One of the most significant characteristics of Kaspi's dictionary is his treatment of homonyms: he preferred to assign a minimal basic meaning to a root family without paying attention to the specific contextual meanings, so that what others saw as homonyms were in his eyes more a matter of polysemy.⁵⁰ Kaspi explicitly criticizes his predecessors 'for dividing up the root into numerous subjects'. For instance, whereas Ibn Janāḥ accords the root *k-m-r* four different meanings – 'a stirring of compassion', 'dryness and shrinking', 'priests', and 'a net for hunting' – Kaspi set a basic meaning 'shrink' for all occurrences.⁵¹ Hence a trap is called *miḵmar*,

⁴⁷ See Jerchow, 'From *Makre Dardeke* to *Sefer 'Arb'ah ve-'ešrim*', and Hary and Ángeles-Gallego, 'Lexicography and dialectology'.

⁴⁸ Se'adiah, *Sefer ha-šorašim* (1996), is an edition; see also Maman, 'Sefer ha-šorašim leRabbi Se'adiah ibn Danan'.

⁴⁹ See Aslanov, *Le provençal des Juifs et l'hébreu en Provence*, and Kahan, *Ha-^cvirit bi-r'i ḥoxmat ha-higgayon*.

⁵⁰ Kahan, 'Aspects of medieval lexicography'.

⁵¹ For the first, second, and fourth senses, Ibn Janāḥ cites 1 Kings 3:26, *niḵmaru raḥameha* 'al *bənah* ('her heart yearned upon her son'); Lamentations 5:10, *orenu kə-lannur niḵmaru* ('our skin is hot as an oven'); and Isaiah 51:20, where the form is *miḵmar*, *miḵmoreṭ*; the word for 'priests' is *komer*, *kəmarim*.

because it catches and causes the hunted animal to contract; yearnings of compassion also cause the flesh to contract, which is where the phrase *niḵmāru raḥamaw* (Genesis 43:30) ‘his heart yearned’ comes from; the priests are called *kəmarim* (2 Kings 23:5) because their skin shrinks due to their asceticism and withdrawal from the passions and pleasures of the world.

Dictionaries for Post-Biblical Hebrew

Saadia composed glossaries to certain Mishnaic tractates, which gave an example to subsequent scholars.⁵² As for dictionaries, there are very few. Rav Hai Gaon’s *Kitāb al-Ḥāwī* was the first extant comprehensive dictionary covering both Biblical and post-Biblical Hebrew and Aramaic. This was a revolutionary dictionary in several aspects, one of which was its anagrammatic arrangement (see below). *Ḥāwī* contained the vocabulary of the whole of Hebrew and Aramaic literature from the Bible to the Talmud and Midrash, and in fact, of anything that was written up to Rav Hai’s time; it also, probably uniquely, encompassed names of peoples and places. Unfortunately *Ḥāwī* was lost in the fourteenth century, and only some of its vestiges were rediscovered in the Cairo Geniza towards the end of the nineteenth century.⁵³

Another post-biblical dictionary is *ʿArukh* composed by R. Nathan ben Yehiel of Rome.⁵⁴ *ʿArukh* seems to be a compilation of the lexicographical materials appearing in the post-biblical rabbinic literature discourse. It certainly relies on Gaonic literature, especially Hai Gaon’s *Responsa* and *Commentaries*, if not on *Ḥāwī* itself, but this question needs further investigation. Compared to *Ḥāwī*, *ʿArukh* is much more compact and arranged in simple alphabetical order; it uses Hebrew for its definitions. Though based on the old-fashioned root theory, these other features explain its survival.

Another post-biblical dictionary, the Hebrew–Arabic *Al-Muršid al-Kāfi*, was composed in the thirteenth century by Tanḥum Yerushalmi.⁵⁵ It covers the lexicon of Maimonides’ *Code* (*Mishne Torah*), based mainly on Mishnaic terminology, but also including Maimonides’ own innovations. In Yemen, a famous Hebrew–Arabic dictionary to Maimonides’ *Mishneh Torah* and his

⁵² See, e.g., the Geniza fragments described in Maman, *Otsrot Lashon*, 447, 459, 467, 474, 480–1, 498.

⁵³ For samples and a thorough description of *Ḥāwī*, see Maman, ‘Lə-darko šel Rav Hai Gaon’, and Maman, ‘Səridei’.

⁵⁴ Nathan, *Aruch completum* (1810–12), is an edition; the dictionary needs to be re-edited.

⁵⁵ Tanḥum, *Al-Muršid al-Kāfi* (2005), is an edition with translation from Judeo-Arabic into Hebrew.

commentary to the Mishnah, *al-Jāmiʿ* ('collection, thesaurus'), was composed between 1483 and 1486 by David ben Yeshaʿ, many copies of which survived in manuscripts from the fifteenth through the nineteenth centuries. Its entries are arranged according to the first letter (not according to their root).⁵⁶ In 1339, in what is now Turkmenistan, Shelomo ben Samuel composed *Sefer ha-Melīzah* (or *Sefer Pitronei Millim*; 'The book of translation'), a Hebrew (and Aramaic)–Persian dictionary covering words from biblical and post-biblical literature. It is similar to today's bilingual dictionaries.⁵⁷

The next important post-biblical dictionary was Eliahu Levita (Baḥur)'s *Meturgeman* which is a lexicon to the Aramaic of the Targums, or Aramaic versions of the Bible (both Targum Onkelos and Targum Jonathan), published in Italy in 1540. This is the first dictionary dedicated to Targumic Aramaic. Another lexicon of Levita's was *Tishbi*, published in 1541, covering 712 Talmudic and post-Talmudic Hebrew entries, some Aramaic, e.g., *ḥas* which renders biblical *ḥalīla* 'God forbid!', or words borrowed from foreign languages, arranged alphabetically. (The number 712 equals the arithmetical value of *Tishbi*, a biblical epithet of the prophet Elija, which Levita used to refer to himself.) The entries are defined in Hebrew and translated into German. Some entries are proper names, e.g., *Tiberia*, while some are medieval, borrowed from Arabic, for instance *ḥoqen* 'enema'. The definitions are short in general, except when the entry is encyclopedic, as in *Vitry* (old French Vitry-en-Perthois, where Simḥa ben Šemuʿel in 1208 completed his famous book of prayers called *Maḥzor Vitry*) and *Hoshaʿana* (hosanna). In 1648, David Cohen de Lara published a dictionary on the foreign words in rabbinic Hebrew-Aramaic literature, such as Greek, Latin, and Persian.

The Lexicography of Grammar, Rhymes, Homonyms, and Synonyms

The two books of grammar in which Yehuda Ḥayyūj dealt with the problem of the Hebrew root (see above) established the grammatical dictionary genre, listing groups of roots of the same type, arranged alphabetically. The discussion in each entry is only grammatical: the second-level classification in each entry is semantic, but Ḥayyūj provides no semantic definition. Other grammatical wordlists were compiled, such as that of R. Yehuda ibn Balʿam's *Kitāb*

⁵⁶ David ben Yeshaʿ, *Millon ʿivriy-ʿaraviy al-Jāmiʿ* (1988), is a facsimile, with detailed description by Y. Tobi (175–87) and U. Melammed (188–9).

⁵⁷ Fragments were published in Bacher, 'Hebräisch-persisches Wörterbuch', 1–76 (Hebrew part, i.e., second sequence of pagination).

Ḥurūf al-Maʿāni ('The book of particles'), which gives an exhaustive description of Hebrew function words. Such lists, arranged alphabetically, are also incorporated in grammar books, such as those of Abū al-Faraj and Ibn Janāḥ. Ibn Balʿam composed also a book on denominative verbs, *Kitāb al-Afʿāl al-Mushtaqqā min al-ʿAsmāʾ*.

Beside the main purpose of the dictionary as a tool for the exegesis of religious or ancient texts, medieval linguistics also intended to regulate the uses of language, one of which was obviously poetry. As mentioned earlier, the first Hebrew dictionary was Saadia's *Egron*, a dictionary designed for liturgical poetry. Over time, this genre developed into sophisticated rhyme dictionaries, many of which have survived in manuscript form, such as Shlomo Ben-Elia of Poggibonsi's *Sefer ha-Ḥaruzim* ('The book of rhymes'), composed around 1500, which is a technical elaboration of Yehuda Hashaʿari da Porta's *ʿaremat Ḥiṭṭim* ('pile of wheat'), of around 1400.⁵⁸ This genre evolved especially in Italy, where Hebrew poetry flourished during the Renaissance, but one finds rhyme dictionaries in other places too, for example in North Africa. Not all rhyme dictionaries reflect the influence of others: some were compiled independently, and their methods seem to be more sophisticated than those of their modern counterparts, because they relied on phrases from the Bible and on verses and authentic rhymes from classical poetry.⁵⁹

Two famous dictionaries of homonyms were composed in medieval Spain, one, *Kitāb al-Tajnīs*, by Rabbi Yehuda ibn Balʿam, and one, *ʿAnaq*, by Rabbi Moshe ibn Ezra.⁶⁰ *Kitāb al-Tajnīs* lists, analyses, and explains 314 groups of biblical homonyms ordered alphabetically. *ʿAnaq* is a series of 1,263 stanzas that Ibn Ezra composed rhyming two to four homonyms and using, altogether, 573 groups of biblical homonyms. The stanzas are organized according to themes, not alphabetically. Ibn Ezra did not explain the homonyms, but at a later stage Elʿazar ben Ḥalfon wrote a commentary to *ʿAnaq* and explained the differences between the homonyms with grammatical and semantic analyses.⁶¹

The first dictionary of synonyms in Biblical Hebrew was *Ḥotam Tokhnit* ('The seal of the well-built edifice'), composed by Abraham ben Isaac Bedersi in the second half of the thirteenth century.⁶² Another, which is very concise,

⁵⁸ Poggibonsi's dictionary is preserved in New York, Jewish Theological Seminary Library, MS 10612, and digitized images are available on the library's website; for da Porta's, see Maman, *Otsrot Lashon*, entries for MSS 4552, 4666, 8109.

⁵⁹ Maman, *Otsrot Lashon*, §9. ⁶⁰ Yehuda, *Kitāb al-Tajnīs* (1975), is an edition of the former.

⁶¹ Elʿazar, *Pitron* (2011), is an edition. ⁶² Bedersi, *Ḥotam Tokhnit* (1865), is an edition.

is *Ohel Mo'ed* of Solomon ben Abraham of Urbino, compiled in 1480 and published in Venice in 1548.⁶³

Definitions and Their Language

From the eighth century on, a huge majority of the Jews lived in Arab lands; that is why Saadia, his younger contemporaries David ben Abraham Alfāsi of Jerusalem and Ibn Quraysh of Tahort, and many others used Arabic for their definitions. Menaḥem ben Saruq, in contrast, though living in Arab-dominated Spain, was a poet and uniquely used Hebrew in his writings, including the definitions in his dictionary *Maḥberet*.

As for the nature of definitions, we have seen that in the glossaries they are very concise, not extending beyond the translations of the lemmata to the local languages. The same pattern is also found in Menaḥem's *Maḥberet*, though it is a systematic monolingual dictionary. Its definitions usually gave a synonym of the entry root or synonyms of each of its subentries. In a few hundred cases, he offered even less, assuming that the reader would derive the meaning from the verses presented in the entry: a pseudo-definer, *kəmašma'o* ('as it sounds', 'in its literal meaning'), or no definition at all. Other lexicographers tended to elaborate by discussing the grammatical and semantic aspects of the entry, though some, such as Ibn Janāḥ in his *Kitāb al-ʿUṣūl*, also used similar pseudo-definers, such as *ma'rūf* ('it is known', 'as is known').

Naturally, lexicographers related some lemmata to others when these presented similarities. The question whether a similarity was the outcome of homonymy or polysemy stood behind the split between two types of philologists: those who tended to list as many meanings as possible for a single word, 'meaning maximalism', in the terms of Richard Steiner, versus those who did the opposite and made every effort to put together as many words as possible (even those which were seen as homonyms by others) under the same basic meaning, 'meaning minimalism'.⁶⁴ Examples of the former are Rav Saadia Gaon and Ibn Janāḥ, and of the latter, Rashi, Menaḥem, and Kaspi.

Etymology and Comparative Semitic Philology

The field of Semitic comparative philology as a whole was clearly established by Jewish scholars, mostly by virtue of their being multilingual. Arabic was

⁶³ A more modern edition was published in 1881. ⁶⁴ Steiner, 'Saadia vs. Rashi'.

the language of their everyday speech, Hebrew and Aramaic filling their diverse cultural needs. Saadia had already used Hebrew–Arabic and Hebrew–Aramaic comparative philology as a means to explain the lemma under discussion, mostly in sporadic notes throughout his translation and commentary of the Bible. But what was sporadic in Saadia’s oeuvre became systematic in that of his younger contemporary, Yehuda ibn Quraysh, who composed three comparative dictionaries as parts of his book *Risāla*: a Hebrew–Aramaic comparative dictionary, a Hebrew–Arabic comparative dictionary, and a Hebrew–Hebrew dictionary comparing biblical forms with their rabbinic counterparts. The reason for writing the Hebrew–Aramaic dictionary, according to the author’s foreword, is that he heard that the Jews of Fez had ceased the ancient traditional custom of reading the Targum Onkelos along with the weekly Torah portion in the Hebrew original, arguing that they did not understand Aramaic any more. Ibn Quraysh sought to convince them that Aramaic could still illuminate Hebrew, as it is its cognate. And since he dealt with one Semitic cognate, Aramaic, he went on to deal with a second cognate, Arabic, and with other aspects of comparative philology.

This area of activity went on developing until it reached its peaks in mid-tenth-century Spain, with R. Yona ibn Janāḥ’s lexicographic work, and afterwards, in the first half of the twelfth century, with R. Isaac ibn Barūn’s Hebrew–Arabic comparative dictionary, *Kitāb al-Muwāzana bayna al-luġa al-Ibrāniyya wal-ʿArabiyya* (‘The book of comparison of the Hebrew language with Arabic’).

But not all Hebrew philologists were willing to practise comparative philology. Some opposed it in principle, such as Menaḥem ben Saruq and his disciples, who argued that if language comparison were permissible, this implied that the compared languages were equal, and thus might complement each other. That is, anything lacking in one language could have been completed from the other. In fact, in the late nineteenth century, Eliezer Ben-Yehuda (for whom see Chapter 20) proposed to use Arabic words to fill the gaps in Hebrew. But though he succeeded in borrowing a few Arabic words, his opinion was rejected by the Council of the Hebrew Language, an institution founded in 1890.

In any case, medieval Hebrew philologists should be regarded as pioneers in using comparative Semitic philology for their lexical definitions. This served as a foundation for nineteenth-century comparative Semitic philology, as exemplified by Wilhelm Gesenius, who incorporated earlier scientific achievements in comparative Semitic linguistics in his dictionary alongside his own contributions to the field.⁶⁵

⁶⁵ Bajema, ‘Jewish grammarians in the lexicography of Wilhelm Gesenius’.

Arrangement of Entries

The lemmata in the glossaries were naturally arranged according to the flow of the text they were explaining. A glossary to the Pentateuch was arranged according to book, chapter, and verse, and so for any other biblical book. In the same way, glossaries to the Mishnah and the Talmud were arranged according to tractate, chapter, and mishnah. However, dictionaries such as Saadia's, Ben Saruq's, and Alfāsi's were arranged in the alphabetical order of the roots of their entries, of course according to each lexicographer's perception of the root.⁶⁶ As we have seen, Saadia's *Egron* entries were arranged two ways, both alphabetical: one according to the first letter of the lemma and one according to its last letter. This served as a model for subsequent rhyme dictionaries (see above). Most of the others only used the former method.

Yet, there was a different lexicographical arrangement. Rav Hai Gaon, the last Babylonian Gaon, edited his *Kitāb al-Ḥāwī* 'The comprehensive book', the anagrammatic way.⁶⁷ For example, at the opening of the *ḥṭl* headword, Rav Hai presents the six following roots in a row: *ḥṭl*, *ḥlt*, *ṭlh*, *ṭhl*, *lṭh*, *lḥt*. This presentation shows that the discussion will not be confined to entries derived from *ḥṭl* alone but will embrace all the entries derived from its permutations (any sequence of three different characters has six possible permutations).⁶⁸ If we applied the anagrammatic arrangement to an English dictionary, we would have, for instance, under the root PRT, the sequences PRT (*part*, *port*, *porter*, etc.), PTR (*patter*, *peter*, *Petra*, etc.), RPT (*repeat*, etc.), RTP (unused), TPR (*taper*, *tapir*, etc.), and TRP (*trap*, etc.). Though the anagrammatic method was already presented in *Sefer Yešira*, an enigmatic short text on which Saadia wrote a commentary, Rav Hai probably adopted it from Arab lexicographers, namely from al-Ḥalīl bin Aḥmad's *Kitāb al-ʿAyn* (see Chapter 8), the first Arabic dictionary ever written.⁶⁹ Yet whereas al-Ḥalīl organized the Arabic alphabet in phonetic order, Rav Hai left the traditional Hebrew alphabetical order intact. In what remains of Rav Hai's dictionary, three direct quotations from *Kitāb al-ʿAyn* are found, leaving no doubt that Rav Hai was familiar with al-Ḥalīl's work. It seems that he also knew of other anagrammatic Arabic dictionaries that used traditional alphabetical order, such as Ibn Durayd's *Ġamharat al-luġa*, though it is not certain that he was familiar with al-Azharī's *Tahḏīb al-luġa*.⁷⁰ This is an example of direct cultural

⁶⁶ For a detailed description see Maman, 'Menaḥem ben Saruq's *Mahberet*'.

⁶⁷ See Maman, 'Sārīdei', 344. ⁶⁸ Téné and Barr, 'Linguistic literature', §2.6.1.

⁶⁹ Kopf, *Studies in Arabic and Hebrew Lexicography*, Hebrew section, 117.

⁷⁰ For both of these, see Chapter 8 and Kopf, *Studies in Arabic and Hebrew Lexicography*, Hebrew section, 119.

influence of Arabs on Jews in this literary genre. When it comes to learning from others, Rav Hai had no reservations about cultural contacts with hetero-religionists, as can be seen also in his relations with the Catholicos of Baghdad.⁷¹ In fact, there was another anagrammatic list of roots composed by Abū l-Faraj Harūn as part of his grammar *Kitāb al-Muštamil* a few decades after Rav Hai, but it had a different purpose and was not meant to be a dictionary.⁷²

To conclude, the Hebrew (and Aramaic) lexicographic project began in the definitions that are scattered throughout the rabbinic discussions in the Mishnah and in the Talmud, and on a small scale even in the Bible itself. In a later period, glossaries were composed to the entire Bible or some of its books, as well as to the Mishnah and Talmud or some tractates thereof. Systematic Hebrew lexicography, however, began to be established in the form of comprehensive lexica from the time of Saadia. Whereas the glossaries, and some dictionaries, were intended to serve the direct understanding of a given text, there were also dictionaries written for specific purposes, such as the grammatical or homonym dictionaries.

⁷¹ Maman, 'Səridei', 368.

⁷² Maman, 'Lə-darko šel Rav Hai Gaon'; Maman, 'Seder u-mašma'ut bə-ʔotiyot ha-šoreš'.

The Chinese Periphery to c. 1800

MÅRTEN SÖDERBLOM SAARELA

This chapter will treat the lexicographies of eastern Inner Asia, Japan, and Korea that were directly influenced by the Chinese tradition.¹ The first section will begin by discussing lexicography in the Inner Asian empires in the early second millennium AD, with a focus on the lexicography of the Tangut and Mongols in North China; and will then turn to the lexicography of Inner Asian languages in the Chinese Míng empire that succeeded the period of Inner Asian political dominance in North China; and finally to the lexicography of Manchu and Mongolian in the Inner Asian Qīng empire, which conquered the Míng and ruled from Beijing. The remaining two sections of the chapter will treat lexicography in Japan and Korea.

China and Inner Asia

The roughly four hundred years from the tenth century to the fourteenth witnessed the invention, adoption, and adaptation of several scripts in Inner Asia. Several of these scripts had fallen into disuse by the end of the period, some being only partially deciphered to this day. Yet their long-term effects were significant; the invention of the Korean alphabet in the fifteenth century is hardly conceivable were it not for these earlier Inner Asian experiences, nor is the adoption of an ultimately western Asian script by the Manchus in the early seventeenth. The only dictionaries that remain for languages written in the new Inner Asian scripts are dictionaries of Tangut produced in the Xī ('western') Xià, or Tangut empire, which lasted from the 980s to 1227. The Tangut script was reportedly created in 1036 on imperial command.² As in Chinese, every character represented one spoken syllable, which also often

¹ I would like to thank Peter Kornicki and Si Nae Park for sharing references and their own unpublished works with me.

² Galambos, *Translating Chinese Tradition*, 120.

corresponded to a morpheme.³ As was probably the case among the Jurchen (for whom see below), lexicography played an important role in establishing the new written language that formed through translations from Chinese and Tibetan.⁴

Literary Chinese had probably been the main written language in use in the Tangut state before the invention of the indigenous writing system, and initially Tangut literates probably also knew to read Chinese. The process through which they did this might have involved paraphrasing in Tangut (similarly to how literary Chinese was famously read in Japan using *kundoku*), meaning that the border between Chinese and Tangut was fluid. Indeed, the Tanguts modelled their own language pedagogy on the Chinese tradition, reportedly translating primers and linguistic works. The latter are supposed to have included a translation of the *Ēryǎ* (see Chapter 3).

The dictionaries compiled by the Tanguts included the bilingual Tangut–Chinese *Timely Pearl in the Palm*, the monolingual Tangut *Sea of Characters*, the encyclopedia (*lèishū*) *The Sea of Meanings Established by Saints*, and a thematically arranged glossary included in the primer *Miscellaneous Characters*.⁵ The *Timely Pearl in the Palm* was a ‘glossary of basic terms and phrases, divided into three major chapters: Heaven, Earth and Man ... It provides a basic grammar through the use of sentences each having about four Hsi Hsia [Xi Xia] characters on the average.’⁶

Lexicographic texts from the Tanguts’ contemporaries, the Jurchen Jīn (Jurchen emperors ruled in China from 1114 to 1234) are extant only from the Míng empire’s interpreters’ bureau, and will be discussed in that context presently.

When the Mongols conquered first their Inner Asian neighbours and then China in the thirteenth century, they committed their language to writing for state-building purposes. Unlike the Tangut and Jurchen, the Mongols of the Yuán empire (1271–1368) at first adopted an existing script used by the Uyghurs to write their language.⁷ Scribes at the Mongol court wrote and translated a lot in many different languages, but no works of lexicography in the Uyghur–Mongol script have survived from the Yuan period.

What does survive are glossaries and a dictionary making use of the ‘Phags-pa script, which is named after its originator, the ‘Phags-pa Lama,

³ I infer this from Kychanov, ‘Tangut’, 228, who states the matter in different terms.

⁴ Galambos, *Translating Chinese Tradition*, 123.

⁵ Galambos, *Translating Chinese Tradition*, 4–5. ⁶ Kwanten, *Timely Pearl*, 3.

⁷ Kara, ‘Aramaic scripts’, 539–40; Kara, *Books of the Mongolian Nomads*, 27–9; Brose, ‘Uyghur technologists of writing’.

a Tibetan cleric in the service of Kubilai (reigned 1260–94). His alphasyllabic script, structurally similar to – and influenced by – Tibetan writing, was intended as a state script for the Mongol empire.⁸ The introduction of the ‘Phags-pa script followed the Inner Asian precedent of marking the creation of a new state with a new written language.

Chinese or Chinese-speaking officials needed lexicographical resources to learn to compose Chinese-language texts written in the ‘Phags-pa script.⁹ These materials served to allow readers of Chinese to find the ‘Phags-pa equivalents of Chinese syllables, each represented by one Chinese character. Accordingly, they arranged Chinese characters in ways familiar to Chinese literates. *Bǎijiā xìng* 百家姓 (‘The hundred surnames’) presented a list of Chinese surnames in a fixed order. The list was memorized by children as they learned to read. It was turned into a ‘Phags-pa glossary by adding a ‘Phags-pa transcription of each Chinese character. Since ‘Phags-pa, unlike the Chinese characters, recorded sound on the subsyllabic level, transcribing the Chinese graphs into it meant defining their pronunciation, which probably represented the educated standard of pronunciation in North China at the time.

A manuscript dictionary transcribing Chinese syllables into ‘Phags-pa script is also extant. It was evidently based on earlier handbooks that are now lost. Some of these books influenced Chinese-script lexicography: *Gújīn yùnhuì jǔyào* 古今韻會舉要 (‘Lifted essentials from the “Confluence of rhymes from past and present”’, 1297), as well as the lost work of which it was an abridgement, was based on canonical Chinese rhyme dictionaries, but recast on the basis of Sino-‘Phags-pa handbooks. The extant ‘Phags-pa-script dictionary, also based on these earlier books, is titled *Měnggǔ zìyùn* 蒙古字韻 (‘The Mongol-script rhymes’). It is undated; the extant manuscript might date from as late as the eighteenth century.

Měnggǔ zìyùn arranged the Chinese syllable lemmata according to the common arrangement in Chinese rhyme books used for the composition of poetry. The sound system on which those books were based was very different from the Chinese spoken in North China in the thirteenth century and represented in the ‘Phags-pa transcriptions. Chinese literates would have memorized the sequence of rhymes and could use it to locate characters. Within sections, syllables were arranged first by the spelling of their final in the ‘Phags-pa transcription. The syllables with the same orthographic finals

⁸ Mote, *Imperial China*, 483–4.

⁹ Unless otherwise indicated, my source for the ‘Phags-pa lexicographical materials is Coblin, *Handbook of ‘Phags-pa Chinese*, ch. 1.

were then divided into groups according to the spelling of the whole syllable, so that syllables that had both the same initial and the same final were listed together. Among these, finally, homophonous syllables were grouped by tone. Tone was not indicated in 'Phags-pa, only by its Chinese name, written at the head of every group. *Měnggǔ zīyún* indicated how to transcribe Chinese characters into the 'Phags-pa script. It did not include any semantic information.

The Chinese Míng dynasty (1368–1644), which took control over China when Mongol power collapsed, compiled glossaries for politically important foreign languages. In 1382–9, scholars working for the court compiled a Chinese–Mongolian glossary, and in 1407–8, foreign-language lexicography was given a new institutional foundation.¹⁰ The glossary that was printed in 1389 had the title *Huáyí yìyǔ* 華夷譯語 ('Sino-xenic translated terms') and contained 840 Mongolian phrases with Chinese translations, arranged by topic. Such glossaries were not unprecedented in China; the medieval tradition of Sanskrit–Chinese glossaries (see Chapter 6) was similar. To better note the pronunciation of the Mongolian words in *Huáyí yìyǔ* using Chinese characters, the compilers supplied them with diacritics in the form of smaller characters that specified aspects of the intended articulation.¹¹ Similar compilations for other languages followed, compiled by both the Translators' and the Interpreters' Institutes. The Translators' vocabulary contained Jurchen script in addition to Chinese translations and transcriptions. '[T]hese vocabularies had a limited aim: to be able to communicate, on a basic level, with "barbarians" on the rare occasions when this was absolutely inevitable, as when they brought tribute to the Court.'¹²

The Manchus instituted a language closely related to Jurchen as their language of state in the early seventeenth century, writing it down using the Uyghur-Mongol script. In 1644, the Manchus conquered Beijing and subsequently all of China, which they ruled until 1911 under the name of Qīng. During the first two centuries of their ascendancy, the Manchus incorporated first the eastern and then the western Mongols into their polity. The major languages of this Inner Asian empire were described in dictionaries.

Manchu was the language of the imperial family and the centre of the new Qīng lexicography. Manchu–Chinese and Chinese–Manchu dictionaries appeared in print for the first time in the 1680s and 1690s. The early

¹⁰ Hucker, *Dictionary of Official Titles*, 263–4 (item 2889) and 448 (item 5656).

¹¹ Nie and Sun, *'Xīfān yìyǔ' jiàolù jí huībīān*, 4–9. ¹² Kane, *Sino-Jurchen Vocabulary*, 100.

dictionaries were arranged either topically, which worked in either language, or according to the structure of the Manchu script, which was possible only if the dictionary's leading language was Manchu. The topically arranged dictionaries do not appear to have been influenced by the Jurchen vocabularies of the Míng interpreters: those works had a limited circulation already at the time of their composition, and they were not known to the early lexicographers of Manchu. The first Manchu dictionary known to have been printed is *Dà Qīng quánshū* 大清全書 | *daicing gurun-i yooni bithe* ('Complete book of the Great Qīng', 1683), which drew on earlier manuscript materials. In it, Manchu words and phrases were arranged in an alphasyllabic order derived from the syllabary commonly used to teach the Manchu script to both Manchus and Chinese.¹³

In 1708, the imperial court published a topically arranged dictionary that, unlike almost every other Manchu dictionary, was monolingual. Instead of translations, *Han-i araha manju gisun-i buleku bithe* ('Imperially commissioned Mirror of the Manchu language') contained simple definitions in Manchu, often accompanied by examples from Manchu translations of the Chinese Confucian classics. The book was intended to regulate Manchu and to elevate its status to that of a literary language on a par with Chinese.¹⁴ In 1717, the monolingual *Buleku bithe* was made into a bilingual dictionary by the addition of Mongolian translations. Several more dictionaries in this tradition appeared during the eighteenth century. A Manchu–Chinese one from 1771–3 circulated widely, whereas the five-language version containing translations into Chinese, Mongolian, Tibetan, and Uyghur is perhaps better understood as a monument to the Manchu vision of a universal empire than a reference work in actual use.¹⁵

Indeed, the imperial language of Manchu was what brought all of these languages together. As a rule, Chinese did not occur together with Mongolian without the mediation of Manchu, and Tibetan often occurred through the mediation of Mongolian. Accordingly, a separate lexicographical tradition developed in the Tibetan–Mongolian interface, with which Manchu–Chinese lexicography had no relationship. Meanwhile, Manchu entered the Chinese lexicographical tradition, with Chinese–Manchu dictionaries appearing in both manuscript and print in the eighteenth century that arranged Chinese characters according to the radical-and-stroke order seen in

¹³ Ji, 'Dà Qīng quánshū yánjiū'. ¹⁴ Jiāng, *Kāngxī 'Yǔzhì Qīngwén jiàn'*.

¹⁵ Corff et al. (eds.), *Aufkaiserlichen Befehl erstelltes Wörterbuch* is an edition and translation.

Chinese dictionaries such as *Zihui* and *Kāngxī zidiǎn*. Manchu lexicography continued to develop in China into the nineteenth century.

Japan

Japanese is an agglutinative, verb-final language that is typologically very different from Chinese. Yet the Chinese script, language, and written tradition were very influential on the development of Japan's literary culture, including its lexicography. Chinese written culture was mediated in part through Korea and through Buddhism.¹⁶ Lexicographic activity is inferred – but there are no extant dictionaries or wordlists – for the eighth century. Extant dictionaries date from the following Heian period (794–1185), with new developments during the later periods of warrior rule, social change, and new interactions with the continent and eventually the European powers.

Reading and writing in Japan is a complicated matter. Literary Chinese could be read as Chinese, character by character and employing the received Chinese pronunciation of the characters. This practice is called *ondoku* 音読, 'reading by voice', and was tantamount to reading Chinese with a heavy Japanese accent. However, the same texts could be, and were, read in Japanese through 'reading by gloss', or *kundoku* 訓読: Chinese characters were associated with Japanese words; the order of the words was changed in accordance with Japanese syntax; and grammatical elements were added to the Chinese text.¹⁷ (There was a similar practice in medieval Korea.) The addition of reading marks helped readers to produce a correct *kundoku* rendering.

Yet from the period of the earliest received texts written in Japan, Chinese characters were also used for their sound value (in most cases derived from their Chinese pronunciation, perhaps through Korean mediation) to write Japanese words. Chinese characters thus used are called *man'yōgana* 万葉仮名, named after an eighth-century collection of poetry, *Man'yōshū* 万葉集, in which Japanese is written in this way.¹⁸ *Man'yōgana* was used in Japanese lexicography from the beginning, and the association of Chinese characters to both Chinese-derived pronunciations and Japanese words remained in the writing system. Early dictionaries were arranged on the Chinese model, either by radical, by (Chinese) rhyme, or by subject matter. With the development of the *hiragana* and *katakana* syllabaries – derived from, but also

¹⁶ Lurie, *Realms of Literacy*, chs. 2–3. ¹⁷ Lurie, *Realms of Literacy*, 175, 183–4.

¹⁸ Müller-Yokota, 'Abriss', 7–14; Osterkamp, 'Sinoform writing', 115.

radically different from, Chinese characters – there were new possibilities of lexicographic glossing and arrangement.

The earliest reference to lexicographic activity in the Japanese historical record is an entry in *Nihon shoki* 日本書紀 ('Annals of the book of Japan'), the country's second-oldest chronicle, saying that in the eleventh year of the reign of Tenmu 天武 (682) the emperor ordered the compilation of 'Niina 新字 in forty-four scrolls'. Translating as 'new names' or 'new characters', the book is not extant. It was probably related to an ongoing project to compile a history and may have represented an attempt to regulate Chinese character forms and Japanese glossing practices.¹⁹

Preceded by phonetic and semantic glossaries (*ongi* 音義) to Buddhist texts, the earliest extant dictionaries compiled in Japan were dictionaries of Chinese characters rather than dictionaries of the Japanese language. The Buddhist monk Kūkai, who travelled in China, compiled *Tenrei banshō meigi* 篆隸万象名義 ('Words and meanings of the myriad manifestations in seal and clerical [in this case meaning *kāi* 楷, standard] script') in c. 830 (extant copy from 1114). The book is a Chinese dictionary, with lemmata consisting of single Chinese characters, based on *Yùpiān* (see Chapter 3) and retaining its arrangement by radical.²⁰

A few generations later, another monk, Shōjū, compiled *Shinsen jikyō* 新撰字鏡 ('Newly composed mirror of characters'; 898–901). Unlike Kūkai's book, it contained Japanese glosses, but only for a minority of entries. The lemmata are, again, single Chinese characters that are arranged under a reduced set of radicals. Within sections, a secondary arrangement, either by meaning or by Chinese pronunciation, is apparent, probably reflecting the arrangement of the sources.²¹

The writing of Chinese verse was an integral part of elite sociability in Heian Japan. The need to respect Chinese poetic conventions motivated the compilation of rhyme books (*insho* 韻書). Unfortunately, from this early period, only the titles remain.²²

Other early Japanese dictionaries were arranged by topic. Not everybody would agree that all of these books are dictionaries in the current sense of the word. In terms of genre, they belong to a continuum that also contains encyclopedias in the Chinese tradition of 'classified writings' (Chinese *lèishū*, Japanese *ruisho* 類書). Encyclopedias, which were read for facts, were indeed

¹⁹ Kawase, *Kojisho*, 15 and 20–1 (cf. 13, where *Niina* is glossed as *Shinji*); Bailey, 'Early Japanese lexicography', 1–2.

²⁰ Sugimoto, *Jisho, jiten*, 20. ²¹ Bailey, 'Early Japanese lexicography', 3–4.

²² Sugimoto, *Jisho, jiten*, 22–3.

also compiled at this time. *Wamyō ruiju shō* 和名類聚抄, finished in 934, was, however, a lexicographic product. This encyclopedic work is a dictionary of Chinese characters, giving character forms and Japanese and literary Chinese glosses, directed at the female literary circles at the imperial court. The thematic arrangement differs between manuscript traditions, but it was probably inspired by the arrangement of Chinese encyclopedias such as *Yiwén lèijù* 藝文類聚 ('Categorized collection of arts and letters'; Japanese *Geimon ruijū* – note the similar title). Entries in *Wamyō ruiju shō* consist of the Chinese lemma; Chinese commentary and occurrences of the word in Chinese sources; Sino-Japanese reading given in *fǎnqiè* (for which see Chapter 6), supplemented by other phonological information; Japanese translation in *man'yōgana*; and occurrences of the Japanese word in Japanese texts. Given the nature of the Japanese sources, the Japanese vocabulary is often literary. Yet at times the book also gives 'vernacular' (*zoku* 俗) glosses drawn from the spoken language of the time. The matching of Chinese words with Japanese shows an awareness that Chinese and Japanese are two separate languages with separate literary traditions. *Wamyō ruiju shō* was very influential, going through many printed editions in the early modern period.²³

The *kana* syllabaries were invented during the ninth century. From the middle of the Heian period, furthermore, official contacts with China ceased. These developments coincided with an interest in Japanese verse (*waka* 和歌) among the elite, which was reflected in lexicography. The syllabaries and Japanese poetry were both consequential for the development of Japanese lexicography.

The two Japanese syllabaries *hiragana* 平仮名, variously translated as 'ordinary [or "completely"] borrowed characters' or characters 'without angles' and *katana* 片仮名, 'partially borrowed characters' or 'simple, incomplete' characters, were both derived from the Chinese script used as *man'yōgana*. Chinese characters – in whole, as simplified cursive forms, or in part – were used to represent only sound. In their contemporary standardized form, the two syllabaries are structurally identical. They can be used to represent words by themselves or to provide phonetic glosses to words written in Chinese characters. In Japanese texts they are used together with Chinese characters according to the demands of the genre.²⁴

²³ Karow, 'Wörterbücher der Heianzeit'; Sugimoto, *Jisho, jiten*, 26–8.

²⁴ Müller-Yokota, 'Abriss', 17–31; J. S. Smith, 'Japanese writing'; Eschbach-Szabo, 'Réflexion linguistique au Japon', 459–60.

By the twelfth century, the syllabaries were well established. The spoken Japanese language had meanwhile developed away from the language of poetry. The syllabaries were presented in several fixed orders, eventually reduced to two in common usage. The arrangement of lexical material might seem an obvious reason to establish a fixed order of the syllabic characters. However, one of the original usages of at least *katana* was the phonetic glossing of Chinese-character texts, and the difficulties of interpreting phonetic information encoded in Chinese characters also motivated the organization of the syllabary into a certain order intended to facilitate that process, so lexicographical arrangement was not necessarily the original motivation.²⁵

Dictionaries of literary Japanese words used in poetry were compiled in the eleventh and twelfth centuries and arranged by theme.²⁶ *Iroha jiruishō* 色葉字類抄 ('Characters classified in *iroha* order and annotated') appeared at the same time, but was different. The book was first compiled in 1144–5 but then expanded upon for several decades; unsurprising in a literary culture that was still dominated by manuscripts. This book was arranged – to a certain extent – according to one of the fixed orders of the syllabary that later became canonical in lexicography. The top level of the dictionary is arranged according to 'Iroha uta' いろは歌, a song or poem (*uta*) thus named for its first three syllables. The poem is not interesting for its literary qualities, but for the fact that each of the syllabic signs in use at the time of its composition in 1079 occurs in it, and occurs there only once. In the context of the poem itself, the three syllables – written in *hiragana* as いろは – are to be interpreted as '[As for] the colours'.²⁷ The two Chinese characters chosen to represent the phrase in the title of the dictionary translate as 'coloured leaves'. In later dictionaries, the phrase was written as 伊呂波, using three Chinese characters that cannot be read for their meaning but only for their sound value, *iroha*.

Iroha jiruishō contains forty-seven sections, each containing Japanese words beginning with the same syllable. Within each of the top-level sections, the dictionary uses a secondary arrangement by topic. This mixed arrangement was very influential in later periods. The dictionary's lemmata consist of Chinese characters, accompanied (but not always) by their Sino-Japanese and native Japanese readings. 'It is not . . . a work used to learn the meanings of Japanese words'; rather, it is 'one for looking up Chinese characters in terms of their Japanese pronunciation'. From the point of

²⁵ Maeda, 'a i u e o to iroha', 16–22; cf. Müller-Yokota, 'Abriss', 35.

²⁶ Bailey, 'Early Japanese lexicography', 11. ²⁷ Müller-Yokota, 'Abriss', 38–9.

view of its intended use, the mixed arrangement made sense. Previously, in order to look up how to write Japanese words using Chinese characters, thematically arranged dictionaries had been the only option, as neither dictionaries arranged by Chinese character shapes nor dictionaries arranged by Chinese rhyme could be used for this purpose. The addition of forty-seven syllabic divisions onto the thematic sections made the latter easier to peruse.²⁸

When linguistic and literary developments in China led to new rhyme dictionaries for poetic composition and appreciation being compiled there (see Chapter 6), these dictionaries entered Japan. Rhyme books made in Japan, such as *Shūbun inryaku* 聚分韻略 ('Outline of rhymes assembled and divided [into categories]'), which was printed in 1307, provided Japanese readers, faced with all this variety, with a standard to abide by when writing Chinese verse (*Kanshi* 漢詩). Similar to some innovative mainland dictionaries of the period, *Shūbun inryaku* had a hybrid phonological-semantic arrangement. On the top level, the book was organized by Chinese rhyme, to which the tone was considered integral. Thus, first 'entering'-tone rhymes were listed, followed by the rising tone, and so on. In later versions, the three tones that were defined by their tonal contour were stacked in three rows on the same page, with the entering-tone rhymes, characterized by their unique stop finals, listed separately at the end (see Chapter 6 for Middle Chinese tones).²⁹ This arrangement earned the dictionary the alternative appellation *Sanjūin* 三重韻 ('The three stacked rhymes'). A similar arrangement was seen in Korean rhyme books. In both cases, the utility of the arrangement came from the fact that the stop finals were, unlike in Mandarin Chinese, retained in Sino-Korean and Sino-Japanese readings, meaning that the entering-tone characters could be identified easily by ear. Within each rhyme, finally, the characters were listed by theme. The simple glosses constituting the entries were written in literary Chinese, as was the whole dictionary.³⁰

Hyōta jiruishō 平他字類抄 ('Characters in the even and other [tones] divided up and annotated'), from 1388 or earlier, was similar to *Shūbun inryaku* in that it served to teach Chinese poetical conventions to Japanese readers. This new book was arranged primarily by theme and in one section by *iroha* order, and contained didactic features that made it useful for a new

²⁸ Quotations from Bailey, 'Early Japanese lexicography', 15–18; see also Sugimoto, *Jisho, jiten*, 31–3.

²⁹ For the terminology surrounding Chinese rhymes, see Li, 'Rime (rhyme)'.

³⁰ Sugimoto, *Jisho, jiten*, 39–40.

generation of aspiring *Kanshi* writers (courtiers, Buddhist clergy, warriors) who wanted to acquire vocabulary for their poetic compositions.³¹

Shūbun inryaku influenced later dictionaries, including *Kagakushū* 下学集 ('Collection of mundane matters') from 1444. This thematically arranged dictionary of Chinese characters with Japanese sound glosses and Chinese notes was 'intended to serve as a small encyclopedia and textbook'.³² It is another testament to the expansion of literacy in the period. Another dictionary, *Setsuyōshū* 節用集 ('Collection of words for everyday use'; second half of the fifteenth century), is similarly reflective of the practically oriented, widespread literacy of the period. This book, arranged first by *iroha* and then, under each syllable, by topic, 'is essentially a dictionary for looking up Chinese characters (*jibiki*), almost devoid of explanations of the lemmata'. The *Setsuyōshū* was printed in the sixteenth century, and over the next three centuries, more than 180 editions of the dictionary were printed. *Setsuyōshū* even became a generic term for 'dictionary'. Novels and satirical verse of these centuries testify to the ubiquity of dictionaries in the *Setsuyōshū* lineage; in scenes from this literature, a landlord might take one out and read difficult characters to his tenants.³³ The ease of reference contributed to the popularity of *Setsuyōshū*. Yet some experimented with new forms of lexicographic arrangement. *Onkochishinsho* 温故知新書 ('Mastery of the old and inquiry into the new') from 1484 was similar to the contemporary *Setsuyōshū* in that it used a composite arrangement of syllables on the top level followed by topic on the second level. Yet whereas *Setsuyōshū* – like both earlier and later dictionaries – arranged the syllables according to the order of the *iroha* poem, *Onkochishinsho* used the 'fifty-syllable table' (*gojūonzu* 五十音図).³⁴

As its name indicates, the fifty-syllable table was not a linear presentation of the Japanese syllabary. It was a two-dimensional grid, where the horizontal axis listed syllable initials and the vertical axis listed rhymes. Isolated Japanese syllables are always open, so the vertical axis listed the language's five vowels. The table was a product of Chinese phonological studies in Japan and the study of Indic scripts there. That might seem like a dual origin, but the Chinese phonological tradition had also at least in part developed under Indic influence, which is definitely in evidence in the development of 'rhyme tables' (Japanese *tōinzu* 等韻図; see Chapter 6).³⁵ These followed from the

³¹ Bailey, 'Early Japanese lexicography', 32–4; Sugimoto, *Jisho, jiten*, 39.

³² Bailey, 'Early Japanese lexicography', 37.

³³ Sugimoto, *Jisho, jiten*, 38, 51; see also Bailey, 'Early Japanese lexicography', 42.

³⁴ Sugimoto, *Jisho, jiten*, 38; Bailey, 'Early Japanese lexicography', 47.

³⁵ Simmons, 'Rime tables', 610.

encounter with tabular presentations of Indic scripts in which basic alpha-syllabic characters were paired with diacritics. In Japan as in China, Chinese rhyme tables were used to interpret *fǎnqiè* spellings, which were difficult for Japanese readers to interpret. The tables were useful in this regard, as they allowed the reader to run through either alliterating or rhyming syllables to find matching pairs. The earliest fifty-syllable tables pre-date the earliest rhyme table, but are similarly structured and fulfil the same function: they allow the reader to interpret spellings of sound glosses. The most striking difference is that the fifty-syllable table is much simpler – similar to the Indic-script tables – and lists *katakana* rather than Chinese characters.³⁶ On the other hand, it was much more analytic and learned than the *iroha*, which was recited by children, and *Onkochishinsho* did not get it to replace *iroha* in lexicographic arrangement. In our day, however, the ‘fifty syllables’ are frequently used.

Most of the Muromachi-period (1336–1573) dictionaries discussed thus far – the widely circulating and influential *Shūbun inryaku*, *Hyōta jiruishō*, and *Setsuyōshū*, as well as *Onkochishinsho* – had an arrangement mixing phonological order, either by Chinese rhyme or by the Japanese syllabary, with organization by subject matter. By contrast, *Wagyokuhen* 和玉篇 (‘Japanese jade chapters’), of which the oldest extant copy dates from 1489, did not.³⁷ Like its sixth-century Chinese namesake *Yùpiān* (see Chapter 3), it arranged Chinese characters by radical. More than thirty early manuscript copies are extant. It was printed in movable type in 1605, followed by more than forty editions thereafter. Entries were simple: single Chinese characters (and a few compounds inherited from earlier dictionaries) were glossed with their Sino-Japanese and Japanese readings. Yet thematic arrangement was so popular that it made its way into the *Wagyokuhen* tradition as well: some manuscripts imposed a thematic arrangement over the radical-based one, and other versions removed the subject headings, but grouped the radicals according to their ascribed meaning.³⁸

Books in several languages continued to be brought to Japan after the period of civil war (second half of the fifteenth century to 1600), as Japanese society, with the onset of peace under the Tokugawa shogunate (1600–1868), saw social and cultural developments that included expanded literacy, a surge

³⁶ Maeda, ‘*a i u e o* to *iroha*’, 18–19; Müller-Yokota, ‘Abriss’, 39–41.

³⁷ The Romanization of the title follows Aoki, ‘*Jisho*, *sakuin*’, 254. Historically, *Wagyokuhen* has also been called *Wagokuhen*: see Bailey, ‘Early Japanese lexicography’, 30; Kawase, *Kojisho*, 683.

³⁸ Aoki, ‘*Jisho*, *sakuin*’, 254–5; see also Sugimoto, *Jisho*, *jiten*, 35–6.

in printing, and the appearance of new genres and linguistic registers in writing. A lot of the writing and publishing activity of these centuries can be understood as forms of translation: between classical Japanese and contemporary forms of the language, or translations into Japanese from literary Chinese, written vernacular Chinese, or Dutch.³⁹ Dictionaries and glossaries were, naturally, of great utility in such a situation. Yet apart from the introduction of new foreign languages such as Dutch or Manchu, the period was perhaps characterized not so much by changes in type or format of dictionaries as by a big change of scale. Earlier dictionaries were printed, for the first time or in unprecedented quantities, reaching a new readership.⁴⁰

Before their activities were successfully halted, the Jesuits at Nagasaki printed both monolingual Japanese and Japanese–Portuguese dictionaries using movable type; however, as products of missionary linguistics, they fall outside the scope of the present chapter (for them, see Chapter 29). These dictionaries were not particularly influential on the development of Japanese lexicography in the Tokugawa period.

Contemporary developments in Chinese lexicography were more consequential for Tokugawa dictionary users. The practice of reading literary Chinese texts ‘by gloss’ meant that bona fide Chinese books could be adapted for an elite Japanese readership simply by the addition of reading marks and phonetic glosses.⁴¹ Through such means, the new type of Chinese dictionary, with consistent arrangement of characters according to both radical and order and number of brush strokes, was adapted for the Japanese market. The new radical-based dictionaries of the late Míng and early Qīng (see Chapter 6) reached Japan. The *Zihui* was especially popular: it was reprinted with the addition of pronunciation glosses, and Japanese writers treated it as an authority.⁴² Revised versions such as *Wa jii* 和字彙 (‘The Japanese collection of characters’) and *Shō jii* 小字彙 (‘The collection of characters, abridged’) appeared.

Dictionaries of Japanese origin were also reworked and printed. *Shūbun inryaku*, the rhyme book for the composition of Chinese verse dating from the early fourteenth century, was rearranged according to the Japanese reading of the Chinese-character lemmata and printed early in the Tokugawa period as *Irohain* 伊呂波韻 (‘Rhymes in syllable order’). Such

³⁹ Clements, *Cultural History of Translation*. ⁴⁰ Sugimoto, *Jisho, jiten*, 42.

⁴¹ Kornicki, *Book in Japan*, 153–5.

⁴² The summary of Tokugawa lexicography that constitutes the remainder of my section on Japan is a paraphrase of Sugimoto, *Jisho, jiten*, ch. 1, sections 3.2–3.4 and section 4, with my other citations referring only to specific points.

a book allowed a reader to look up a character she wanted to use in a Chinese poem without having to know or guess its Middle Chinese pronunciation in advance. *Kagakushū*, similarly, was published in an enlarged edition in 1669. *Setsūyōshū* was subject to greater reworking, testifying to its popularity. Some new editions expanded the book's contents to make it into an encyclopedia. Other editions were small, pocket-sized or thin and oblong books with pages shorter than they were wide, allowing for more but shorter columns that could easily be browsed. *Hayabiki setsūyōshū* 早引節用集 ('Collection of words for everyday use, easily retrievable') is an example. In such books, the arrangement was modified to facilitate word-retrieval. In addition to arranging the Japanese words according to their first syllabary in *iroha* order, the new *Setsūyōshū* further arranged them according to the number of syllables in ascending order. Other *Setsūyōshū* targeted specific groups of readers, such as children or females. These books became 'guides to everyday linguistic life'.⁴³

As in earlier periods, literary trends affected lexicography. Word repositories for the writing and appreciation of classical Japanese verse (*waka*) were published in the Tokugawa period. Meanwhile, the popularity of a new genre, *haikai* 俳諧, using more vernacular language, led to the compilation of vocabularies of that linguistic register. A broadened interest in the lexicon is also evidenced by the publication of dictionaries of idioms, aphorisms, and dialect expressions. *Sewaji setsūyōshū* 世話字節用集 ('Collection of colloquialisms for everyday use') from 1692 was, despite its title, not a dictionary arranged according to *iroha*, but one of colloquial words and expressions arranged by subject matter. It was similar to *Jigen benmō shō* 通言便蒙抄 ('Vernacular expressions annotated for young students', 1682), later pirated as the illustrated *Zatsuji kinmō zui* 雑字訓蒙図彙 ('Illustrated collection of miscellaneous words for instructing young students'). The expressions collected and explained in this topically arranged book were vernacularisms from contemporary literature, including words written with Chinese characters used only for their phonetic value, a latter-day variant of the *man'yōgana*.

The interest in the foundational Japanese literary and historical texts among Tokugawa scholarship retrospectively called 'national studies' contributed to the compilation of new dictionaries. Collections of vocabulary drawn from *waka* and other old genres, these books anticipate modern Japanese dictionaries in their *iroha*-based arrangement. *Ōmushō* 鸚鵡抄

⁴³ Sugimoto, *Jisho, jiten*, 44.

(‘The annotated parrot’), completed in 1685, is one example. It listed more than 22,000 words, cited examples and various opinions from ancient and modern sources, and arranged the lemmata not only by their first syllable, but also by their second. The arrangement began to approach that of contemporary dictionaries in other ways too. Etymological dictionaries such as *Nihon shakumyō* 日本釈名 (‘The Japanese “Explaining names”’, 1700), named in reference to *Shiming*, an ancient Chinese work arranged by topic (for which see Chapter 3), superposed an arrangement according to the fifty-syllable table onto the topical organization, anticipating the general shift to *gojūonzu* order in the nineteenth century. *Tōga* 東雅 (‘The eastern elegance’, 1717) referenced another ancient Chinese work (*Ėryā*) in its title, and innovated by considering Sanskrit and Korean in its discussion of the origins of Japanese words. Tanikawa Kotosuga’s *Wakun no shiori* 倭訓栞 (‘Guide of Japanese glosses’), published in parts after the author’s death in 1776 and continuing well into the nineteenth century, was another work bridging the Tokugawa period with the later linguistic order of the Japanese nation-state. Arrangement according to the fifty-syllable table; the inclusion of ancient, refined, vernacular, and dialect vocabulary (the latter two collected in the last part published in 1877); and, finally, the citation of sources make this dictionary similar to contemporary dictionaries. Collections of dialect vocabulary, often in the form of regional *materia medica*, also appeared in the Tokugawa period.

Foreign-language lexicography in Tokugawa Japan primarily meant dictionaries of contemporary vernacular Chinese and Dutch, even though scholars of the period also took an interest in other languages such as Manchu, Russian, and English. Yet even in the case of Dutch, and more so with the other European languages and Manchu, the dictionaries resulting from this foreign-language research appeared only in the nineteenth century, and thus lie beyond the scope of this chapter.

In addition to commercial interaction with China, the popularity of vernacular Chinese fiction in Japan and the intellectual trend retrospectively known as ‘ancient learning’ (*kogaku* 古学) provided incentives to compile dictionaries of the language of the great continental neighbour. Tokugawa scholars distinguished the Confucian learning of antiquity from the so-called Neo-Confucian revival in the first centuries of the second millennium. The partially vernacular language of the Neo-Confucian texts was treated in glossaries from the late seventeenth century onwards. In the eighteenth century, dictionaries appeared that explicitly focused on the language of Chinese novels, such as *Shōsetsu jii* 小説字彙 (‘Characters collected from

the novels', 1791), whose title made reference to the influential *Zihui*, and which provided Chinese words with Japanese synonyms. Some dictionaries, such as the one just mentioned, focused exclusively on vernacular Chinese vocabulary and its translation into current Japanese, whereas others included vernacular words as part of a larger investigation of the Chinese lexicon as a resource available to users of Japanese. *Jōgosō* 常語藪 ('Collection of ordinary expressions', 1795), arranged topically according to a system common in China, is an example of such a work. Just as the introduction of Chinese vernacular fiction stimulated the literary description of the everyday world of Tokugawa Japan, so too did the lexicographical description of vernacular Chinese stimulate an interest in the vernacular Japanese language, which was used to translate the Chinese expressions.⁴⁴

The work of Dutch and later other Western languages in the Tokugawa period was not, as was the case in many other places in Asia and in Japan in the sixteenth century, an example of missionary linguistics – for Christianity was strictly forbidden – but a result of Japanese intellectuals' engagement with European learning.⁴⁵ The first published dictionary of Dutch, *Haruma wage* 波留麻和解 ('Halma with Japanese explanations'), appeared in 1796, at the end of the period under consideration in this chapter. In a testament to the spread of Dutch learning from the interpreter milieu at Nagasaki to the intellectual mainstream, the book was published in Edo, location of the shogunal government. Arranged in alphabetical order, the book presented Dutch lemmata translated into Japanese. It was based on a Dutch–French dictionary by François Halma, introduced to Japan in 1754, hence the title of the Japanese book.⁴⁶ In the early nineteenth century, other dictionaries followed, as Japanese contact with Western countries increased.

Korea

Korea had close contacts with China, and the Korean elite were proficient in literary Chinese and its Confucian tradition. The Korean language, however, is very different from Chinese, its grammar being similar to that of Japanese or Manchu. The genres that have been identified as antecedents of modern dictionaries in Korea are rhyme books and dictionaries listing Chinese characters according to their constituent parts.⁴⁷ However, the many

⁴⁴ On fiction, see Pastreich, *Observable Mundane*.

⁴⁵ For background, see Goodman, *Japan and the Dutch*.

⁴⁶ Mac Lean, 'Introduction of books', 14; Proust, 'De quelques dictionnaires', 21–2.

⁴⁷ Park, *Hanguk jajeon*, 191.

vocabularies compiled in the Joseon period (1392 onwards) were also part of the world of Korean lexicography.

Considering only the major genres that can without qualification be called reference works, that is, phonologically arranged rhyme books (Chinese *yùnshū*; Korean *unseo*) and graphologically arranged character books (Chinese *zishū*; Korean *jaseo*), then the most striking characteristic of the development of Korean lexicography compared to Chinese lexicography is that, in Korea, character books developed out of indexes to rhyme books. In China, by contrast, the major character books of the Qīng period, such as the paradigmatic *Kāngxī zìdiǎn*, developed by incorporating elements of rhyme books, such as phonological indexes.⁴⁸

Compared to the situation in Japan, the early history of Korean lexicography is poorly documented. As far as I know, no Korean compilations, or even Korean copies of Chinese dictionaries, are extant from the first millennium. All we have are a few mentions. King Muyeol 武烈 (602–61) of Silla is said to have studied the *Ēryǎ* with the scholar Gangsu.⁴⁹ It has been proposed (but not convincingly) that the famous statement in *Samguk sagi* 三國史記 ('Historical record of the three kingdoms', 1145) that Seolchong 'read the Nine Classics using local language [*bang-eon*] and taught it to his students', commonly interpreted to mean that Seolchong used the existing techniques for adapting literary Chinese text to Korean by inserting grammatical particles, in fact refers to Seolchong using the Chinese dictionary of synonyms *Fāngyán* (Korean *Bang-eon*; for it, see Chapter 3) to interpret the Confucian classics.⁵⁰

Beginning in the Goryeo period (918 onwards), there is evidence of Korean reprints of Chinese lexicographical books. The oldest extant copy of the Chinese *Yīqièjīng yīnyì* 一切經音義 (Korean *Ilche gyeong eum-ui*; see Chapter 6), a Buddhist glossary from the Tang period, is a reprint from Goryeo.⁵¹ The great upswing in the popularity of rhyme books in this period is due to the civil service examination system, in use from 958, which tested candidates in the composition of Chinese verse and prose.⁵² In the 1030s, *Lǐbù yùnlüè* (see Chapter 6) was imported into Korea and reprinted soon thereafter. It remained the standard for the Korean examinations until the Joseon period.⁵³ Yet, after its publication in 1229, the so-called *Píngshuǐ* version of *Lǐbù*

⁴⁸ Park, *Hanguk jajeon*, 203–4. ⁴⁹ Park, *Hanguk jajeon*, 92.

⁵⁰ Lee, *Geschichte der koreanischen Sprache*, 55; No, 'Hanguk godae', 21–2.

⁵¹ Sim, 'Hanguk ui unseo', 143–4. ⁵² Duncan, 'Examinations', esp. 72–4.

⁵³ Sim, 'Hanguk ui unseo', 144.

yǔnlüè was adopted in Korea, with a reprint dating from 1297.⁵⁴ The importation of Chinese rhyme books into Korea continued thereafter.

The invention of the Korean alphabet, now called *hangeul*, in 1443 unsurprisingly had an enormous influence on Korean lexicography. The new script was immediately used in several royally sponsored rhyme books. Exactly who at court did what and why might be debated, but it is clear that without the recent experience of the Mongols' 'Phags-pa script and advances in Chinese phonological studies, there would have been no *hangeul*.⁵⁵ The Joseon court was well aware of 'Phags-pa and other scripts used in the region and, as mentioned above, 'Phags-pa dictionaries influenced the Chinese rhyme books *Gǔjīn yùnhuì* from 1292 and *Hóngwǔ zhèngyùn* (see Chapter 6), both of which reached Korea. So, on the one hand, the inventors of *hangeul* might have been directly influenced by 'Phags-pa when designing the new letters and, on the other, the Chinese and 'Phags-pa-influenced rhyme books played an important part in the early history of the new alphabet.⁵⁶

The royal court was interested in promulgating knowledge of the reading pronunciation current in Míng China as well as defining proper Sino-Korean reading pronunciation. In 1444, work began on several rhyme dictionary projects: *Saseong tonggo* 四聲通考 ('Comprehensive examination of the four tones'; completed before 1450), *Hongmu jeong-un yeokhun* 洪武正韻譯訓 ('Correct rhymes of the Hóngwǔ emperor, translated and glossed'; completed 1455), and a translation of *Gǔjīn yùnhuì jǔyào*. The last project was never completed, but in 1447 another rhyme book was finished. Titled *Dongguk jeong-un* 東國正韻 ('Correct rhymes of the eastern state'), it appears to have been compiled in lieu of translating *Gǔjīn yùnhuì jǔyào*.

Dongguk jeong-un was made to standardize Sino-Korean character readings, notably in the context of providing Buddhist scriptures with Korean sound glosses.⁵⁷ Whereas *Dongguk jeong-un* was intended to standardize usage within Korea, *Hongmu jeong-un yeokhun* was intended to promulgate knowledge of normative reading pronunciation. It provided Korean alphabetic transcriptions of the Chinese pronunciation of the rhyme book's lemmata. The Korean translation of the Hóngwǔ rhymes kept the original's arrangement, which, following Middle Chinese tradition, grouped all lemmata first by tone and then by rhyme. It has been shown that this book, too, made important use of *Gǔjīn yùnhuì jǔyào*.⁵⁸

⁵⁴ Park, *Hanguk jajeon*, 155–7.

⁵⁵ Lee, 'Inventor of the Korean alphabet'; cf. Lee, 'Hunmin jeong-eum'.

⁵⁶ For the possibility of direct influence, see Ledyard, 'International linguistic background'.

⁵⁷ Sim, 'Hanguk ui unseo', 145.

⁵⁸ Dormels, 'Hunmin-chōngum and the transcription of Chinese pronunciation'.

Saseong tonggo was an abbreviated version of the Hóngwǔ rhymes made for easy consultation. It listed Chinese characters irrespective of tone under their transcription into the Korean alphabet. The three pitch tones of Middle Chinese were not retained in Sino-Korean pronunciation, so this arrangement made it easier for Korean users to look up characters.⁵⁹ Such an arrangement had precedents in Mongol and post-Mongol China, being used in *Zhōngyuán yīnyùn* (see Chapter 6).⁶⁰ The great lexicographic endeavours of the fifteenth century were continued in the early sixteenth, with the revision of *Saseong tonggo* as *Saseong tonghae* 四聲通解 ('Comprehensive explanation of the four tones', 1517) and the emendation of the translated Hóngwǔ rhymes as *Sokcheom Hongmu jeong-un* 續添洪武正韻 ('Subsequent additions to "Correct rhymes of the Hóngwǔ emperor"').⁶¹

Saseong tonghae retained the arrangement first by rhyme and then by tone, providing the Chinese character headwords with simple glosses in literary Chinese and in a few hundred cases also in Korean. The Korean glosses are related to the contents of *Hunmong jahoe* 訓蒙字彙 ('Collection of characters for the instruction of children').⁶² This work was published by the same author ten years later and is famous for its instruction of the names that are still used for the Korean letters. It was a topically arranged vocabulary that transcribed and glossed Chinese characters. Yet the vocabulary was often abstract and poorly suited for elementary teaching. The character 燠, 'warm', for example, was glossed in Korean as 'warm *uk*' (*deoul uk*), with *uk* being the Sino-Korean reading of the character meaning 'warm'.⁶³

The invention of the Korean alphabet, then, immediately led to great developments in phonological lexicography. Graphologically arranged dictionaries, character books, were not part of the court's agenda. In 1573, however, *Unhoe okpyeon* 韻會玉篇 ('Jade chapters to the Confluence of rhymes') was produced as an index to *Gǔjīn yùnhui jǔyào*.⁶⁴ The development of character books through rearranging lemmata from rhyme books continued in later centuries.

The compilation and reprinting of rhyme books certainly also continued. One of the more widely used books was *Samun tonggo* 三韻通考 ('Comprehensive examination in three rhymes'), which might date from the fifteenth century but is of unclear authorship and date. This dictionary represented the rhyme system used for the poetry component of the civil

⁵⁹ Ledyard, 'International linguistic background', 48–9; Park, *Hanguk jajeon*, 172–7.

⁶⁰ Ning, *Hànyǔ yùnnshū*, 7. ⁶¹ Park, *Hanguk jajeon*, 172–7. ⁶² Park, *Hanguk jajeon*, 177–80.

⁶³ See reproduced image in Ogura, *Zōtei hochū Chōsen gogakushi*, 195, and description 195f.

⁶⁴ Park, *Hanguk jajeon*, 160.

examinations, but rearranged it so that rhyming characters in the three pitch tones of Middle Chinese were listed on the same page, with the entering-tone rhymes relegated to a separate section at the back. Again, as was the case in the *Saseong tonggo*, the reason for this arrangement was that the entering tone was retained and identifiable in Sino-Korean pronunciation, whereas the others were not. The Japanese rhyme book *Shūbun inryaku* (also, as we have seen, known as *Sanjūin* ‘The three stacked rhymes’ because of its treatment of the three pitch tones together) had a similar arrangement for the very same reason. Indeed, one theory that flourished already in the Joseon period held that *Samun tonggo* in fact originated in Japan, whence it had been imported to Korea.⁶⁵ The book was complemented in the eighteenth century with the addition of summaries of recent Qīng research on ancient Chinese rhymes.⁶⁶

The eighteenth century, the end point of this survey, saw further developments in rhyme books and also character books. In 1691, the encyclopedic, bipartite dictionary *Okhwuiun* 玉彙韻 (‘Jade collection rhymes’) was published. The first part lists literary Chinese idioms (*sugeo* 熟語) drawn from historical sources, arranged in rhyme order. The second part lists single Chinese characters according to rhyme, distributed across the four tones all appearing on the same page, providing them with pronunciation glosses and a definition.⁶⁷

The great rhyme books of the eighteenth century were *Hwa-Dong jeong-eum tongseok un-go* 華東正音通釋韻考 (‘Examination of rhymes in the correct pronunciation of China and the East [viz. Korea], with comprehensive explanations’) from 1747, *Samun seonghwi* 三韻聲彙 (‘Collected sounds of the three rhymes’) from 1751, and *Eojeong Gyujang jeonun* 御定奎章全韻 (‘Royally commissioned complete rhymes of the palace library’) from 1796. The first of these books was based on the expansion of *Samun tonggo* from the early eighteenth century. It innovated by supplying two kinds of sound glosses, one Chinese reading, drawing on *Saseong tonghae*, and one Korean reading. The book was reprinted in 1787 with a royal preface, and was thereafter bestowed on successful civil examination candidates.⁶⁸ *Samun seonghwi* was made upon consultation of *Hwa-Dong jeong-eum tongseok un-go* and other earlier books. It is noteworthy for having two parts: in addition to the rhyme book proper, there is a graphological index (*okpyeon* 玉篇),

⁶⁵ Ogura, *Zōtei hochū Chōsen gogakushi*, 685–6; Sim, ‘Hanguk ui unseo’, 148; Park, *Hanguk jajeon*, 163–8.

⁶⁶ Fang, *Asami Library*, 44–5 (item 12.7). ⁶⁷ Park, *Hanguk jajeon*, 180.

⁶⁸ Fang, *Asami Library*, 42–3 (item 12.1); Park, *Hanguk jajeon*, 181–3.

allowing the reader to look up characters according to their form and learn under which rhyme they are to be found.⁶⁹ *Eojeong Gyujang jeonun*, finally, was written on royal command and served to compose Chinese regularized verse. Within the rhymes, characters are arranged further according to their Korean alphabet transcriptions. The Korean transcriptions, however, are normative and do not reflect current vernacular usage. The book underwent some revision in the nineteenth century and was extremely popular throughout the Joseon period. *Eojeong Gyujang jeonun* was also transformed into a graphological dictionary as *Jeon-un okpyeon* 全韻玉篇 ('Jade chapters for the *Complete rhymes*') at some point in the nineteenth century. This was more than an index to the rhyme book, as definitions were provided for its lemmata, not only indications on the rhyme under which they could be found in the original rhyme book.⁷⁰

In addition to rhyme books and the graphological indexes and dictionaries based on them, Joseon scholars compiled a variety of other lexicographical works. Glossaries to linguistically difficult corpora of particular interest to Korean intellectuals, such as the Neo-Confucian 'recorded sayings' (Chinese *yǔlù* 語錄), which are partially in vernacular Chinese, are one example.⁷¹ Many topically arranged collections are on the border between lexicography and encyclopedic compilation. *Jasan eobo* 茲山魚譜 ('List of fish in Jasan'), for example, is a list of terms for marine life used on the island Heuksan-do 黑山島, where its author had been exiled.⁷²

I should mention, finally, the important tradition of foreign-language lexicography that flourished in the Joseon period. Contemporary Chinese character readings were included in some of the rhyme books discussed above, and vernacular Chinese terms occurred in glossaries. The Joseon government, furthermore, maintained a staff of interpreters for facilitating contacts with neighbouring countries. The interpreters compiled topically arranged dictionaries for vernacular Chinese, Japanese, Mongolian, and Manchu. *Yeogeo yuhae* 譯語類解 ('Translated terms classified and explained') from 1690 was a dictionary for vernacular Chinese. It was followed by similar works for the other languages in the eighteenth century.⁷³ Dictionaries imported from the Qīng empire were used in the compilation of such books.

⁶⁹ Park, *Hanguk jajeon*, 183–6.

⁷⁰ Sim, 'Hanguk ui unseo', 155–64; Park, *Hanguk jajeon*, 186–90, 206–8.

⁷¹ Fang, *Asami Library*, 43–4 (item 12.4). ⁷² Lee, 'Silhak sidae', 126.

⁷³ Song, *Study of Foreign Languages*, 72–3.

The Turkic Languages and Persian to c. 1700

MAREK STACHOWSKI

Names for Dictionaries in the Turkic Languages and in Persian

In Europe, the word *sözlük* (from *söz* ‘word’) is the best-known Turkish designation of a dictionary because it regularly appears in the titles of tourist dictionaries of Turkish. However, this word actually first appeared in the title of a Russian–Turkmen (!) dictionary of 1929, and was adopted by the Society for the Investigation of the Turkish Language (Türk Dili Tetkik Cemiyeti) only in 1932.¹ Earlier, that is both in the Ottoman and in the early republican periods, two other words were in use: *lüğat* or *lügat* (from Arabic) and, more rarely, *divān* (from Persian). While the latter word is used for a collection of poems in Persian, the former also appears as *lughat* ‘dictionary’ in that language. Today, Persian *farhang* is the most common word with this meaning. In addition, in both the Ottoman and the Persian traditions, the word *qāmūs* (Arabic from Greek *ōkeanós* ‘ocean’) was employed for great works combining a dictionary with a sort of encyclopedia; this practice originates with the Arabic dictionary of al-Fīrūzābādī (see Chapter 8). In the other Turkic languages usually either a local adaptation or translation of Turkish *sözlük* or, sometimes, of *lughat* or else a loan from Russian *slovar* ‘dictionary’ is in use.

The Beginnings of Persian Lexicography

It is true that the oldest Persian dictionaries have not survived, and we are sometimes not even able to date them exactly; thus, the oldest Persian wordlist whose name is known to us is *Risālah* by Sughdī, but attempts at

¹ Ölmez, ‘Tarihi Türk’, 110.

dating it vary between the seventh century and the eleventh.² Nevertheless, although in such cases we are informed about the first dictionaries only by some later Persian authors who mentioned them and sometimes also adduced some words or passages from them in their own works, there seems to be no reason to doubt their connections with poetry and literary works in Persian. The oldest Persian lexicographers aimed at providing help with accurate versifying and correct understanding of poems, explained dialectal and archaic words, elucidated rare idioms, and so on.

Simultaneously, multi-language vocabularies (for instance, Avestan–Persian and Aramaic–Persian) were compiled.³ In those days they were probably used as teaching aids because their word material was divided into semantic groups, such as ‘numerals’, ‘parts of the body’, ‘God and religion’, and others, so that each was actually a collection of specialized wordlists. Nowadays they are mostly studied for rare words unattested in other sources.

The earliest extant wordlist is *Lughat-i furs* (c. 2,000 words) compiled about 1066 by a poet and copyist from Tus in eastern Persia, and thus called Asadī Tūsī, that is, Asadī from Tūs. He introduced a few methodological devices that would be followed by his successors. First, he organized his word material according to the alphabetical order of the last letter of a given word, thus making his list a sort of rhyming dictionary (this is usually considered to be a result of Arabic influence). Secondly, he adduced phrases or passages from different poems written by more than a hundred poets in order to illustrate various meanings of the words. Thirdly, from time to time, he indicated dialectal areas in which a specific word was used.⁴ Tūsī’s main aim was to make vocabulary used in the east of Persia and therefore attested in court poetry comprehensible to readers in the west: ‘Thus the earliest Persian monolingual dictionaries were compiled by poets of the east for poets and readers of the west.’⁵

The Arrangement of Persian Dictionaries

As mentioned above, Asadī Tūsī introduced chapters with headwords arranged according to their final letters. That system must have occurred very naturally to him as a poet because verses in poetical collections called *dīwān* were also grouped in the same way.⁶ Thus, it was very easy to find all

² MacKenzie, ‘Persian lexicography’, 525.

³ See Malandra, ‘Frahang ī oīm’, and MacKenzie, ‘Frahang ī pahlawīg’.

⁴ Kapranov, ‘Lexicography’, 354–5. ⁵ Maciuszak, ‘Persian lexicography’, 183.

⁶ Baevskij, *Rannijaja persidskaja*, 109.

words with identical closing letters. However, navigating within a chapter was made difficult by the fact that words were not arranged in any way within the chapter. Nevertheless, Tūsī's system was only modified more than 250 years later, by Moḥammad Nakhjawānī, who used a more elaborate two-level system in his dictionary *Šiḥāḥ al-furs* (1327). He first collected all words ending in a specific letter into a chapter called *bāb* ('door, gate'). Then he divided them according to their initial letters into subdivisions called *faṣl* ('section, part'). Later, the terms *ḡunah* ('sort') and *bakhsh* ('part') were also sometimes used.

In the following centuries the opposite system would occasionally, albeit rather rarely, be applied: words in chapters were organized by the first letter, those in subdivisions by the last one. In addition, a fifteenth-century lexicographer, Qāzī Khan Dihlawī, often referred to as Dhārwal because he lived in the town of Dhār in India, compiled a two-part dictionary, *Adāt al-fuṣalā*, of which the first part contained words arranged according to their first letter while the second part had compounds and phrases alphabetized by first and last letter.⁷ This method would in the course of time evolve into a fully alphabetical system, which would become the only usual one in the second half of the seventeenth century. The seventeenth century would also witness yet another attempt. In his dictionary dedicated to Jahāngīr, the fourth Mughal emperor, and therefore called *Farhang-i Jahāngīrī*, Ḥusayn Injū would divide words into chapters by second letter and further into sections by first letter. However, that system never became really popular.

Finally, a typical feature of Persian wordlists was that they were written out in continuous text, and the specific headwords were singled out by being underlined or written in red.⁸

Maḥmūd al-Kāšgarī's Turkic Comparative Dictionary (1077?)

The Persian Asadī Tūsī and the Uyghur Maḥmūd al-Kāšgarī, who are the oldest lexicographers known in the Persian and the Turkic traditions respectively, were near-contemporaries. Their dictionaries were completed within about a decade of each other: Tūsī's around 1066 and Kāšgarī's around 1077. In other respects, however, they differed a lot from each other, as we shall see in the next section.

⁷ Akhtar, 'Dhārwal'. ⁸ Baevskij, *Rannjaja persidskaja*, 112.

Not only is Kāšgarī's year of birth (c. 1008) uncertain; his place of birth is also disputed. The most popular way of thinking is of course to connect him with Kāšgar because of his *nisba* Kāšgarī. However, the geographical name of Kāšgar refers both to a city and to a prefecture whose centre this city is. Mehmet Ölmez says on the basis of his investigations in China that, according to the local tradition, Maḥmūd al-Kāšgarī was born in the small Uyghur town of Oypal in the Kashgar Prefecture in the north-west of China.⁹ But Robert Dankoff says in his edition of Kāšgarī's dictionary that Kāšgarī was 'born in Barsyān near Lake Issik-kul' (now in Kyrgyzstan).¹⁰

The title of the dictionary is *Dīwān Luġāt at-Turk* (or, in the modern Turkish spelling often used nowadays, *Divân-u Lügâtî't Türk*). Thus, the title is Arabic, and so are all the commentaries, the explanations, the introduction, and so on. A significant detail is the vowel length in the word *luġāt* because it signals the plural, that is 'languages, speeches', the singular form being *luġat*. In other words, what Kāšgarī offers us is not just a dictionary of Turkish or Turkic. It is, instead, an Arabic-language comparative dictionary, with encyclopedic elements, of various Turkic dialects. Even if its comparative (and sometimes, albeit very rarely, etymological) character cannot be observed in absolutely every entry, it is undeniable. Two examples may be given here:

sēn 'You . . . ' Kāncāk dialect. The Turks say: **sān**. Thus the speech of Kāncāk is corrupt . . . since they always prefer *kasra* (I). The Čigil, Yağma, and Tuxsi, as far as Upper Šin, prefer *ḍamma* (U). The Oğuz, Qifčaq, and Suvārīn, as far as Rūm, prefer *fatha* (A). . . . This is the kind of difference among these groups.¹¹

yumyāq Anything 'rounded or circular . . . ' Thus 'coriander seed . . . ' is called **yumyāq tana** in Uč dialect. As for **tana** I reckon it to be the Persian word for 'seed', *dāne*, which has been Turkicized to **tana**.¹²

Kāšgarī is said to have peregrinated in Turkic-speaking areas of Central Asia over a period of sixteen years, collecting very different phonetic, semantic, and stylistic variants of Turkic words, as well as samples of folk songs, riddles, proverbs, and other pieces of the oral literature of Turkic tribes.¹³ Because Kāšgarī always marks, as precisely as was possible to him, the geographical or tribal area of specific words and phrases, his dictionary is, indeed, more a comparative lexicon than a simple Turkic–Arabic wordlist. That is why

⁹ M. Ölmez, in Yeşildal, 'Bir Türklük atlası', 9b.

¹⁰ Dankoff, in Maḥmūd al-Kāšgarī, *Compendium*, I.4, following Hazai, 'Al-Kāshgharī', 699b.

¹¹ Maḥmūd al-Kāšgarī, *Compendium*, II.218–19. ¹² Maḥmūd al-Kāšgarī, *Compendium*, II.170.

¹³ Kaçalin, in Yeşildal, 'Bir Türklük atlası', 10b.

the title of his work was translated by Dankoff as 'Compendium of the Turkic dialects'.

Usually, 1071, 1072, or 1074 is given as the year in which the *Dīwān* was completed. However, pieces of information concerning the dates of writing the dictionary can be found at four points, and they do not seem to be fully compatible. In the modern and nowadays most popular edition of the *Dīwān*, all these dates are compared and discussed, the final result being the opinion that the work on the dictionary most probably began on 25 January 1072 and the fourth and final redaction was completed on 9 January 1077.¹⁴ If these dates are accurate, Kāšgarī not only carried out field research collecting Turkic word material and folklore samples for sixteen years, but then worked on their redaction for five years. Such an effort needs a very good motivation. Probably the most satisfying opinion is that Kāšgarī wanted 'to show that the Turkic dialects keep pace with Arabic like two horses in a race'.¹⁵

The original copy of the *Dīwān* has not survived to this day. What we have at hand now is a handwritten duplicate of a 1266 copy of the original manuscript. It was found, more or less by accident, in a used bookstore in Istanbul in 1915, by Ali Emîrî, a Turkish bibliophile and extremely ardent reader who, a year later, would establish a National Library in Istanbul (Turkish Millet Kütüphanesi, not to be confused with the National Library in Ankara, called Millî Kütüphane). He also gave his collection of 16,000 rare books and manuscripts as a gift to that library and was its director in the last eight years of his life. The copy of the *Dīwān* was in this collection. Immediately after the purchase of the manuscript, he decided to offer it for editing to Kilisli Rifat, a Turkish philologist in Istanbul. Rifat's edition appeared in three volumes in 1917–19.

The character of the lexicon is best understood by considering a geographical map which was included in the *Dīwān*. It was made, like so many maps in those days, with east at the top while the Turkic tribes were located in the central part of the map. In other words, Kāšgarī took pains to show his readers the geographical distribution of the tribes he was discussing. A special curiosity is that, at the very top of the map, that is, in the very east of the depicted world, Japan is marked, with the name *Ġābarqā*, which is a phonetically distorted variant of the Chinese name for Japan.¹⁶

¹⁴ Dankoff, in Maḥmūd al-Kāšgarī, *Compendium*, I.6–7, 23.

¹⁵ Dankoff, in Maḥmūd al-Kāšgarī, *Compendium*, I.41, quoting *ibid.*, I.71.

¹⁶ The map is redrawn and its text translated as a fold-out in Maḥmūd al-Kāšgarī, *Compendium* I, between 82 and 83; discussion by M. Ölmez, in Yeşildal, 'Bir Türklük atlası', 12a.

About 8,000 words are presented by Kāšgarī in his lexicon. The arrangement model he applied was a modified version of a system elaborated by the tenth-century Arabic lexicographer al-Fārābī in his *Dīwān al-Adab* (see Chapter 8) and probably never used by anybody else. Kāšgarī first divides all the words into eight chapters according to their phonological structure. There exists, for instance, a chapter for ‘words having a geminate’, one for ‘words with consonant clusters’, one for ‘words with nasals *ŋ* or *nč*’, and so on. (By the way, the treatment of *nč* just as a nasal rather than as a consonant cluster is fairly intriguing, and seems possibly to point to Kāšgarī’s acquaintance with the Old Turkic runic script, in which one letter was actually used to mark the consonant cluster *nč*.) Each of the chapters is further divided into two sections: ‘Nouns’ and ‘Verbs’. Within each section words are arranged according to the number of their consonants. The next rule is organizing the words in accordance with Arabic structural patterns of the type *ʿfaʿl*, *fuʿl*, *fiʿl*.¹⁷ Within each structural section words are further arranged by last letter.¹⁷

The most important strengths of Kāšgarī’s *Dīwān* are as follows: (a) it is a new type of dictionary in the Islamic world, a comparative one; (b) it also includes geographical data, namely the map on the one hand and dialectal remarks on the other; (c) the base of the dictionary is samples of living spoken Turkic dialects; (d) it is a substantial wordlist – a total of 8,000 words is a great achievement.

A Comparison of the Dictionary by ʿTūsī with That by Kāšgarī

As was said at the beginning of the previous section, ʿTūsī and Kāšgarī were contemporaries. Besides, they were both acquainted with principles elaborated by the Arabic school of lexicography. That is why one might suppose that also their dictionaries would be very similar to each other, apart from the languages they concern. However, that is not at all the case. The main human and professional difference is that ʿTūsī was in the first place a poet rather than a scholar, while Kāšgarī was a proper philologist and field dialectologist. The following criteria, some of them resulting from that fact, can be used to examine the main features distinguishing the two dictionaries:

- (a) **Priority.** Even though we usually describe the works by these two authors as the oldest ones in the Persian and the Turkic lexicographical

¹⁷ See Dankoff, in Maḥmūd al-Kāšgarī, *Compendium* 1.33; Kāšgarī’s own list of chapters is *ibid.*, 1.71.

traditions respectively, only Kāšgarī actually deserves to be called the father of a lexicographical tradition, while Ṭūsī just had good luck with his dictionary, which did not vanish like those written by his many predecessors.

- (b) **Aim and target.** Ṭūsī wrote his dictionary as a poet, and for poets as well as poetry readers, so that his motivation was social rather than national. The situation with Kāšgarī is very different, because his motivation was clearly patriotic, as he first of all wrote for Arabs whom he wished to convince about the high quality of the Turks, their dialects, and their culture.
- (c) **Documentation.** A feature common to both dictionaries is the documentation of the usage of words and idioms by adducing passages from poetry, written (Ṭūsī) or oral (Kāšgarī).
- (d) **Normativity.** Ṭūsī's dictionary seems to display a stronger inclination towards linguistic prescription than is the case with Kāšgarī who, however, generally advocated Turkic linguistic purism. The situation can be readily explained because Ṭūsī's aim was literary, while Kāšgarī tried to demonstrate the richness of the Turkic lexis he had archived, so that he relatively seldom felt the need to express his personal opinion on what was better or more correct in the Turkic dialects.
- (e) **Type of dictionary.** Ṭūsī's *Lughat* was, as far as we can formulate an opinion on the lost works of his predecessors, a traditional type of lexicographical aid for the correct understanding and writing of poetry, and, at the same time, a wordlist of one language only, even though it bridged the gap between the eastern and western dialects. Unlike Ṭūsī, Maḥmūd Kāšgarī introduced a new type of wordlist: a comparative dictionary with encyclopedic elements (the term 'comparative' is of course not to be understood exactly in our modern sense; it was more a dictionary with a strong comparative tendency). His work was not intended to be a poetical aid and, moreover, it concerned a whole family of languages rather than just one. In addition, the *Lughat* defined Persian words in Persian while the *Dīwān* explained Turkic words in Arabic.
- (f) **Capacity.** About 2,000 words in Ṭūsī's work versus about 8,000 words in Kāšgarī's.
- (g) **Material base.** Ṭūsī depended on poetical works. Thus, his dictionary is what we could, today, call a literary product of an armchair researcher. Kāšgarī undertook protracted fieldwork, and his dictionary in the first place concerns the living spoken dialects of Turkic tribes.

- (h) **Arrangement.** While Ṭūsī works with one arrangement rule only (last letter), Kāšgarī applies as many as five rules: phonological features; morphological classification; number of consonants; Arabic word-formation patterns; last letter.
- (i) **The map.** Kāšgarī made a special map of the tribes whose lexis is cited in his work. Ṭūsī probably did not deem such a map necessary because his readers were educated Persians, and one could, therefore, expect them to know the geography of Persian-speaking areas. In contrast, Arab readers of Kāšgarī's *Dīwān* would be unlikely to know the localization of specific Turkic tribes. The fact that Kāšgarī also marks Japan is an additional advantage of the map and points to his wide knowledge of the geography of Asia.
- (j) **Tradition and continuation.** While Ṭūsī arose from a tradition and became a methodological pattern for his followers, Kāšgarī, as a real initiator, had no predecessors and, unfortunately, no followers. Ṭūsī's methodology remained in force and was developed in Persia. Kāšgarī's methodology and his comparative dictionary were forgotten for centuries – the continuation of his dialectal fieldwork and comparative methodology as well as that of his encyclopedic treatment of the vocabulary would only be found (of course in a better and more modern sense) in Wilhelm Radloff's comparative dictionary at the end of the nineteenth century as well as in Edward Piekarski's Yakut dictionary in the first half of the twentieth century, and even the manuscript of the *Dīwān* would, as we now know, be rediscovered only in the first quarter of the twentieth century.

Dictionaries of Persian with Arabic and Turkish

Persian dictionaries other than the literary dictionaries arranged by final letter which were mentioned above may be divided into at least two typological groups: bi- and multilanguage ones on the one hand and specialized ones on the other. Iranian metalexicography more often than not regards Persian wordlists compiled in India as a separate group, although they, too, may represent different typological variations. An outlying group are European wordlists, generally multilingual, that include the Persian lexis among others. In what follows I am going to characterize some works constituting each of these groups.

There can be no doubt that Arabs exerted huge influence on Persian lexicographic culture. Several descriptive Arabic dictionaries were even

compiled by ethnic Iranians. That was the case, for instance, with Adīb Naṭanzī's work called *Dastūr al-lughā* (or sometimes *Kitāb al-khalāṣ*), finished in 1090 and containing about 7,000 Arabic words, mostly glossed in Arabic. The work was composed for educated bilingual intellectuals like translators, poets, scribes, or officials so that Persian explanations were not absolutely necessary. The fact, however, that the headwords were arranged alphabetically by initial letters displays the author's detachment from the typical Arabic patterns. It was also in this dictionary that a few additional diacritical symbols were for the first time introduced in order to better render consonants in the adduced Persian words.¹⁸

Because of the specific morphological structure of Arabic words, the presentation of nouns (Arabic *asmā'* and *asāmī*, both plural forms of *ism* 'noun') and verbs (Arabic *maṣādir* ~ Persian *maṣādir*, plural of *maṣdar*, an Arabic verbal noun used for exemplifying verbal morphology) in the form of distinct wordlists became usual in the Arabic tradition and was also taken over by Persian lexicographers. The earliest *asāmī* list, titled *al-Sāmī fi'l-asāmī*, was compiled by Abū'l-Faḍl al-Maydānī in 1104. Here, the nouns are divided into four 'books' (*kitāb*): religion; animals; the celestial; the terrestrial. They are further subdivided into chapters (*bāb*) and sections (*faṣl*). An interesting fact is that 'the Persian glosses are absent from some of the manuscripts'.¹⁹ This seems to point to a two-phase system of compiling the wordlist: the Arabic part was probably written first as a monolingual list intended for Arabic-speakers, and the Persian glosses were added afterwards, changing the original monolingual list into a bilingual one. The first dictionary of the *maṣādir* type was one called *Kitāb al maṣādir*, written in the eleventh century by Qāḍī Ḥosayn Zawzanī. It included about 5,000 entries, but their arrangement following the Arabic patterns made it rather cumbersome and impractical for inexperienced users. A far more handy tool was *Tāj al-maṣādir* by Zawzanī's younger contemporary Abū Ja'far al-Bayhaqī, who divided his work into morphological sections, each with headwords organized by first letter. The *Tāj* included 10,000 entries, and it soon became very popular. Neither of these genres survives today. The *maṣādir* dictionaries disappeared in the fifteenth century, whereas the *asāmī* lists survived until the early eighteenth century.

As far as topically specialized dictionaries are concerned, we should start with the above-mentioned bilingual wordlists from the earliest times, such as

¹⁸ Perry, 'Early Arabic–Persian lexicography', 251–2.

¹⁹ Perry, 'Early Arabic–Persian lexicography', 254.

the Aramaic–Persian *Farhang ī pahlawīg*, dating from the tenth (?) century, with its thirty-one chapters arranged topically. However, it was, as a matter of fact, an explanatory book of Aramaic ideograms rather than a dictionary proper. Unlike that dictionary, the first Qurʾān glossary in Persian, *Tarjuman-i Qurʾān* (eleventh century), compiled by Qāzī Ḥosayn Zawzanī, can be deemed representative of topical dictionaries in the Persian tradition.

It is only natural that Turkish was, after Arabic, the second most important language for Persians, even though both nations could in many cases easily communicate with each other in Arabic. The situation resulted in producing two types of wordlists: Persian–Turkish on the one hand and Arabic–Persian–Turkish on the other. The oldest Persian–Turkish glossary was *Şihāḥ al-ʿAjam* by Fakhr-al-Dīn Nakhjawānī (c. 731/1330).

Persian Lexicography in India

As the result of a great migration of Persian intellectuals, attracted from the twelfth century at the latest by very good living and working conditions at Muslim courts in northern India, a new and important centre of lexicographical thought came into being in this area at the end of the thirteenth century. Persian-language lexicography continued to be undertaken there until the early twentieth century (for the eighteenth, nineteenth, and twentieth centuries, see Chapter 17, and cf. Chapter 18). A usual writing convention, especially in the initial period, was to inform readers, in the introductory part of a work, that the author had been ‘asked by his friends to compile a dictionary . . . so that they could easily read Ferdousi’s *Shāhnāmāh* (most frequently) and other outstanding works of Persian literature’.²⁰

Fakhr-al-Dīn Mubārakshāh Qawwās Ghaznawī is considered the father of Persian lexicography in India because he compiled the oldest extant Persian dictionary in this country. The date of his *Farhang-i Qawwās* is generally thought to lie between the end of the thirteenth century and the second decade of the fourteenth. The special significance attributed to this work in the context of the history of Persian lexicography results from the fact that it is chronologically only the second wordlist (after Tūsī’s dictionary) whose original manuscript has survived until today. The structure of the *Farhang* was already dated at the time of its compilation because the author divided his word material – somewhat more than 1,300 words – into five topical chapters (*bakhsh*) rather than alphabetically. That is why it was also called

²⁰ Maciuszak, ‘Persian lexicography’, 185–6.

Farhang-i panj bakhsh or *Panj bakhshī*. Another characteristic feature of this work is that, due to the author's puristic attitude, it almost exclusively contains Persian words, and even in the introduction his tendency to avoid Arabic words can be observed.²¹

The two-part dictionary *Adāt al-fuṣalā'* (compiled 1419 or perhaps as early as 1409) was mentioned in the discussion of the arrangement of Persian dictionaries above. An interesting detail concerning this work is the fact that its author, Qāzī Khan Dihlawī, was a scholar and previously a pupil of leading philologists of the time rather than just a poet.²²

One of the best-known dictionaries compiled in the first phase of Persian lexicography in India is *Farhang-i zafāngūyā wa jahānpūyā*, written by Badr-al-Dīn Ibrāhīm in the late fourteenth or early fifteenth century. The author starts his dictionary with an explanation of his task and a presentation of the principles according to which he organized the entries. This fact clearly points to an evolution of the macrostructure of Persian wordlists towards a rational and well-thought-out dictionary plan. Badr-al-Dīn's *Farhang* contains more than 5,000 words divided into seven topical parts (*bakhsh*) that are subdivided into chapters (*gūnah*) in which the words are arranged by first letter. Each *gūnah* is further divided into sections (*bahr*) organized according to the last letter of a specific word. For decades the dictionary was thought to have been lost, but between 1960 and 1990 four manuscript copies were discovered, three of them by one and the same researcher, Solomon I. Baevskij.

In the golden age of Persian lexicographical activity in India, numerous wordlists of various types were compiled, and it is no easy task to choose the most important ones. The importance of the *Farhang-i Jahāngīrī* (1608) by Ḥusayn Injū, mentioned in the discussion of the arrangement of Persian dictionaries above, comes – apart from its size of almost 10,000 entries – principally from its grammatical parts and its well-developed macrostructure: the dictionary proper is preceded by an introduction and followed by a supplement. The introductory part contains a philological presentation of the geography of Persian, its graphematic and phonological features, a sketch of its grammar, and a list of other dictionaries used by Injū, while the supplement is a collection of five wordlists, among others one with metaphorical and poetical expressions and one with foreign words, borrowed from Turkic, Greek, and the Indian languages. *Burhān-i Qāṭi'* ('Decisive proof', 1651) was compiled by Muḥammad Ḥusayn Tabrīzī, known as

²¹ Baevskij, *Rannjaja persidskaja*, 52. ²² Maciuszak, 'Persian lexicography', 190.

Burhān, apparently in Hyderabad (present-day India) and, unlike other Persian lexica of the time, it did not provide poetic quotations. Although not always accurate, it was a touchstone of the later lexicographical tradition (see Chapter 17). Another much-valued dictionary of those times was one written by ‘Abd-al-Rashīd, and thus called *Farhang-i Rashīdī* (1654). Its significance results, inter alia, from the fact of its being used by Orientalistic philologists and lexicographers in Europe such as Johann August Vullers and Francis Joseph Steingass.

Persian Lexicography in Europe

In this context, some words should also be said about early European dictionaries with a Persian section, even though they were never as numerous and substantial as the Turkish ones because Persian was much more distant and much more rarely needed in Europe. The oldest wordlist in this group is a *Specimen lexici Arabico–Persico–Latini* (Leiden, 1645) by Christian Ravius.²³ German by birth, Ravius tried to find jobs in various countries, failed to obtain the chair of Arabic at Oxford, and then moved to Sweden where he was employed as professor of Oriental languages in Uppsala. The quality of his knowledge of Arabic and Persian is sometimes contested and, it seems, with good reason. Nobody doubts that Ravius learned to speak Turkish during his stay in Istanbul. However, in his letters he fantasized about his alleged visits to monastery libraries in Aleppo because he tried to impress on his European colleagues his knowledge of Arabic, and it was for the same reason that he preferred to withhold the fact that he spoke Turkish rather than Arabic with his Arab friend, presented by him to his fellow writers as an Arab scholar from Aleppo.²⁴ Doubts about his real command of Arabic – as well as his illegal export of manuscripts from Istanbul which caused other Europeans serious problems with Ottoman authorities when they sought to obtain manuscripts – might have caused the failure at Oxford.²⁵ The originality of Ravius’ *Specimen* is therefore disputable and calls for further investigation.

Another wordlist which may be mentioned here is a collection consisting of two Persian glossaries, as well as a glossary of Persian numerals, in the third edition (London, 1664) of Sir Thomas Herbert’s *Some Years Travels into*

²³ A portrait made c. 1645, around the time of the publication of his *Specimen*, is reproduced in Kilpatrick and Toomer, ‘Niḡlāwus’, 122.

²⁴ Kilpatrick and Toomer, ‘Niḡlāwus’, 10.

²⁵ Kilpatrick and Toomer, ‘Niḡlāwus’, 12 and n. 41.

Divers Parts of Africa and Asia the Great.²⁶ Unlike Ravius, Herbert certainly collected his Persian materials from Persians in Persia. In 1627 he participated in a diplomatic mission to Persia. Even though the diplomatic aspect of the mission ended up as a failure, he took the opportunity to make extended tours in Persia. Having come back to England, he prepared a first version of his *Description of the Persian Monarchy* (1634) based on his field collections.²⁷

As early as 1669, an English Orientalist and professor of Arabic at the University of Cambridge, Edmund Castell, published a dictionary in London, at his own expense of £12,000. It included, among others, a Persian section, and was titled *Lexicon Heptaglotton Hebraicum, Chaldaicum, Syriacum, Samaritanum, Aethiopicum, Arabicum, et Persicum*.²⁸ The work on this dictionary lasted almost twenty years, from 1651 to 1669, and 'became Castell's obsession, taking precedence over all other tasks and other duties'.²⁹ Castell, being a Semitic – and first of all an Arabic – philologist, made up his own grave inscription in Latin with one line in Arabic.³⁰ The entries in his 4,000-page *Lexicon* were arranged according to the Hebrew alphabet. He never appeared to be especially fascinated by the Persian language. And, indeed, the Persian part of his *Lexicon* was actually a *Dictionarium Persico-Latinum*, found as a manuscript among the papers of Jacobus Golius, a Dutch Orientalist at the University of Leiden, after his death. Castell only supplemented the *Dictionarium* at some places and then included it in his own *Lexicon*.

Fifteen years later, Ange de Saint Joseph, a French missionary, linguist, and pharmacologist, published his *Gazophylacium linguae Persarum* (Amsterdam, 1684; some sources indicate 1685 but the note 'Anno 1684' can be easily read on the title page). This work is a grammar with an Italian–Latin–French–Persian wordlist and corresponding word indexes. Its very special feature is an attempt to show etymological connections of Persian with some European languages. Thus, in the third chapter, 'De linguae Persicae cum Europaeis analogia', one can find, among others, the following comparisons: Persian *amma* 'but' ~ Italian *mà*; Persian *beheter* 'better' ~ Dutch *beeter*; Persian *chizi* 'thing' ~ French *chose*; Persian *C-hoda* [= *khudā*] 'god' ~ Dutch *Godt*, and so on.³¹ A few equations (such as Persian *berader* ~ Dutch *broeder*) are correct, but Ange de Saint Joseph had no way of separating these from the chance resemblances which he observed.

²⁶ Considine, 'Wordlists', 368. ²⁷ Aune, 'Review', 231–3.

²⁸ Thomas, *Universal Dictionary*, 1.538. ²⁹ Norris, 'Edmund Castell', 71.

³⁰ A photograph of the inscription can be seen in Norris, 'Edmund Castell', 73.

³¹ Ange de Saint Joseph, *Gazophylacium*, 5–6.

A conspicuous point is that most of these wordlists – unlike their Ottoman Turkish counterparts, as we shall see – are no simple bilingual glossaries, compiled for those who needed to speak Persian, for instance for mercantile purposes. Nevertheless, the fact that those four dictionaries were published within forty years clearly shows that the interest in Persian was rather considerable in seventeenth-century Europe, while their content and structure hint at the scholarly and cognitive rather than practical character of that interest.

The Codex Comanicus

The Codex Comanicus is the oldest-known collection of documents in Cuman (Kipchak Turkic). The copy extant in the Biblioteca Nazionale di S. Marco in Venice dates back to the first half of the fourteenth century, but it has been suggested that it was preceded by an older one, made probably in 1303 in Turkey.³² Thus, the original text is likely to have come into being in the second half of the thirteenth century.

Part of the name of the Codex is an adjective mostly spelt *Cumanicus*, and less commonly *Comanicus*. This tradition seems to go back to Géza Kuun, the Hungarian editor of the 1880 edition of the Codex, a very popular one in his day, who favoured the form with *-u-* because it matched the Hungarian pronunciation *kún* ‘Cuman’ (he was himself descended from a Cuman family: *Kuun* is an old spelling of what is written *kún* today). He disregarded the fact that this ethnonym is written with *-o-*, that is *Comani*, in the Codex itself, for instance in the opening lines of the first page: ‘In hoc libro continentur persicum [sic] et Comanicum per alphabetum.’ Initially, Kuun’s form was not really accepted; so, for instance, Wilhelm Radloff used the *-o-* form in his study based on Kuun’s edition.³³ But then the *-u-* form became more and more popular. Despite the negative opinion of some modern Turkic linguists and the fact that newer editions use the *Comanicus* variant, the title of the Codex is usually written with *-u-* even today (cf. also the German word *komanisch*, used along with *Codex Cumanicus*, for instance by the Dutch Orientalist Martijn Th. Houtsma, just under fifteen years after Kuun’s edition, as well as nowadays, for example in the title of Kaare Grønbech’s book and that of Ingeborg Baldauf’s study; for English cf. ‘the Codex Cumanicus, a handbook of the Koman . . . language’).³⁴

³² Houtsma, *Türkisch-arabisches Glossar*, 5; repeated in Golden, ‘Codex Cumanicus’.

³³ Radloff, *Das türkische Sprachmaterial*, *passim*.

³⁴ For objections to the form in *-u-*, see Drimba, *Codex Comanicus*, 11–13, 35, and for the use of the form in *-o-* see the title of Drimba’s edition; for *komanisch*, see Houtsma, *Türkisch-*

The Codex comprises a few parts but not all of them are wordlists. The first part is a Latin (Italian variant)–Persian–Cuman glossary of 110 pages. Its authors were probably Italian Franciscans and merchants operating in Crimea, and that is why it is often called an Italian part. The manuscript is carefully written, with three clear columns, and the word material is divided by first letter, each time with a heading like ‘Hec sunt Verba et nomina de litera [sic] A’, and so on. The other part of the Codex, sometimes called the German one, consists of religious texts and a few riddles in the Cuman language with glosses in Latin, German, or both, as well as an eleven-page Cuman–(Middle High) German wordlist. However, this is hardly anything more than a working collection of words, not really divided into uniform sections, and written in somewhat irregular columns. The working technique was presumably to write the left column in Cuman first and to add German equivalents subsequently. The order of the words is rather chaotic: horizontal and diagonal lines connect some Cuman words with their German glosses written at different places on the page, vertical lines separate irregular columns from each other, some words are scored through or corrected, some are added afterwards. This is certainly, in contrast to the Italian part, a collection of lexical notes, compiled by a few different persons, rather than a ready and elaborate glossary.³⁵

It is somewhat unusual that the Venice copy contains not only a fair Italian manuscript but also a first draft of the German wordlist. One would rather expect a copyist to have made a fair copy of the German draft, too. Baldauf’s claim that the final edition of the *whole* Codex was made by a German-speaking person is, thus, more than questionable (apart from the fact that the German wordlist bears no traces of any editing at all).³⁶ Rather, a fair copy of the Italian part and the religious texts and riddles fell into the hands of some German merchants (this conjecture is based on the lack of religious terms in the German wordlist), who attached a few German glosses at various places of the texts and, furthermore, added their draft notes to the earlier compiled parts, thus combining originally different texts under one cover to produce what we today call the Codex Comanicus. However this may be, the Codex is very important for philologists and linguists because it is the only source of the thirteenth/fourteenth century in which Kipchak words are spelt with Latin characters so that their vowels can be seen clearly.

arabisches Glossar, 5, Grønbech, *Komanisches Wörterbuch*, title, and Baldauf, ‘Missionshilfe’, title; the English quotation is from Clauson, *Etymological Dictionary*, xxiv, par. 47.

³⁵ Baldauf, ‘Missionshilfe’, 16–17. ³⁶ Baldauf, ‘Missionshilfe’, 16a.

Other Turkic and Ottoman Turkish Dictionaries

As a matter of fact, exceptionally numerous wordlists, both long and short, were produced in the past to help Turks who spoke various dialects all over Central Asia and Anatolia to learn Arabic and Persian. It is, to be sure, a generally accepted opinion that the Turkic dialects very much resembled each other so that, for instance, a glossary of a given Kipchak dialect could have successfully been used by speakers of other Kipchak dialects as well as even by those of non-Kipchak dialects. The opinion is, to a certain extent, justified. Nevertheless, as we shall shortly see, sometimes Chagatay wordlists were compiled to enable the Ottoman Turks to correctly understand Chagatay literature.

A general feature of Turkic dictionaries was that they mostly had a didactic and practical character. The Turks thought of compiling instruments for teaching or learning foreign languages as well as translating Arabic and Persian works rather than of collecting and explaining their own vocabulary. It was only in the nineteenth century that the Ottoman Turks would start working on their descriptive dictionaries (see Chapter 17).³⁷

It is impossible to present all the Central Asian and Ottoman wordlists (from both Anatolia and Rumelia; *Ottoman* is used below in its chronological rather than its stylistic sense) of Turkic dialects, even though they are all generally considered a common part of, so to say (in a cultural rather than a political sense), a 'pan-Turkic' lexicographical tradition. I will try to concentrate on a few works viewed as most significant for different parts of the Turkic-speaking world.

One of the particularly famous dictionaries in this group is that compiled by Ġemāl-al-Dīn ibni Mühennā (or ibn Muhannā) in the late thirteenth or perhaps the early fourteenth century, and mostly called *İbni Mühennā Lügati* today. The work comprises three parts: Arabic–Persian, Arabic–Turkic, and Arabic–Mongolian. An especially interesting fact here is that ibni Mühennā continued, in a sense, Maḥmūd al-Kāšġarī's tradition of comparative study of Turkic in that he added various remarks concerning pronunciation along with, albeit more rarely, inflectional elements such as diminutive or participial suffixes typical of different Turkic dialects.³⁸

When Martijn Houtsma, edited, in the year 1894, an anonymous Arabic–Turkic glossary whose only copy has been preserved in Leiden, he mistakenly dated the manuscript to the year 1245, and it was only in 1968 that another

³⁷ Tietze, 'Lexikographie', 2400a.

³⁸ Gül, 'İbni Mühenna Lügati'nin Türk ve Moğol dil araştırmalarındaki yeri ve önemi', 89–92. For the newest edition of the Mongol part, see Ġemāl-al-Dīn ibni Mühennā, *Moğolca*.

Orientalist working in Leiden, Barbara Flemming, corrected this dating to 1343.³⁹ The title of this manuscript is rather long – *Kitāb-i maǧmūʿ-i terǧūmān türki ve ʿaǧamī ve moǧolī ve fārsī* – and that is why the work is usually just called Houtsma’s Glossary. It has seventy-six sheets; sixty-two are a Turkic–Arabic wordlist, and fourteen are a Mongolian–Persian one. Flemming’s correction changed the chronological status of the glossary in the history of Arabic–Mongolian lexicography. When it was dated to 1245, Houtsma’s Glossary was viewed as the oldest Arabic wordlist of Mongolian. After the correction it was demoted to second place, after ibnī Mühennā’s work. The author of these wordlists is unknown. Phonetic and orthographical features of the Arabic parts point to a person who was well acquainted with Egyptian Arabic.⁴⁰ Nevertheless, he need not have necessarily been an ethnic Arab, and his awareness of differences between at least two varieties of Turkic – namely Kipchak and ‘Türkmāni’, as he puts it, which we would identify as Oguz today – suggests instead that he was probably of Turkic, or more probably, Mamluk descent which would mean that his grandparents were Kipchak slaves. This conjecture fits very well with what he says about Turkic: Kipchak is, in his eyes, the correct and very pure form of Turkic whereas ‘Türkmāni’ is a corrupted variant, full of Arabic and Persian words.⁴¹ A comparison of this source with the Codex Comanicus leads to the conclusion that the language of Houtsma’s Glossary is actually nothing but a newer form of the dialect displayed in the Codex. In those days Kipchak, as a language of the Mamluk rulers, was the dominant language in Egypt, while Arabic was only a vernacular used by local peasants and craftsmen. Under these circumstances Houtsma’s Glossary may be considered to have been written as a teaching instrument for Arabic-speaking citizens who would have to communicate with the Kipchak-speaking Mamluk authorities. The aim of preparing the Mongolian–Persian wordlist is less obvious. Its shortness may possibly be interpreted as a signal of its working or draft character.

An important place among dictionaries popular in the Turkic-speaking world has been occupied by a work titled *Muḳaddimat al-adab*, which was originally an Arabic–Persian dictionary compiled by al-Zamahṣarī in the twelfth century, though its numerous readers more often than not used various later reworkings. Al-Zamahṣarī, though himself of Persian descent, promoted Arabic as a common language of the Islamic world because it was the language that had been selected by God. A very gifted and prolific author

³⁹ Houtsma, *Türkisch-arabisches Glossar*, 2; Clauson, *Etymological Dictionary*, xxv, par. 49.

⁴⁰ Houtsma, *Türkisch-arabisches Glossar*, 6. ⁴¹ Houtsma, *Türkisch-arabisches Glossar*, 4.

of linguistic and theological studies, al-Zamahšarī presumably wrote his *Muḳaddimat* for Persians who wished to improve their knowledge of Arabic. However, later on, some Turkic intellectuals added their own Turkic translations and thus made different trilingual copies of Zamahšarī's dictionary that even so continued to circulate under the same original title among the Turks. It was only in 1926 that a new copy of this work could be found in the former palace library of the emir of Bukhara. Unlike its predecessors, that one had four parts: Arabic, Persian, Chagatay, and what after some time proved to be Mongolian. The copy was made by a secretary (*kātib*), Darvīš Muḥammad, in 1492.⁴²

The so-called Chagatay language was, as a matter of fact a sort of a graphical *koine* of various Turkic dialects rather than a homogeneous language. Thanks to some specific features of the Arabic script (like unmarked or poorly marked vowels and ambiguity of some consonant letters) as well as relatively regular phonetic correspondences between Turkic languages and dialects, a text written in the Chagatay *koine* orthography could easily be read by speakers of various vernacular idioms. The problem became somewhat greater when speakers of more distant dialects, for instance those from the Ottoman empire, wished to read Chagatay poetry, such as poems by ʿAlī Šīr Nevāī, the more so as not every Chagatay word was equally well known to them. This fact explains the popularity of Chagatay dictionaries in the Ottoman Turkish milieu. The best-known wordlist in this group is a Chagatay–Ottoman Turkish dictionary of more than 2,000 loanwords. It was compiled in Anatolia around 1534. Since the name of the compiler is unknown to us the dictionary is usually called *Abuška Lüḡati* or *Abuška Sözlüğü* after the first word it explains: Chagatay *abuška* 'old man'. The primary aim of the compiler will, in all likelihood, have been to help Ottoman Turks to correctly understand Nevāī's poetry. Nevertheless, he also treated vocabulary from works by thirteen lesser-known Chagatay poets.⁴³ Thus, *Abuška Lüḡati* seems to be the first Turkic dictionary written for literary purposes. Furthermore, it is a dictionary of two Turkic idioms (not Arabic or Persian and Turkic). And in its primary focus on the works of Nevāī, *Abuška* is a forerunner of later Turkic dictionaries devoted to one author only. It is therefore easily understandable that it has aroused considerable interest among Turkic philologists.⁴⁴

⁴² Poppe, 'Viersprachige Zamaxšarī-Handschrift', 301–2. ⁴³ Kaçalin, 'Abuška sözlüğü', 24.

⁴⁴ For the most important editions, see Ölmez, 'Çağatayca sözlükler', 13–14.

About a hundred years later another very special non-Arabic and non-Persian wordlist was compiled in the Turkish world. At the opposite end of the Ottoman empire, in Bosnia, the Slavic population was rather rapidly converted to Islam after their kingdom, left without help from European countries, fell in its unequal fight against the Turks in 1463 and became a new part of the empire. One of the results of this historical fact was that, a century and a half later, one of the Muslim Bosniaks, Muḥamed (Bosnevī) Hevāi Uskufī, compiled a Bosnian–Turkish dictionary, *Maḳbūl-i ʿarīf* (with the Arabic script used for both languages). It was ready in 1631, and this date is considered to mark the beginning of Bosniak and, more generally, of Balkan Slavic–Ottoman Turkish lexicography.

That is, however, not true of Balkan Slavic lexicography in general. For instance, Pietro Lupis Valentiano published as early as in 1527 a work titled *Opera nuova che insegna a parlare la lingua schiavionesca*, being a sort of an introduction into Čakavian Croatian, and this glossary may be viewed as the oldest lexicographical record of a Balkan Slavic idiom. It is noteworthy that the same Lupis published, probably also in 1527, an introduction to Turkish: *Opera nova . . . La qual insegna a parlare Turchesco*.⁴⁵ Also Hieronymus Megiser's *Dictionarium quatuor linguarum* of 1592 had, besides German, Latin, and Italian, also a separate Slovene part. Another work by Megiser, namely *Thesaurus polyglottus vel dictionarium multilingue* (1603), also offered Turkish lexis.⁴⁶ Interest in Turkish was tied, as we can see, more than once to interest in Slavic languages; as we shall see, this is also true of François à Mesgnien Meninski's activity.

As to the pen name *Uskufī* (whose Turkish pronunciation is *Üsküfī*) its meaning is still discussed. If the old interpretation, that it is a distortion of an original *Üsküpī* 'descendent from Skopje', should prove correct, his family will originally have come from Macedonia to Bosnia.⁴⁷ *Uskufī*'s wordlist is sometimes called a Turkish–Bosnian rather than a Bosnian–Turkish dictionary. Actually both are correct, because it is a versified dictionary in which each of 330 verses (with more than 700 lexemes) can start either with a Bosnian or with a Turkish headword, so that the work could better be called a dictionary of Bosnian and Turkish.⁴⁸ In this respect the *Maḳbūl-i ʿarīf* belongs to a rather long tradition of Near Eastern didactic glossaries in verse. The oldest-known Turkish one from Anatolia is an Arabic–Turkish wordlist compiled in the late thirteenth century by ʿAbd-al-laṭīf (Abdüllatif) ibni

⁴⁵ Rocchi, 'Turkish historical lexicography', 209.

⁴⁶ For more details see Considine, *Dictionaries in Early Modern Europe*, 291–3.

⁴⁷ Okumuş, 'Muhammed Hevâi Üsküfî', 825. ⁴⁸ See Okumuş, 'Muhammed Hevâi Uskufî'.

Melek, and the last one is a French–Turkish *Mitfāh-ı Lisân* (1849) by Yūsuf Ḥālîş.⁴⁹ Since it is somewhat unusual for a European reader to see a dictionary with French words written in the Arabic script, the first *beit* of that dictionary may be of interest here, in transliteration and with italic equivalents in modern French orthography:

Allah	diyö,	gökler	siyö,	yer	ter,	komanse	ibtidâ,
<i>dieu</i>		<i>cieux</i>		<i>terre</i>		<i>commencer</i>	
Dâ'im	tužur,	bâki	eternel,	enfini	bî-intihâ		
<i>toujours</i>		<i>éternel</i>		<i>infini.</i> ⁵⁰			

Here and elsewhere, some of the French is difficult to recognize: for instance, the word *Lebsom*, which is given as a 'French' equivalent of Ottoman Turkish *Zebûr* 'King David's Book', stands for French *Les Psaumes*.⁵¹

Ottoman Turkish Transcription Dictionaries

The English term 'transcription texts', rendering the original German expression *Transkriptionstexte*, is used to denote all texts written in the Ottoman Turkish language but in a non-Arabic script before 1928, which is the year when a Latin-based Turkish alphabet was introduced. The term has been criticized many times, because these texts are in actual fact directly written in a non-Arabic script rather than transcribed, but it is commonly accepted nowadays. The importance of this group of linguistic sources is especially great because texts in non-Arabic scripts could, in various ways, render actual Turkish pronunciation better than those in the Arabic script. Besides, they were generally authored by Europeans who had to communicate with various Turks, often non-elite speakers, on everyday subjects. This means that transcription texts display dialectal words and daily phrases such as, for instance, names of tools, formulas used in bargaining, or even swear words, which stood almost no chance of turning up in Ottoman poetry.⁵² Of course, not every transcription text was a wordlist but many of them included one.

The Codex Comanicus should be viewed as the first-known transcription dictionary of a Turkic language, but that term is generally used in regard to sources in Ottoman Turkish only. So, apart from single glosses in fifteenth-

⁴⁹ Okumuş, 'Anadolu sahasında', 152. For a list of other Turkish versified dictionaries, see Çınar, *Mitfāh-ı Lisân*, 19–30.

⁵⁰ Çınar, *Mitfāh-ı Lisân*, 79, Beit 1. ⁵¹ Çınar, *Mitfāh-ı Lisân*, 79, Beit 3.

⁵² For a general introduction, see Rocchi, 'Turkish historical lexicography', 206–7.

century travelogues, the first transcription dictionary in the full sense of the word was a three-volume *Regola del parlare turcho* [sic] compiled in 1533 by Filippo Argenti, a chancellor in the Florentine embassy in Constantinople. The *Regola* includes a Turkish grammar and two Italian–Turkish wordlists: nouns in volume I, and verbs in volume II. Together the lists contain about 4,200 entries.⁵³

Hieronymus Megiser's *Thesaurus Polyglottus* (1603) was mentioned above. But it was Italians who were doubtless the most active authors of Turkish wordlists and other language materials, especially in the seventeenth century. This fact was mostly connected with Italian diplomatic, mercantile, and missionary activities in the Ottoman empire. Cases of slavery as a motive for lexicography were relatively rare but they also occurred. Such is the instance of Pietro Ferraguto who lived for six years as a prisoner of war among Turks. After release he joined the Jesuits and then compiled a three-part work intended to be a linguistic tool for missionaries, comprising a grammar, a phrasebook, and a *Dittionario della lingua Turchesca* (1611).⁵⁴

In the same year of 1611 a somewhat curious *Notebook* was compiled which is now preserved in the university library in Wrocław. Neither its author nor its title is known, so that the title *Notebook* is just a matter of convention. It contains several parts, one of them being a Spanish–Turkish wordlist. Here, a specific feature is that Spanish (and other European words, added at various places) are written in brown, whereas the Turkish parts are in black. The mysterious author stayed, together with his brother, for almost two years in Turkey where he learned Turkish from a Turk and, separately, from an Armenian. Spanish was probably not his mother tongue, as he relatively often added Italian, French, Latin, or German equivalents of the Spanish words in his glossary, and verbal forms occurring in dialogues display Italian suffixes added to Spanish verbs.⁵⁵

Giovanni Molino, who compiled a *Dittionario della lingua Italiana–Turchesca* (1641), only partially belongs to the Italian group of lexicographers. His dictionary with its c. 15,000 entries divided into 494 columns was an important and huge wordlist in Molino's day. Its specific feature was the use of colloquial Turkish only. But the strangest thing about this *Dittionario* was that, for a long time, no European Turcologist knew who Molino was, apart from the fact that he earned a living as a so-called dragoman, or interpreter, working for the Republic of Venice. Today we know that he was not an

⁵³ Rocchi, *Ricerche*, 1–3; Rocchi, 'Turkish historical lexicography', 209–11.

⁵⁴ Rocchi, 'Turkish historical lexicography', 213–14; Rocchi, in Ferraguto, *Dittionario*, 7–9.

⁵⁵ Majda, *Rozwój języka tureckiego*, 24–6.

Italian at all. Instead, he was descended from an Armenian family, and his true name was Yovhanes Ankurac'i (or, Ankivrac'i), that is 'from Ankara'. He was born in 1592 but then he lived and studied for some time in Rome before moving to Venice.⁵⁶ His good knowledge of colloquial Turkish and simultaneous ignorance of cultivated Ottoman Turkish point to his living in Turkey outside Ottoman educational institutions. That conjecture fits very well with what one can learn of Ankurac'i from Armenian sources: namely, that he was a great enthusiast of Armenian books and made every effort (albeit to no effect) to induce the Vatican to print an Armenian Bible. The failure did not discourage him from dreaming of printing Armenian books. When in Venice, he was at the same time busy with both dragoman's duties and his printing matters, and he actually did manage to print some other books in Armenian.⁵⁷

Arcangelo Carradori was another Italian cleric author. After an eight-year stay in Egypt he returned to Italy and wrote two dictionaries: an Italian–Turkish and an Italian–Nubian one. Because the title page of the Turkish wordlist is lost we can only say, on the basis of a colophon, that the extant work is a third copy, corrected and personally rewritten by Carradori in 1650. The structure of the dictionary is worth mentioning. Each page is divided in three columns. The central one is in Italian, the left one is for Turkish in Arabic script, while the right one presents a Latin transcription of the given Turkish word.⁵⁸

Even though a full inventory of all wordlists, including also short glossaries in principally non-lexicographical sources, would be rather long, the depiction of what Europeans actually knew of Turkic (mostly, Tatar and Turkish) matters could be complete only if it included commentary on manuscripts and printed works that contain Turkic words but no Turkic wordlists. That is a task for the future.

Meninski's *Thesaurus* (1680)

The fact alone that a Turkish Turcologist, Mehmet Ölmez, decided to publish a photomechanical reprint of Meninski's dictionary in 2000 – 320 years after its first edition in 1680 – because it was missing from Turkish university libraries points to the importance of this extraordinary work. Its author, François à Mesgnien Meninski, is an exceptional personality. He can

⁵⁶ Świącicka, 'Interpreter', 329, 339, 341. ⁵⁷ Iszchanian, *Książka ormiańska*, 63.

⁵⁸ Rocchi, 'Turkish historical lexicography', 214–15; Rocchi, in Carradori, *Dizionario*, 9–15.

briefly be characterized as a Polish linguist of French (or Lorrainian) descent, who spent somewhat more than half of his life in Vienna (he has wrongly been called a 'German Orientalist').⁵⁹

François à Mesgnien was born in 1620 or 1623 in Totainville in Lorraine.⁶⁰ He studied in Rome with a Jesuit scholar, Giovanni Battista Giattini. As well as subjects typical of those days such as logic, mathematics, and theology, he also learned foreign languages. The reason for his moving, in about 1646, to Poland is not precisely known. Be that as it may, he learned Polish and published a Polish grammar as early as 1649. In the same year a French and an Italian grammar also appeared, and a new version of the Italian grammar was published in 1651. In all these books the author's name is given as Franciscus Mesgnien Lotharingus. In the decade between the last Italian grammar and his moving to Vienna in 1661 or 1662, he Polonized his name into *Meninski* and regularly used this form afterwards, even in his Vienna years. Actually the Polish form should be written with a diacritical mark: *Meniński*. Its omission was probably a technical device aiming at making the spelling of the name easier for Austrian and other Western European readers.

Towards the end of 1653, Meninski participated in a Polish legation to Istanbul. He stayed for two years there as an official delegate of Poland, and practised his Turkish both in conversation with Turks and making use of advice and help from a Pole, Wojciech Bobowski, who, under the name of Ali (Ufkī) Bey, acted, inter alia, as the First Dragoman to the Ottoman Divan. Soon after Meninski's return to Poland in 1656 he was appointed as a First Interpreter of the Polish Royal Household and visited Turkey a few times in the course of international affairs. He also worked out a project of a school for interpreters and diplomats for the Polish Oriental service. Unfortunately, this plan could not be implemented because of financial problems. It may be conjectured that the failure of his project irritated Meninski so much that he decided to move to Austria and enter the service of the house of Habsburg. That is, however, nothing more than a conjecture since no documents concerning Meninski's decision seem to be known.

The elaboration of his Turkish-language materials and compilation of the dictionary lasted for ten years. Arabic types were cast according to his design. The three-volume *Thesaurus* and a one-volume *Grammatica Turcica* appeared in 1680. A fifth volume was probably ready when the printing house was burnt during the Battle of Vienna in 1683. That is why the last volume, titled

⁵⁹ Chalmers, *General Biographical Dictionary*, XXII.52.

⁶⁰ Biographical data are mostly gathered from Stachowski, 'François à Mesgnien Meninski', *passim*.

Complementum . . . seu . . . index verborum, could appear only in 1687. An extended version of the *Grammatica Turcica* was republished in Leipzig in 1756, and a second, partially extended edition of the *Thesaurus* appeared in Vienna between 1780 and 1802. The size of the dictionary is evident from the number of consecutively paginated columns: 6,080 in all. A single column averages seven entries, which makes a total of about 42,500 entries. By contrast, the total number of lexemes is hard to determine because many Turkish *izafet* nouns are placed under the main noun. For instance, *deniz dili* ‘gulf’ (literally, ‘sea tongue’) is only adduced under *dil* ‘tongue’ (col. 2,114). But a far more imposing feature is the informativeness of the entries as well as the quality of the lexicographical accomplishment.

A typical entry contains a Turkish word rendered in the Arabic script (various orthographic variants are allowed for) and transcribed with Latin characters. Then, morphological information and more often than not synonyms follow. The meanings are given in Latin, German, Italian, French, and Polish. What follows are examples of usage in the form of sentences or idiomatic phrases. Both specific words and phrases are, if needed, additionally explained or provided with supplementary remarks. The Latin transcription used is no simple transliteration. Instead, Meninski always aims at showing the real pronunciation. For instance, under the Arabic sequence ⟨bk⟩ that can be read in several ways (first of all, *bek* or *beng*) he puts the transcription *ben* ‘birthmark, naevus’ (col. 853). Meninski’s transcription system is sometimes a little misleading: for instance, the use of ⟨æ⟩ for a short *a*, and the double function of both ⟨ü⟩ and ⟨y⟩ are certainly not quite intuitive.⁶¹ At some places Meninski’s notations, remarks, and explanations may astonish a reader. Nonetheless, they should never be ignored. Quite the contrary, a future long series of analyses of the Turkish data in Meninski’s *Thesaurus* is only starting to appear in print at the time of writing.

⁶¹ Stachowski, ‘Remarks’, 195–6.

Byzantine Greek

STEFANO VALENTE

The Byzantine millennium is customarily defined as beginning with the closure of the Neo-Platonic school in Athens ordered by Emperor Justinian in the year 529, and ending with the capture of Constantinople in 1453 by the Ottoman army led by Sultan Mehmed II. Of course, Greek lexicography existed before Byzantium (see Chapter 5). Furthermore, the Byzantine legacy extends into the European Renaissance, affecting both the Greek-speaking communities and European scholars who learned Greek. The present chapter covers those regions of the Byzantine empire where Greek was spoken, that is to say *grosso modo* Greece, Asia Minor, and southern Italy. Constantinople and its institutions formed one of the main centres for the preservation and evolution of Byzantine Greek lexicography throughout this time. Some other cities and regions – such as Thessalonica and the ‘Terra d’Otranto’ in southern Italy – also played an important role in the production and dissemination of certain lexicographic works. The texts which will be discussed in this chapter are preserved in parchment or paper manuscripts, usually in codex form, dating from around the ninth century up to the sixteenth.

In the Byzantine world, lexica usually served two main purposes: to help explain and understand the diction of biblical and literary texts, and to help learn and correct the Greek (literary) language. Among the huge lexicographic production that has been preserved until today, different typologies and various grades of complexity can be found in terms of content, matched to the respective intended readership, whatever its level of expertise. In fact, examining Byzantine lexicography entails dealing with a crossover of educational and learned traditions. Its complexity is further due to the fact that Byzantine scholars and grammarians, while producing new lexica, used to incorporate and rework material from different existing lexica which they considered authoritative. To

I would like to thank Klaus Alpers, Christian Brockmann, Ian Cunningham, Daniel Deckers, and Leonardo Fiorentini for their valuable remarks and suggestions on a first draft of this paper.

the authors, innovation did not consist in adding first-hand information on the basis of fresh investigations, but rather in a careful selection and utilization of the available sources. In this respect, a remarkable feature of Byzantine Greek lexicography is the mere amount of conservation and re-use of pre-existing materials. Something similar happened when lexica of the past were copied into new manuscripts: since lexica were perceived as working tools for educational, scholarly, and practical goals, every copyist felt free to alter the text of his model, for instance by cutting some information – especially quotations from literary texts or the attribution of the sources, be they grammarians or literary authors. Furthermore, supplying explanations from other grammatical, lexicographic, or literary sources was also customary. Therefore, almost every manuscript copy of every Greek lexicon can to some extent be considered a new issue or even a new edition of a given work. The original text was perceived as a sort of commons, whose contents every copyist and user felt entitled to modify and improve.

So, when studying a given lexicon, each copy and each manuscript should be studied both in its own right and in relation to tradition – and, in the latter case, not only in relation to the textual tradition of a given work, but also in relation to other lexicographic traditions. An analysis of the context of its production and use is also required. This complexity is one of the reasons why one of the most experienced connoisseurs of this subject, Klaus Alpers, defined Greek and Byzantine lexicography as a sort of jungle into which – to paraphrase – many scholars dared to enter only reluctantly, and in which, without a good guide and adequate preparation, many of them managed to get lost.¹ The study of the textual tradition of a Byzantine lexicon can be quite tiresome, and the final result – a critical edition of the lexicon in question – may sometimes be seen as a ‘piece of dry scholarship: no-one will read it for pleasure, but some may find it useful’, to quote from the preface to the edition of the *Synagoge* by Ian Cunningham.² Nevertheless, diving into the lexicographic materials and manuscripts can be an exciting journey through a multifarious field of investigation covering different areas of expertise, such as cultural, linguistic, and philological studies.

¹ Alpers, ‘Griechische Lexikographie’, 14: ‘Die griechische Lexikographie des Altertums und des byzantinischen Mittelalters ist wegen der nur bruchstückhaften Erhaltung und der überaus komplizierten gegenseitigen Beeinflussungen und z.T. unklaren Abhängigkeiten eine Art Dschungel, in den sich auch die meisten klassischen Philologen nur sehr ungern hineintrauen und in dem manch einer, der sich ohne nötige Vorbereitung und Führung durch Spezialisten hineingewagt hatte, kläglich gescheitert ist.’

² *Synagoge* (2003), 1.

In most cases, the typical structure of an entry in a Byzantine Greek lexicon is composed of a lemma followed by an explanation of differing length and content, depending on the purposes of each work or copy. Concerning the organization of the entries, a fully alphabetical arrangement is by no means the rule, but rather an exception, in Byzantine lexicography.³ We can find lexica whose entries follow a first-letter arrangement, others using a two- or up to four-letter ordering (for more on levels of alphabetization, see Chapter 13). Later interpolations and reworkings of the original text usually led to an alteration of the alphabetical arrangement, not necessarily resulting in a single consistent ordering.

Obviously, this chapter is not intended as an exhaustive presentation of the whole lexicographic production in the Byzantine world, but rather as a panorama of the different typologies and main achievements in this field of knowledge during the Byzantine millennium. Bilingual lexica, such as Greek–Latin ones, will not be taken into account (Hebrew–Greek lexica are mentioned briefly in Chapter 9, and Greek–Latin lexica in Chapter 13).⁴ As far as possible, lexica will be presented according to their typology and, within each group, in chronological order.

General Lexica

This section intends to give an overview of the main Byzantine lexica. Their extent and general content made them reference works for the study and learning of the Greek language from a grammatical, lexicological, and rhetorical point of view.

One of the most widely circulated lexica during the Byzantine age was probably composed at the very end of late antiquity, that is to say, around the fifth/sixth century: the so-called lexicon of St Cyril.⁵ In some manuscripts, it bears a title which means ‘Alphabetic lexicon of our Holy Father Cyril, archbishop of Alexandria’. Cyril, patriarch of Alexandria from 412 to 444, was an influential personality, later considered as one of the Church Fathers.

³ On alphabetization in Greek and Byzantine lexicography, see Daly, *Contributions to a History of Alphabetization*; Alpers, ‘Review of Daly, *Contributions*’; Valente, ‘Alphabetical dictionaries’.

⁴ Some Latin entries, especially juridical ones, can be traced back to Greek lexica, such as that of Hesychius.

⁵ See Drachmann, *Überlieferung des Cyrillglossars* (which includes excerpts); Hesychius, *Lexicon*, I.xliv–li; Hunger, *Hochsprachliche profane Literatur der Byzantiner*, II.37–9; Alpers, ‘Griechische Lexikographie’, 24–5 with n. 52; Alpers, ‘Handschriftenfund zum Cyrill-Glossar’; *Synagoge* (2003), 43–9; Dickey, *Ancient Greek Scholarship*, 100–1.

Whether the lexicon was composed by him, or under his guidance, or if it originally collected lexicological explanations to supplement some of his works, is still an issue under discussion. In any case, the lexicon may have been produced in Alexandria. Neither has the lexicon been edited as a whole yet, nor has its extremely complex textual tradition been conclusively studied. More than seventy codices are still preserved: their texts are very heterogeneous, and modern scholars group them into different families. One of the oldest surviving complete manuscripts, now preserved in the State and University Library of Bremen, has been recently rediscovered and studied, and a transcription of its text is available online.⁶ This manuscript dates back to the late ninth century and was probably copied in the Monastery of Stoudios in Constantinople, an important religious and cultural centre during the Middle Byzantine age. In general, the lexicon is of quite elementary character and served as a basic dictionary. The bulk is composed of explanations of the vocabulary of the Old and New Testament, and of Christian authors – among them St Cyril. Its core body has constantly been augmented, with glosses to Homer and to other classical writers, as well as with Atticist materials (we shall return to Atticist lexicography below). In the surviving manuscripts, the text usually exhibits mistakes and double glosses. The entries are arranged according to the first three letters, but there are numerous misplaced groups of items, most of them due to what appear to be later additions. Its diffusion can be measured not only through the huge manuscript tradition, but also because it was used as a source in many later Byzantine lexicographic works. Paul Maas, one of the most eminent twentieth-century scholars of ancient and Byzantine Greek, referred to it as ‘the most terrible lexicon ever’.⁷ However, it may become a valuable source for understanding the educational practices in the Byzantine world over centuries and within different social groups, provided that its whole manuscript tradition is studied and the text of the different versions is edited according to modern standards.

Probably around the same period, that is on the threshold of the Byzantine age, between the fifth and sixth centuries, Hesychius of Alexandria composed his huge lexicon.⁸ It is probably the best-known and most-studied Greek

⁶ Bremen, Staats- und Universitätsbibliothek, msc 0011, edited as *Lexicon of Cyril* (2005).

⁷ Maas, ‘Review of Drachmann, *Überlieferung*’, 380: ‘an diesem greulichsten aller Wörterbücher’.

⁸ Hesychius, *Lexicon* (1953–2009), is the standard edition (vol. I has now been replaced by the second edition of 2018); see also Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 35–6; Alpers, ‘Griechische Lexikographie’, 25; Dickey, *Ancient Greek Scholarship*, 88–90; Valente, ‘Hesychius’.

lexicon, and it still exerts an enormous influence on both the literary and the linguistic study of ancient Greek. It is preserved in a single manuscript, copied in Constantinople around the year 1430, now kept in the Marciana Library of Venice.⁹ After reaching Italy, the codex constituted the basis for the *editio princeps* by Aldus Manutius, which appeared in Venice in 1514.¹⁰ In preparation for the printing of this edition, the Cretan copyist and scholar Marcus Musurus revised the text on the manuscript itself, marking it up with his corrections and annotations. Many studies on the text of the lexicon and on its tradition within the Greek and Byzantine lexicography have demonstrated that this codex preserves a shortened and reworked version, in which about a third of the entries have been interpolated from the lexicon of Cyril, and other additions include glosses to classical authors such as Homer and Euripides, to the Bible, and to the Church Fathers. Moreover, the text contains many troubling mistakes and problems. The original form of the lexicon is therefore definitely lost. In accordance with usual practice in Greek and Byzantine lexicography, Hesychius introduces the lexicon with a prefatory letter, addressed to Eulogius, whose identity unfortunately cannot be ascertained. In this letter, Hesychius states his compositional methodology and the sources he used: alongside lexicographic works collecting glosses on Homer as well as on the tragic and comic poets, mainly those of the classical age, his principal source is the second-century lexicon by Diogenianus (itself an abridgement of that of Pamphilus; for both of which see Chapter 5), whose work Hesychius enthusiastically celebrates. In particular, Diogenianus transmitted both literary and proverbial expressions as well as dialectal glosses from different sources. The alphabetical arrangement according to the first three or four letters is also a legacy from Diogenianus' lexicon. In the surviving version of Hesychius' work, this arrangement has, however, been obscured by interpolations. Furthermore, in the dedicatory epistle Hesychius remarks that he engaged in correcting the deficiencies he detected in the work of his predecessor, such as the lack of explanations of proverbs and polysemic words and the omission of the names of the authors quoted in the entries. Interestingly, the lexicographer stresses that he carried out the task of copying the work 'with his own hand', taking particular care with the orthography and accentuation according to the doctrines of Aelius

⁹ Venice, Biblioteca Marciana, ms gr. Z 622; see Smith, 'Note on Holkham Gr. 88'; Wilson, 'Transmission', 372–3; Hesychius, *Lexicon*, III.xx–xxi, and IV.xi–xiii; Speranzi, 'Copista del lessico di Esichio'.

¹⁰ See the Latin preface, with an English translation, in *Aldus Manutius: The Greek Classics*, 258–61.

Herodianus, probably the most influential Greek grammarian of the second century. To us, the lexicon is a primary source for the study of Greek dialects and for the text of classical and Christian authors.

The *Synagoge* ('Collection' or, more fully in Greek, 'Collection of useful words') is another fundamental scholarly work, produced in the eighth or ninth century, very likely in Constantinople.¹¹ The unknown compiler or compilers aimed at producing an enlarged and fully alphabetized version of the lexicon of Cyril. Only four manuscripts survive, dating from between the ninth and the fourteenth centuries. Shortly after its composition, the lexicon was substantially expanded with additions from many other readily identifiable sources, such as Atticist and rhetorical lexica, more than one expanded version being extant.¹² The additions were most likely made in Constantinople: the latest editor of the *Synagoge*, Ian Cunningham, has suggested the school in the Magnaura, near the Great Palace, where another crucial lexicon, the so-called *Etymologicum genuinum* (for which see below), was possibly composed. The *Nachleben* of the *Synagoge* is in turn impressive since it constitutes the foundation of some of the most important Byzantine lexica of the ninth to the eleventh centuries, such as the lexicon of Photius, the *Suda*, and the *Etymologicum genuinum*.

Photius, twice patriarch of Constantinople in the ninth century, is also a person of interest in the history of lexicography. On the one hand, he owned and studied many lexicographic manuscripts: descriptions and synopses can be read in his *Bibliotheca*, a sort of collection of book reviews *ante litteram* composed during his later years, where he describes books he and his scholarly circle have read. On the other hand, Photius also composed his own lexicon during his youth.¹³ It is a compilation from different sources, some of them no longer extant. The lexicon of Diogenianus, mentioned above, and two different versions of the expanded *Synagoge* constitute the core of this work. Photius brought together scholarly materials from these sources and enlarged them by inserting glosses taken from other lexica. To judge from the extant manuscripts, Photius' work enjoyed little circulation

¹¹ *Synagoge* (2003) is the standard edition; see also Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 37; Alpers, 'Griechische Lexikographie', 25–6; Dickey, *Ancient Greek Scholarship*, 102.

¹² One of them is preserved in Paris, Bibliothèque Nationale, ms Coislin 345; on this manuscript, see Valente, *Antiatticist*, 6–12.

¹³ Photius, *Lexicon* (1982–2013), covers the range alpha to phi at the time of writing, and is supplemented by Photius, *Lexicon* (1822); see also Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 39–40; Wilson, *Scholars of Byzantium*, 89–119; Dickey, *Ancient Greek Scholarship*, 101–2; Pontani, 'Scholarship in the Byzantine empire', 331–7.

during the Byzantine age, probably because of its private character – that is, its being a tool for ‘rhetorical and linguistic training’ within the scholarly circle of the author himself.¹⁴ Later in his life, Photius looked back at his own lexicographic activity, describing it in these terms:

one would compile a long book, not only if one wished to embrace all the polysemic words (a laborious task, next to impossible), but even if one wished to collect in one place the most common of them, those surfacing more often in literature: precisely this I did, as you know, when I was quitting the age of childhood.¹⁵

The main codex, direct or indirect ancestor of the whole surviving textual tradition, is a tenth-/eleventh-century copy now preserved in Trinity College, Cambridge, known as Codex Galeanus from the name of a former owner, the seventeenth-century English scholar Thomas Gale.¹⁶ Since it is now incomplete due to the loss of some quires, later copies are of great importance for recovering the full text.¹⁷ In November 1959, the Greek scholar Linos Politis discovered a complete manuscript of the lexicon in the monastery of Zavorda in Grevena, in the Greek region of West Macedonia. Thanks to this new codex, dating to the thirteenth/fourteenth century, the edition of the lexicon now relies upon a better and more reliable manuscript basis. The lexicon of Photius enjoyed a wider circulation in Europe during the seventeenth and eighteenth centuries in the form of manuscript copies of the Codex Galeanus, scrutinized mainly as a source for fragments of lost poetry and prose writers from the classical and imperial ages.¹⁸ A recent project, *Photios On Line*, is now beginning to produce an online English translation with some annotations of the Greek text.

Also related to the expanded *Synagoge*, but with a much more general character, is the encyclopedic lexicon *Suda*.¹⁹ Its influence on Byzantine culture – and later during the European Renaissance – is impressive. This lexicon was composed in the tenth or eleventh century, probably around

¹⁴ Pontani, ‘Scholarship in the Byzantine empire’, 332.

¹⁵ Photius, *Amphilochia*, 21.132–6; transl. Pontani, ‘Scholarship in the Byzantine empire’, 333.

¹⁶ Cambridge, Trinity College, MS O. 3. 9; see Photius, *Lexicon*, I.xxx–xxxii. It was dated to the eleventh century by Fonkič, *Grečeskie rukopisi evropejskich sobranij*, 50–2; I am rather inclined to date it to the tenth.

¹⁷ Tsantsanoglou demonstrated that all the extant manuscripts of this lexicon derive ultimately from the Galeanus (see Tsantsanoglou, *To lexiko tou Photiou*, 76–102; Photius, *Lexicon*, I.xxx–xxxiv).

¹⁸ Tsantsanoglou, *To lexiko tou Photiou*, 38–60.

¹⁹ *Suda* (1928–38) is the standard edition; see also Adler, ‘Suidas’; Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 40–2; Alpers, ‘Griechische Lexikographie’, 26–7; Wilson, *Scholars of Byzantium*, 145–7; Dickey, *Ancient Greek Scholarship*, 90–1.

the year 1000. The meaning of the title is still a mystery.²⁰ An alternative title, attested in literary and manuscript sources from the twelfth century onwards, is *Suidas* – it was formerly supposed that this was the name of the compiler of the work – and this title was adopted by the Danish scholar Ada Adler, who produced a splendid edition of the text, which is still the standard one. The *Suda* is the most extensive surviving Byzantine lexicon, with 31,342 entries of different length, arranged according to contemporary Byzantine pronunciation, in the so-called antistoechic order.²¹ (So, for instance, while the last two letters of the Greek alphabet are normally psi and omega, in Byzantine pronunciation, the contrast between omega [o:] and omicron [o] had been lost, so that in the antistoechic order, omega follows immediately after omicron, leaving psi at the end of the alphabet.) Its place of production was likely Constantinople. The compilers had many sources at their disposal: from those, in an almost entirely mechanical fashion, they collected and extracted the useful materials and arranged them into a new work. Among the sources, lexicographic works can be found – for instance the extended *Synagoge* and a lexicon related to that of Diogenianus – and collections of proverbs. Commented editions of Greek literary authors were also excerpted, together with philosophical, religious, and historical sources. An important further source for the *Suda* is the so-called *Excerpta Constantiniana*, that is collections of excerpts produced under the impulse of the tenth-century emperor Constantine VII Porphyrogenetos. Biographical entries were collected from a ninth-century reworking of the *Onomatologos* ('Inventory of famous learned people') of the sixth-century scholar Hesychius of Miletus. Overall, the *Suda* represents one of the high points of Byzantine scholarship and lexicography. An annotated English translation of the Greek text as edited by Adler is available online as *Suda On Line*.

The last general lexicon created during the Byzantine age is the one falsely attributed to the twelfth-century historian and jurist Joannes Zonaras.²² It is also known as *Lexicon Tittmannianum* from the name of its first and only editor, the German theologian and philosopher Johann August Heinrich Tittmann, who published the Greek text with a rich apparatus of notes in 1808. This edition is of some importance because of the methodology used by

²⁰ Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 40–1 with n. 44.

²¹ Adler, 'Suidas', 679–80; Daly, *Contributions to a History of Alphabetization*, 68; Alpers, 'Review of Daly, *Contributions*', 26.

²² *Iohannis Zonarae lexicon* (1808) is still the standard edition; see also Alpers, "'Zonarae' Lexicon"; Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 42–3; Alpers, *Das attizistische Lexikon des Oros*, 11–55; Alpers, 'Griechische Lexikographie', 29–30; Dickey, *Ancient Greek Scholarship*, 102.

Tittmann: it is quite modern and, at some points, anticipates some editorial principles now considered essential, such as the effort to edit the text of the lexicon on the basis of the manuscripts, avoiding the correction of mistakes which the author inherited from his sources. The edition is based upon three manuscripts. Nevertheless, it is now known that more than 120 codices of this work survive, making it the most widespread and best-attested Greek and Byzantine lexicon. Such an abundant manuscript tradition attests to the appreciation this work enjoyed over centuries, from its composition to the sixteenth century. A fresh study of the whole manuscript tradition and, consequently, a new edition of the text are still *desiderata*. As Alpers conclusively demonstrated, this huge lexicon dates back to the first half of the thirteenth century. The name of the compiler – or of one of the compilers – was probably Nikephoros.²³ In the manuscripts, the lexicon bears a general title which means ‘Collection of words and expressions from different books of the Old and New Testament as well as from profane ones’: as Herbert Hunger remarks, ‘the title points to the universality of the lexicon’.²⁴ Its core content is indeed composed of lexical and etymological explanations of the vocabulary of Christian and pagan authors. The lexicon’s structure shows a double, concurrent arrangement: first, the entries are alphabetized by the first two letters and, secondly, they are distributed according to grammatical categories, namely masculine nouns, feminine nouns, neuter nouns, verbs, and adverbs, in that order. The sources are heterogeneous, and include, as well as the *Suda*, the etymologica (both *Etymologicum genuinum* and *Etymologicum Symeonis*), the *Ethnica* by Stephanus of Byzantium, and Atticist lexica (for all of which see below). Grammatical, philosophical, and theological terms taken from other scholarly and literary sources are also explained in detail.

Etymologica

From the fifth century onwards, there is evidence for a well-defined subgroup within the Greek lexicographic production labelled and self-defined as ‘etymologica’.²⁵ These lexica are quite extensive works containing

²³ On this basis, Alpers cautiously suggested the theologian, philosopher, and scholar Nikephoros Blemmydes as a possible candidate for authorship.

²⁴ Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 42: ‘Er [der Titel] weist mit Altem Testament, Neuem Testament und profaner Literatur auf die Universalität des Lexikons hin.’

²⁵ Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 45–8; Alpers, ‘Griechische Lexikographie’, 27–32; Valente, ‘Etymological dictionaries’.

grammatical, linguistic, and literary material. The core idea at the basis of their compilation is the explanation of the origin and of the true (in the sense of correct) meaning of a given word. The tradition of etymology which they represent is not directly connected to the modern linguistic study of the history of words: it was rather 'a tool for thinking', as Ineke Sluiter aptly described it.²⁶ This practice can be discovered as early as the Homeric poetry. It can in particular be seen in classical Athens, especially within the sophistic movement. One of the best examples of how it was applied in philosophy is the dialogue *Cratylus* by Plato and, notably, many Platonic etymologies are still traceable in Byzantine etymologica.²⁷ Later on, the Stoics extended the studies in this field. Etymology was thus a linguistic and philosophical practice which was employed for discovering the reasons why a given word had a given form and a given meaning. This could be achieved by analysing its composition and by discovering which terms and concepts had been merged into a given word. In Byzantine etymologica, the explanation of a given lemma follows this pattern; quotations from literary and scholarly works were then used to support the respective explanations.

The first lexicon containing the term 'etymology' or 'etymologies' in its title was composed during late antiquity by the grammarian Orion of Thebes.²⁸ There is still no modern and reliable edition of it; its textual tradition is divided into three branches showing different degrees of alteration and is still to be clarified in its entirety. This lexicon is arranged by first letter, and the entries strictly follow the order of the philosophical, lexicographic, medical, and grammatical sources that Orion used in compiling the text. The importance of this lexicon lies not only in the fact that it offered the model for future Byzantine etymologica, but also because much of its content was copied word for word into the new compilations.

The earliest of these, possibly dating from the eighth or ninth century, is the so-called *Etymologicum parvum*.²⁹ It is a short lexicon preserved in the last leaves of a single tenth-century codex.³⁰ Because of the loss of at least a quire of this manuscript, the surviving version of the lexicon ends with the letter

²⁶ Sluiter, 'Ancient etymology'.

²⁷ Valente, 'Porphyrios und der "Kratylos"', and Valente, 'From Plato to the Byzantine Etymologica'.

²⁸ Orion, *Etymologicon* (1820), is an edition; see also Dickey, *Ancient Greek Scholarship*, 100; Reitzenstein, 'Etymologika', 810–11; C. Theodoridis in Philoxenos, *Fragmente*, 15–41; Valente, 'Etymological dictionaries', 575.

²⁹ *Etymologicum parvum* (1973) is the modern edition; see also Reitzenstein, 'Etymologika', 812; *Epimerismi Homerici*, II.41.

³⁰ Florence, Biblioteca Medicea Laurenziana, Codex Sancti Marci 304, which also includes a text of the *Etymologicum genuinum* (see note 33 below).

omicron. The entries show an arrangement by first letter which may reflect the order in which the compiler or compilers excerpted the different sources.

The most significant and influential etymologicum is the *Etymologicum genuinum*, mentioned above, which was composed around the middle of the ninth century, probably in Constantinople.³¹ It received its Latin name at the end of the nineteenth century, when Richard Reitzenstein published his seminal studies on Greek and Byzantine etymologica in 1897. The transmitted title of the lexicon in Greek means 'Great etymologicum' – in Latin translation, *Etymologicum magnum* – but that name had already been applied to a later etymologicum, printed in 1499, which will be discussed below. As Alpers has recently suggested, the learned circle of Leo the Mathematician and of Cometas within the context of the school in the Magnaura at Constantinople may be responsible for the compilation of the *Etymologicum genuinum*.³² Only two manuscript copies of this huge work survive, both dating to the end of the tenth century; one of them also includes the sole copy of the *Etymologicum parvum*.³³ The manuscripts are almost complete, but they transmit an abridged text, and some quires are now lost or severely damaged. A fuller version of the text was available to the compilers of later etymologica and to other Byzantine scholars, and these works are therefore of capital relevance for the reconstruction of the original text. As Filippomaria Pontani has put it, the *Etymologicum genuinum* offers

an impressive alphabetical list of terms derived from prose and poetry of all centuries, analysed in their etymology, orthography and meaning, with the help of commentaries to ancient texts, and often with references to non-grammatical sources that may help clarify proper names or historical realities.³⁴

Besides grammatical and lexicographic texts – among others, Orion and the extended *Synagoge* – important sources were some commented editions of poets such as Homer, Hesiod, Aeschylus, Euripides, and Aristophanes. The

³¹ Partial editions include *Etymologicum magnum genuinum* (1976–92), covering the range alpha to beta, and Alpers, *Bericht*, 25–57, a specimen edition of the range comprising words beginning with lambda (others are enumerated in Dickey, *Ancient Greek Scholarship*, 92); see also Reitzenstein, *Geschichte der griechischen Etymologika*, 1–69, and Reitzenstein, 'Etymologika', 812–14; Alpers, 'Difficult problems'.

³² Alpers, 'Byzantinische Enzyklopädie', 260f., 267f.

³³ Rome, Biblioteca Apostolica Vaticana, ms gr. 1818, and Florence, Biblioteca Medicea Laurenziana, Codex Sancti Marci 304 (in which the copy of the *Etymologicum genuinum* was completed on 13 May 994, according to the colophon); see Alpers, 'Marginalien', 523–30, and Alpers, 'Difficult problems', 294–5.

³⁴ Pontani, 'Scholarship in the Byzantine empire', 338.

compilers often had access to better – or different – manuscript sources than those that have been preserved until today. Leaving aside its importance for our understanding of the scholarly context of ninth-century Byzantium, the lexicon is also an invaluable repository of information and quotations from lost works. Unfortunately, only some parts of the text of the *Etymologicum genuinum* have so far been edited. The fascinating and complex story of the preliminary work towards a first critical edition of the lexicon begins at the end of the nineteenth century with Reitzenstein: his materials were inherited by Ada Adler (see above) and are now being completed by Klaus Alpers.³⁵

About a century after the composition of the *Etymologicum genuinum*, that is, during the second half of the tenth century, another pivotal lexicon, the so-called *Etymologicum Gudianum*, was created.³⁶ The name refers to the seventeenth-century German scholar Marquard Gude, the owner of one manuscript, produced in 1293 in Terra d'Otranto in Apulia (Italy) and now at Wolfenbüttel, which formed the basis of the first edition of the text, published in 1818.³⁷ Its editor, Friedrich Wilhelm Sturz, worked from a modern transcript of this manuscript rather than collating it in person; moreover, the manuscript itself is by no means the most suitable witness for the text because of its many additions from other sources and the abbreviation of numerous entries. Luckily, the original working copy of the lexicon is still preserved in a manuscript of the second half of the tenth century, now in the Vatican library, the importance of which was first realized by Reitzenstein.³⁸ This manuscript is one of the very few cases where a working copy of a Byzantine lexicon has survived to this day: different scribes and scholars belonging to the same circle enhanced the main text by adding further materials in the margins, gathering them from different sources, which, interestingly, they sometimes identified with sigla – a rather uncommon practice in the Byzantine world.

It is generally assumed that the Vatican manuscript of the *Etymologicum Gudianum* was produced in the region of Otranto, as was the Wolfenbüttel

³⁵ See Alpers, *Bericht*, and Alpers, 'Difficult problems'.

³⁶ *Etymologicum Gudianum* (1909–20) is the best edition of the range it covers, from alpha to *zeiai*; for the remainder of the alphabet, *Etymologicum graecae linguae Gudianum* (1818) must be used, with caution; see also Reitzenstein, *Geschichte der griechischen Etymologika*, 70–155, and Reitzenstein, 'Etymologika', 814–15; Cellerini, *Introduzione all'Etymologicum Gudianum*; Alpers, 'Difficult problems'.

³⁷ Wolfenbüttel, Herzog August Bibliothek, ms Gudianus Graecus 29/30; see Reitzenstein, *Geschichte der griechischen Etymologika*, 87 n. 1; Harlfinger, *Griechische Handschriften*, 35–7 with pl. 11; Sciarra, 'Note sul codice Vat. Barb. gr. 70', 388 with nn. 108–11.

³⁸ Rome, Biblioteca Apostolica Vaticana, ms Barberinianus gr. 70; see Maleci, *Il codice Barberinianus Graecus 70*; Alpers, 'Difficult problems', 295–6.

manuscript, but it has recently been suggested that it might instead have been written in the eastern territories of the Byzantine empire, possibly in Constantinople.³⁹ Be that as it may, it was kept in southern Italy from the eleventh century onwards, and was used – directly or indirectly – there as the model for further copies. In some of them, the text has been extended through addition from other lexicographic traditions, such as from the lexicon of Cyril. These copies are still of great importance not only for the study of Greek culture in Terra d'Otranto from the eleventh to the fifteenth centuries, but also for the *constitutio textus* of the lexicon itself. In fact, since some quires in the Vatican manuscript are now missing, the copies allow the reconstruction of lost parts of the text. Many of the sources used by the compilers of the *Etymologicum Gudianum* were employed for the *Etymologicum genuinum* – indeed, the latter was used for some expansions of the text of the former – but they were used independently and exhibit some differences.

The picture becomes even more nebulous for the later Byzantine etymologica. Two compilations dating back to the eleventh/twelfth century are particularly relevant: the *Etymologicum Symeonis* and the *Etymologicum magnum*. Independently of each other, their compilers used both the *Etymologicum genuinum* and the *Etymologicum Gudianum*, and of course supplementary sources can be detected as well. There is still no complete modern critical edition of either of them: this complicates a proper study of their texts and their relationship to the sources. Probably in the first half of the twelfth century, a grammarian by the name of Symeon composed an etymologicum that is now preserved in five manuscripts dating to the thirteenth and fourteenth centuries.⁴⁰ Two of them preserve a version closer to the original, the others an extended version, the so-called *Magna grammatica*. The other, later work, the *Etymologicum magnum* already mentioned, was composed in the eleventh or twelfth century.⁴¹ According to Reitzenstein, the author was a competent scholar of impressive diligence.⁴² He combined different sources in a creative way in order to give birth to a new lexicon fitting contemporary needs. He was also able to correct the text of his sources

³⁹ Ronconi, 'Quelle grammaire à Byzance?', 86.

⁴⁰ An edition of the range alpha to beta is presented in *Etymologicum magnum genuinum* (1976–92), and *Etymologicum Symeonis* (2013) adds the range gamma to epsilon; see also Reitzenstein, *Geschichte der griechischen Etymologika*, 254–86, and Reitzenstein, 'Etymologika', 816–17.

⁴¹ The last complete edition was *Etymologicum magnum* (1848); see also Reitzenstein, *Geschichte der griechischen Etymologika*, 212–53, and Reitzenstein, 'Etymologika', 815–16.

⁴² Reitzenstein, *Geschichte der griechischen Etymologika*, 253: 'es war ein Gelehrter von achtungsgebietendem Fleiß und – soweit man bei diesem Stoff davon reden kann – nicht ohne Urteil'.

in some passages through felicitous conjectures. The *Etymologicum magnum* was the first Byzantine etymologicum to be printed: the *editio princeps* appeared in Venice in 1499, from the press of Zacharias Kallierges, a famous scribe and printer from Crete, with a preface by Marcus Musurus.⁴³ The last complete edition was published in 1848 by the English scholar Thomas Gaisford. The text is still an invaluable, sometimes unique, repository of literary fragments as well as of grammatical and linguistic doctrines.

Atticist Lexicography

For many centuries, Atticism had been the most important rhetorical and linguistic trend in Greek and Byzantine literature. The imitation of the language and style of literary authors active in Athens during the fifth and fourth centuries BC was used as a mark of distinction in the higher spheres of Byzantine culture and administration. It had its origins in the Roman empire of the first century AD, flourished during the second century as an element of the so-called Second Sophistic era (see Chapter 5), and lasted until the end of the Byzantine world.⁴⁴ Nigel Wilson describes this trend as follows:

it became the accepted fashion to neglect the spoken language, and instead to follow the style, syntax, and vocabulary of authors who had lived five hundred years earlier . . . The new fashion of Atticism . . . instead of being the passing vogue of a single generation . . . took root to such an extent that it lasted until the end of the Byzantine period.⁴⁵

Imitating the literary Attic language implied at the same time avoiding contemporary language. The re-creation of the style and of the vocabulary of Attic literary models for new compositions had become a requirement for the claim to belong to the cultivated upper class: this resulted in the creation of an artificial Atticist sociolect. Furthermore, the phenomenon of Atticism contributed to an increase in the dichotomy between two layers – or literary sociolects – within the Greek language: a higher and a lower register. It is impossible to understand the evolution of Greek literary

⁴³ *Etymologicum magnum graecum* (1499); see Alpers, 'Difficult problems', 293f. (for the point that although Musurus wrote the preface, he did not edit the text, see 293 n. 3, with further bibliography).

⁴⁴ Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 4–5, 7; Wilson, *Scholars of Byzantium*, 4–8; Rollo, "'Greco medievale' e 'greco bizantino'"; Matthaios, 'Greek scholarship', 290–6; Pagani, 'Language correctness', 828–30.

⁴⁵ Wilson, *Scholars of Byzantium*, 5.

language up to today without a comprehensive consideration of this phenomenon.

For the practical needs of rhetoric and literary composition, Atticist lexica and reference works had constantly been produced and transmitted in manuscript form throughout the centuries. Such a rich and multifaceted production included a lively exchange between different exponents of this cultural movement and strong polemical attacks against each other. Starting from the second century, Atticist lexicographic collections aimed at educating their users on vocabulary and expressions that were considered worthy of being employed in contemporary literary production (see Chapter 5). Atticist lexicographers considered adequate only such words, syntactic constructions, and expressions as they could trace back to the linguistic usage of Attic classical authors. The selection of the canon of authors to be taken as a model was perhaps the main issue of contention between grammarians. Because of their richness in citations, Atticist lexica are of paramount importance to the reconstruction of the works of classical authors, especially for texts no longer preserved as a whole. Moreover, they offer a primary heuristic tool to analyse the evolution of literary – and in part current – language.

During the Byzantine age, two different trends in Atticist lexicography can be observed: first, the copying of Atticist lexica of the past into new manuscripts, with different degrees of fidelity; secondly, the production of new Atticist lexica on the basis of the recombination of older ones with new (literary) sources. The intent of both trends was prescriptive, making a clear distinction between Attic forms to be used in the contemporary literary and rhetorical production and those non-Attic ones which should be avoided. Atticist entries can be discovered, to varying extents, in almost every Byzantine lexicon. Many Byzantine manuscripts preserve copies of Atticist lexica which had originally been compiled from the second century up to the fifth: the lexica by the grammarians Phrynichus, Pollux, and Moeris are only a few examples of this rich production. Eventually, the Palaeologan age – especially the period from the end of the thirteenth century up to the fourteenth – saw an important revival of Atticism within the context of a new age of classicism.

The *Ecloga* by Thomas Magistros is certainly one of the most seminal lexica dating to this period.⁴⁶ The author – ‘a perfect example of a “gentleman

⁴⁶ Thomas Magistros, *Ecloga* (1832), is the last complete edition; see also Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 44; Gaul, ‘Twitching shroud’, 296–328; Gaul, ‘Moschopoulos’, 184–90; and Gaul, *Thomas Magistros*, 141–4, 401–2.

scholar”, to use a definition suggested by Pontani – lived in Thessalonica.⁴⁷ In compiling his lexicon, Thomas combined different Atticist lexa dating from the second and third centuries with materials derived from the grammatical and lexicographical works of his contemporary Manuel Moschopoulos.⁴⁸ Entries coming from the same source are grouped together. The lexicon is alphabetically arranged by first letter. Thomas improved the text of his lexicon during his life, as the manuscript tradition attests (unfortunately, this process cannot be appreciated in the last complete edition of the lexicon, published in 1832). Pontani summarized the main characteristic of the *Ecloga* as follows:

his bulky work, displaying a clear prescriptive goal in orthographical and morphological issues, derives its examples partly from a first-hand reading of some of the leading Attic writers (from Aristophanes to Thucydides), and partly from pagan authors of the imperial age (and from Gregory of Nazianzus); its greatest innovation on the background of contemporary linguistic tools consists in the massive use of the Atticist lexa of the imperial age, from Phrynichus to Moeris.⁴⁹

Conceived in Thessalonica, the work reached Constantinople some decades later, where the text received new expansions within the scholarly circle of Moschopoulos.

Contemporary to Thomas Magistros and Manuel Moschopoulos were other important scholars and teachers. One was Andreas Lopadiotes, the author of the so-called *Lexicon Vindobonense*. The *editor princeps*, August Nauck, chose this name because the sole manuscript known at that time was preserved in the National Library of Vienna (*Vindobona* is the Latin name for Vienna).⁵⁰ The work is a collection of lexicographic excerpts which preserve some otherwise unattested fragments of classical Greek literature.⁵¹ Another scholar of the same period was George Phrankopoulos, whose lexicon is preserved in a single manuscript copied in the year 1310.⁵² The main text turns out to be a re-elaboration of the

⁴⁷ Pontani, ‘Scholarship in the Byzantine empire’, 422. On the label ‘gentlemen scholars’, see Browning, ‘Teachers’, 105; Gaul, ‘Moschopoulos’, 184.

⁴⁸ On Moschopoulos, see Gaul, ‘Moschopoulos’, 166–77, 185.

⁴⁹ Pontani, ‘Scholarship in the Byzantine empire’, 422.

⁵⁰ Vienna, Österreichische Nationalbibliothek, ms phil. gr. 169.

⁵¹ The *editio princeps* is *Lexicon Vindobonense* (1867); a new critical edition is in preparation by Augusto Guida. See Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 43; Gaul, ‘Moschopoulos’, 182–4; Pontani, ‘Scholarship in the Byzantine empire’, 419.

⁵² Rome, Biblioteca Apostolica Vaticana, ms gr. 7; see Ucciardello, ‘Nuovi frammenti’; Gaul, ‘Moschopoulos’, 178–81.

Etymologicum Magnum and of the lexicon attributed to Zonaras. In the margins and on supplementary sheets, the scribe – possibly Phrankopoulos himself, as Niels Gaul argues – added a rich apparatus of various notes.⁵³ The manuscript contains only the first five letters of the alphabet, from alpha to epsilon, and may thus represent a working copy of a lexicon that was never completed.

Lexica of Other Kinds

Among the lexicographic production, some works chiefly dealing with specific areas of expertise can be isolated. Only one geographic lexicon survives from the ancient and Byzantine Greek world: the *Ethnica*, written by the grammarian Stephanus of Byzantium.⁵⁴ Shortly after the composition, the original fifty or more books were abridged by one Hermolaus, who is otherwise unknown. This shortened text is still extant in some fifteenth- and sixteenth-century manuscripts. Traces of a more complete version are preserved in the form of quotations in other Byzantine authors and in a quire of a manuscript of the tenth/eleventh century.⁵⁵ A fuller version than the transmitted one was also known to other Byzantines, such as Constantine Porphyrogenetos and the twelfth-century commentator Eustathius. The lexicon is fully alphabetized: the entries concern toponyms, explained mainly from a grammatical, linguistic, and geographical point of view. The text is also rich in quotations from literary authors. Mythological and historical accounts are usually provided as well. Since many of the reference texts that Stephanus could consult are no longer extant, the lexicon represents for us a primary – and sometime the sole – source of information on literary, historical, and grammatical data.

Particularly rich was the production of synonymica, that is to say lexica in which synonyms are collected in order to ascertain their different meanings.⁵⁶ The larger part of this production is still unedited. Most such lexica produced in the Byzantine age – the extant works attributed to Ammonius and Herennius are the best-known examples – are descended from a lost work by an ancient author, Herennius Philo of Byblos, which probably had a title

⁵³ Gaul, 'Moschopoulos', 182–4.

⁵⁴ The modern edition is Stephanus, *Ethnica* (2006–17); see also Valente, 'Dictionaries of onomastics', 484; Pontani, 'Scholarship in the Byzantine empire', 314–15.

⁵⁵ Paris, Bibliothèque Nationale, ms Coislin 228.

⁵⁶ Hunger, *Hochsprachliche profane Literatur der Byzantiner*, 48–9; Dickey, *Ancient Greek Scholarship*, 94–6; Valente, 'Synonymica'.

meaning ‘On similar and different words’ (see Chapter 5).⁵⁷ The grammarian Symeon, author of the above-mentioned *Etymologicum Symeonis*, also composed a synonymic lexicon, also unpublished.⁵⁸ The typical syntax of a lemma in a synonymic lexicon is: ‘x and y (and z) are different’. Then, the explanation accounts for meanings and usages on the basis of grammatical and literary sources. Single entries with this peculiar structure can also be found in other Byzantine lexica.

Lexica to be used in the first educational stages to learn grammar and to study texts by classical and Christian authors are widespread. Among many others, the *Epimerismi Homerici* offer interesting insights into the teaching praxis of the Byzantines. The grammatical genre of the *epimerismi* is typical of a scholastic context, in which a poetical text is analysed word for word and commented on primarily from a grammatical point of view.⁵⁹ Quotations from classical and Christian authors enrich the explanations of grammatical doctrines. The surviving collection of *epimerismi* to the Homeric *Iliad* offers the best example of this teaching tool. Originally, the entries followed the wording of the poetic verses. Later on, they were rearranged in alphabetical order in order to create an autonomous lexicon. The authorship is unknown, but the last editor, Andrew Dyck, suggested George Choeroboscus, a grammarian who probably lived in the second half of the eighth century. He was the author of a similar work devoted to the Psalms, the *Epimerismi in Psalmos*, described concisely by Pontani as ‘a pedantic exercise of word-for-word parsing of the Septuagint’s text with the aim of singling out the morphological and grammatical category of each term, independently of its syntactic function’.⁶⁰

‘Overtime’ (or Aftermath)

During the European Renaissance, many Byzantine Greek lexica saw printed editions: I have remarked above on those of the *Etymologicum magnum* and the *Suda* (both 1499) and Hesychius (1514). Moreover, as Alpers observes, the

⁵⁷ Herennius, *De diversis verborum significationibus* (1988), is an edition of what has been transmitted under the name of Herennius; see also K. Nickau in Ammonius, *De adfinium vocabulorum differentia*, esp. xxviii–lxiii.

⁵⁸ Reitzenstein, *Geschichte der griechischen Etymologika*, 256; Nickau, ‘Neues zur Überlieferung’, and Nickau, ‘Altes und Neues zur Überlieferung’.

⁵⁹ *Epimerismi Homerici* (esp. I.3–5). Robins, *Byzantine Grammarians*, 125–7, 130–8; Pontani, ‘Scholarship in the Byzantine empire’, 319.

⁶⁰ Choeroboscus, *Epimerismi in Psalmos* (1842), is an edition; the description is in Pontani, ‘Scholarship in the Byzantine empire’, 319.

first new Greek dictionary produced in Italy during the sixteenth century was composed in the style of a Byzantine work.⁶¹ Its author was the Italian Benedictine monk Guarino Favorino.⁶² Around 1480, he went to Florence, where he became first a pupil of Poliziano and later the tutor of Giovanni di Lorenzo de' Medici, the future pope Leo X. From the 1510s, Guarino's service to the latter took him to Rome, where he acted as librarian of the manuscript collection of the Medici family. During this period, he also tended to the composition of his great Greek lexicon, the *Magnum ac perutile dictionarium*, printed in Rome in 1523 by Zacharias Kallierges, who had printed the *editio princeps* of the *Etymologicum magnum* a quarter of a century earlier. In 1538, a new edition, revised by Joachim Camerarius, appeared in Basel. The lexicon is a clever assembly of entries extracted from different sources: among others, the printed editions of Hesychius and of the *Etymologicum magnum* as well as some lexica that had not yet been published, such as the lexicon attributed to Zonaras, the *Lexicon Vindobonense*, the *Etymologicum Gudianum*, and Thomas Magistros' *Ecloga*. With few exceptions, the work is fully alphabetized. The lexicon enjoyed significant popularity and was the first attempt to offer a reference work of the Greek language, based on a wide variety of Byzantine scholarship, to a larger European readership (for Greek–Latin dictionaries of the sixteenth century, see Chapter 14).

Conclusions

This concludes a short journey over the thousand years of history of Byzantine Greek lexicography. To sum up, the main feature of this production is probably the extent to which the Byzantine grammarians and lexicographers accumulated older learned materials they could find in their sources. Their main purpose was not originality, but the preservation and diffusion of the knowledge of the past. In doing so, they both copied old lexica into new manuscripts and composed new ones combining different pre-existing sources. Byzantine studies on language – be it literary or current – were chiefly based upon available lexica. The two main aims of Byzantine lexica were the understanding of the language of literary Christian and pagan authors and a personal rhetorical training for contemporary literary composition. Concerning the new lexicographic creations, their originality resides first of all in an updated recombination of the materials of the sources. Such

⁶¹ Alpers, *Das attizistische Lexikon des Oros*, 42–7, and Alpers, 'Griechische Lexikographie', 30.

⁶² Alpers, *Das attizistische Lexikon des Oros*, 42 n. 3; Ceresa, 'Favorino'.

instruments were fitted to the needs of the respective users. Where this was not feasible, new compilations were created to meet changed expectations. Arriving at a proper understanding of Byzantine lexicography is a recent development within Byzantine Greek studies, developed from new investigations into the materiality and textuality of many of the lexica. This journey continues.

Medieval Latin Christendom

JOHN CONSIDINE

The western Roman empire left a double cultural legacy: Christianity, of a kind which acknowledged the authority of the bishop of Rome, and Latin, the language of Roman Christianity. This chapter tells the story of lexicography in the lands where Roman Christianity was practised and the Latin language was read. These lands correspond roughly with modern western and central Europe, but the concept 'Europe' was not in general use until the very end of this period, and need not distract us here. The Islamic and Orthodox neighbours of Latin Christendom had their own lexicographical traditions, which are treated in Chapters 8, 11, and 12. Within Latin Christendom, Jews contributed to the lexicography of at least three vernacular languages (French, Italian, and Czech), and these contributions are documented here, but the learned Jewish tradition of Hebrew lexicography is treated separately in Chapter 9. The pagans on the edges of Christendom were illiterate, and became literate only as they became Christian.

The time of the story to be told here is bounded by the fall of the western Roman empire and the end of the fifteenth century; the story up to the fall of the western empire is told in Chapter 5, and the story after the end of the fifteenth century is told in Chapter 14 and later chapters. Rather than moving century by century through our subject matter, we will begin with an account of monolingual Latin lexicography, with a brief passage on wordlists of Latin together with the learned languages Greek and Hebrew, before turning to the lexicography of the vernacular languages of medieval Latin Christendom. These will be discussed in four groups: the Celtic languages; the Germanic languages; the Romance languages; and others including Basque and the languages of Roman Slavdom.

All the extant manuscripts of all the wordlists discussed in this chapter were written on parchment or paper, and all but the very shortest were transmitted as codices: they were books or parts of books, not scrolls, or bundles of loose sheets. Until the very end of the period, printing from

movable type was unknown in Latin Christendom. Manuscript production was, however, undertaken on a large scale and, by one estimate, something like half a million medieval manuscripts survive from this cultural area, excluding letters and the shortest records of property ownership.¹ One way of indicating the relative importance of different kinds of wordlist is to give figures for the numbers of manuscripts in which they have been preserved, and this has been done when the figures were available.

As we shall see, there are more than a hundred extant manuscripts for each of the most widely circulated dictionaries (the *Elementarium* of Papias, the *Expositiones* of Brito, the *Derivationes* of Hugutio, the *Catholicon* of Giovanni Balbi, the *Abstractum* glossary, and the *Vocabularius ex quo*), and many others have of course been lost. So, there were undoubtedly thousands of dictionary manuscripts in Latin Christendom by the end of the fifteenth century, not to mention the thousands of copies of dictionaries which were printed in the closing decades of the century. These thousands of dictionaries were in the hands of thousands of dictionary users, a reminder that we should not underestimate medieval literacy. It has been argued that in England, by the beginning of the fourteenth century, a minimal reading knowledge of Latin 'was common among the gentry and may not have been rare among peasants'.² There is no reason to suppose that this situation was exceptional: although literacy rates were probably lower in the seventh century than in the fifteenth, and were surely lower in wild country than in urban centres, literacy in Latin was part of many medieval people's lives. Even if they had not all used dictionaries themselves, dictionary use was an important part of the learned culture of which they had tasted.

Latin: Early Glossaries

A number of ancient Latin lexicographical texts continued to be read and copied in and beyond late antiquity. So, for instance, a collection of explanations of words relevant to administration and the law, perhaps made in the early fifth century, is extant under the title *Expositio notarum* in a manuscript of around AD 1000.³ The fourth-century commentary on Vergil by Servius, which naturally includes lexicological material in the form of remarks on Vergil's vocabulary, was expanded in or around the seventh century, and was

¹ Buringh, *Medieval Manuscript Production*, 100.

² Clanchy, *From Memory to Written Record*, 246.

³ Dionisotti, 'Nature and transmission', 228–41.

copied throughout the medieval period.⁴ The extant manuscripts of the ancient Greek–Latin schoolbook called *Hermeneumata*, copies of which might include an alphabetical wordlist of up to 3,000 entries (a lost manuscript appears to have had a wordlist of 11,309 entries), date from the eighth century onwards.⁵ The abridgement by Paul the Deacon of the *De verborum significatu* of Pompeius Festus likewise belongs to the eighth century.

But the real energy of Latin lexicography from late antiquity to about the eleventh century went into the making of glossaries.⁶ These are a kind of wordlist for which there is no evidence from the ancient world but increasingly ample evidence from the fifth century onwards. They were typically connected to the annotation of texts. For instance, a reader of a particular manuscript of Vergil's *Georgics* came upon the instruction to give nourishing food to swift dogs and to the *acrem Molossum* ('alert mastiff'). *Molossum* is not a common word, and so the reader made a marginal or interlinear note on the manuscript, writing in pen and ink or scratching with a dry point, that it refers to a kind of dog, *canis*. Such a note is called a gloss, and consists of a *lemma*, the word to be interpreted, and an *interpretamentum*, the word which interprets it. (The word *gloss* is also used of other kinds of text such as the brief annotations on legal works and on the Bible which were collected and transmitted under titles such as *Glossa ordinaria*.⁷ These were not primarily lexicographical.)

Once a gloss on *Molossum* had been written, in a form such as 'Molossum[:]
canis', it was then copied with other notes on the same text into a list of *glossae collectae* which could be used by readers of the *Georgics*. The original list may have presented words in the order in which they occurred in the *Georgics*, before being recopied with material from other lists in a single master-list. This master-list, which might preserve its constituent batches of *glossae collectae* intact or might merge them in a single alphabetical or topical sequence, was a glossary. Alphabetical order was widely used, although we shall see exceptions: having been chosen, it demanded little reflection on the part of the compiler or the reader of a glossary, and the codex form lent itself well to multiple alphabetical searches (imagine searching for *velox*, *Sparta*,

⁴ See the numerous fifteenth-century manuscripts catalogued in Murgia, *Prolegomena*, 45–71.

⁵ *Colloquia of the Hermeneumata*, I.16–24 and II.145–8; Dionisotti, 'From Ausonius' school-days?', 87.

⁶ Overview in McKitterick, 'Glossaries and other innovations', 40–76. The major editions of the primary texts are *Corpus glossariorum Latinorum* and *Glossaria Latina*, for both of which see Dionisotti, 'Nature and transmission', 207–25.

⁷ See Gilles, 'Glosses, legal', and Lobrichon, 'Glosses on the Bible, Latin'.

catulus, *acer*, and *Molossus* one after another in the pages of an alphabetized dictionary in book form, and then imagine performing the same search in the multiple scrolls of the same dictionary in scroll form). If alphabetization was chosen, it might be by the first letter only (A- order), or by the first two, three, or four letters (AB- order, ABC- order, ABCD- order). The explanation of *Molossus* as *canis* would, once this gloss had become part of a glossary, be detached from its original context. So, around AD 700, the Anglo-Saxon bishop Aldhelm wrote in Latin of the *Molossus* returning to its vomit as if *Molossus* were a synonym of *canis* rather the name of one special breed, and he must have done this because he had found the word in a glossary.⁸

Early glossaries did not originate exclusively in the glossing of texts, for they might draw on ancient lexical collections such as the *Hermeneumata* wordlists and the *Expositio notarum*.⁹ However, the relationship between reading and the making of glossaries is clear enough, and batches of entries in early glossaries can sometimes be traced to particular literary, patristic, or biblical texts. So, for example, an A- ordered glossary fragment begins with the headword *arma*; this is the first word of Vergil's *Aeneid*, an important school text, and the first entries in the glossary are indeed glosses on the works of Vergil.¹⁰ Another glossary manuscript begins with a statement of its connection with the schoolroom: a full-page picture of a teacher.¹¹

The first extant Latin glossaries belong to the fifth or sixth century, including one associated with the name of Placidus, which was the source for a comic poem of the sixth century.¹² Placidus may be the earliest named lexicographer of Latin Christendom, and is in fact rather exceptional in being named, for many early glossaries are anonymous. They are often named with reference to their first lemmata, for instance *Abba–Ababus*, of which the extant manuscript – mentioned in the previous paragraph on account of its frontispiece – was written in northern Italy, in the late seventh century or the eighth, making it one of the oldest extant glossary manuscripts.¹³ The early glossaries are numerous, and their relationships are intricate.

⁸ The example is from Lindsay, 'Note on the use of glossaries', 16–18; cf. *Corpus glossariorum Latinorum*, VI.708.

⁹ Dionisotti, 'Nature and transmission', 247; cf. McKitterick, 'Glossaries and other innovations', 41.

¹⁰ Dionisotti, 'Nature and transmission', 225, citing *Glossaria Latina* II.1.

¹¹ St Gallen, Stiftsbibliothek, MS 912, p. 3; McKitterick, 'Glossaries and other innovations', 52, reports another example.

¹² Versions of the Placidus glossary are in *Corpus glossariorum Latinorum*, V.3–158 and *Glossaria Latina*, IV.12–70; for the man, see Kaster, *Guardians of Language*, 341–2; for the poem, see Lendinara, 'Contextualized lexicography', 133.

¹³ Lowe, *Codices Latini Antiquiores*, VII.36 (no. 967a).

The greatest early collection of glosses was the so-called *Liber glossarum*, also known as the *Glossarium Ansileubi*, which has its origins in seventh-century Visigothic Spain.¹⁴ A remarkable early manuscript is written in three columns, elegantly laid out on pages 540 mm tall.¹⁵ The *Liber glossarum* draws on Isidore and on earlier glossaries to register some 27,000 Latin lemmata, almost perfectly alphabetized in a single sequence.¹⁶ This high degree of alphabetization, sustained over such a long list, is most unusual for the period, and must have been costly in hours of labour. Perhaps it was achieved by using erasable drafting surfaces such as wax tablets, or perhaps the entries were originally written on separate sheets or slips of parchment which were then sorted into alphabetical order.¹⁷ If the latter is true, the alphabetization of the *Liber glossarum* is an extremely early example of this effective but expensive way of managing information. The difference between the *Liber glossarum* and a dictionary is that the former does not attempt to bring glosses on a given lemma together in unified entries: so, for instance, on the first page, a first entry for *abdican* ('they resign/disown/repudiate/expel') is followed by another entry for the very same form with a different gloss, and then by an entry for *abdicare*, the infinitive form of the same verb, and then by four entries with different interpretamenta for *abdicat*, the third-person singular. Abridgements of the *Liber glossarum* were made. One of these, the tenth-century *Glossarium Salomonis*, goes some way towards uniting its original's multiple glosses for a given lemma: there is, for instance, a single entry for *abdican*, followed by three entries for *abdicat* and one for *abdicare*.¹⁸ In this way, the *Glossarium Salomonis* is less like a collection of glosses than its original, and more like a modern dictionary.

Dictionary, however, is a tricky word to use in this context, for the slightly blurry distinction made in modern English between dictionaries, which provide information about words, and encyclopedias, which provide information about things, was not made explicitly in the medieval period. The *Etymologiae* compiled in the early seventh century by St Isidore of Seville,

¹⁴ *Liber glossarum* (2016) is an edition (lemmata and excerpts are presented in *Glossaria Latina*, vol. I); McKitterick, 'Glossaries and other innovations', 44–5, and the notice in Stammerjohann (ed.), *Lexicon Grammaticorum*, 905–6, are helpful short introductions; see also Ganz, 'The "Liber glossarum"'.
¹⁵ BN MS lat 11529 and 11530.

¹⁶ Entry count from Stammerjohann, *Lexicon grammaticorum*, 905.

¹⁷ The use of slips is suggested by McKitterick, 'Glossaries and other innovations', 45, and in *Glossaria Latina*, I.7; cf. the scepticism of Daly, *Contributions to a History of Alphabetization*, 86.

¹⁸ BSB, Clm 13002, fo. 8v. See also McGeachy, 'Glossarium Salomonis', 313. There is no printed edition of the *Glossarium Salomonis*.

which as its name suggests is full of lexical material, drawing, for instance, on Servius and on the glossary of Placidus, is now seen as an encyclopedia rather than a dictionary.¹⁹ Conversely, there is encyclopedic information in the *Liber glossarum*, some of it from the *Etymologiae*, and it has therefore been called an encyclopedia.²⁰ A twelfth-century manuscript of the *Glossarium Salomonis* which is followed by a text of the *Hermeneumata*, and shares the *mise-en-page* of this language-oriented text, had a series of beautiful drawings, notably of the human body as microcosm, prefixed to it very shortly after the text was written, and these present it to the reader as a book full of information about all the things in the world.²¹

Latin: Later Dictionaries

From the eleventh century onwards, a Latin dictionary tradition emerged.²² The works in this tradition tended to give a general account of the sense of a given lemma rather than focusing on its meaning in one particular passage, and the information which they provided tended to be primarily about the properties of words rather than the properties of things. The tradition begins with the *Elementarium* of the eleventh-century Lombard grammarian Papias.²³ This text had a long career: more than a hundred manuscripts are extant, fifteen of them as late as the fifteenth century, and there were four printed editions between 1476 and 1496; the first printed dictionary of Spanish is based on it.²⁴ The *Elementarium* drew heavily, as Papias pointed out in his interesting preface, on earlier lexicography – in fact on the *Liber glossarum* or one of its abridgements – being ‘a work which has, to be sure, already been elaborated for a long time by many others and has also been added to and amassed by me more recently as best I could’.²⁵ The title, more fully *Elementarium doctrine erudimentum* (‘Basic introduction to education’), pointed out that whereas the *Liber glossarum* and the *Glossarium Salomonis* had been

¹⁹ See Isidore, *Etymologies*, ed. Barney et al., 10–17, esp. 12.

²⁰ Ganz, ‘The “Liber glossarum”’, 127, 132; cf. McGeachy, ‘Glossarium Salomonis’, 310.

²¹ Cohen, ‘Making memories’, 135–42.

²² There is an overview in Weijers, ‘Lexicography in the Middle Ages’, and there are some good comments on individual texts in Hunt, *Teaching and Learning Latin*, 1.371–94; see also Stammerjohann, *Lexicon grammaticorum*, s.nn. Giovanni Balbi, Hugutio, Osbern, and Papias.

²³ Papias, *Elementarium* (1977–80), is an edition of the entries from A to azoni. For Papias’ identity, see Black, *Humanism and Education*, 49 n. 88.

²⁴ Bursill-Hall, *Census*, 315.

²⁵ Papias, edited in Daly and Daly, ‘Some techniques’, 230, ‘opus quidem a multis aliis iam pridem elaboratum, a me quoque nuper . . . prout potui adauctum et accumulatum’, translated *ibid.*, 232.

books for advanced users in centres of learning, the *Elementarium* was meant for less advanced readers in settings such as the cathedral schools of northern Italy, at one of which Papias himself had probably been educated; grammatical texts are therefore among its sources.²⁶ (Some of the extant manuscripts of the *Elementarium* were, however, used for more advanced scholarly work than Papias seems to have expected; for instance, a thirteenth-century scholar annotated his copy with ‘extensive citations of usage of rare words by ancient authors’.)²⁷ Although the *Liber glossarum* is the most important source of the *Elementarium*, Papias was not interested in preserving its full alphabetization: he usually kept the order of the entries which he took over from it, but he sometimes inserted other entries between them without strict regard even to ABC- order. So, the availability of the fully alphabetized *Liber glossarum* had not transformed the lexicographical techniques of Latin Christendom.

After Papias, the story of medieval Latin lexicography is dominated by larger dictionaries with named compilers rather than by small, anonymous glossaries like *Abba–Ababus*. Three of the largest will be discussed below, but there were many others, and some of these were widely read: for instance, a dictionary for reading the Latin Bible, the thirteenth-century *Expositiones vocabulorum Biblie* of Guillelmus Brito (Guillaume le Breton), circulated in some 130 manuscripts and was printed in one fifteenth-century edition.²⁸

In the mid twelfth century, the Benedictine monk Osbern Pinnock of Gloucester compiled the ‘huge, avant-garde, and influential’ dictionary called *Panormia* or *Liber derivationum*, which registers about 18,000 words, and is extant in 27 complete manuscripts.²⁹ It is less than a century younger than the work of Papias, to which it shows no indebtedness. One of its ‘avant-garde’ features is its use of ancient texts which were only coming to be rediscovered at the time of its compilation, for instance the comedies of Plautus. Another is its structure, in which each letter is represented by a section of *derivationes* in which the lemmata are arranged in derivational groups (so, for instance, *auris* ‘ear’, *obediens* ‘obedient’, and *subaudio* ‘take as implied’ are all under *audio* ‘hear’), followed by a section of *repetitiones* in which words that do not fall into those groups are listed together with the less common words from the derivational groups.³⁰ This was a powerful statement of a lexical orientation

²⁶ For these schools, with reference to Papias, see Witt, *Two Latin Cultures*, 132.

²⁷ Rouse and Rouse, *Authentic Witnesses*, 167–70. ²⁸ Brito, *Summa* (1975), is an edition.

²⁹ Osberno, *Derivazioni* (1996), is an edition (the entry count is mine, based on its index); description from Thomson, ‘England and the twelfth-century Renaissance’, 8, 13.

³⁰ Weijers, ‘Lexicography in the Middle Ages’, 142; see also Osbern’s own account, in *Derivazioni* (1996), I.3.

as opposed to an encyclopedic one: Osbern was presenting words in their relationship to other words.

One of Osbern's continental European readers was Hugutio of Pisa (Uguccione), whose own even larger dictionary, the *Derivationes* (or *Magnae derivationes*, as it came to be called), was compiled in the second half of the twelfth century.³¹ About 210 manuscripts are extant, and it has been seen as initiating a revival of grammatical studies in thirteenth-century Italy.³² Its principal source is the *Panormia*, although it also draws on Papias, Isidore, and others.³³ It built on Osbern's interests in derivational grouping, seeking to bring all the words it treated, a total of more than 26,000, together in derivational groups, of which there are 2,834, some of them lengthy.³⁴ So, the *Derivationes* was difficult to use as a ready reference tool, and alphabetical indexes were prepared to make up for this. Although it needed to be used with patience, it is a mistake to dismiss it on that account as having enjoyed 'a reputation and use out of all proportion to its real worth' or as 'surely not a practical instrument'.³⁵ Its worth for its readers and copyists was as a highly coherent presentation of the vocabulary of Latin, and they evidently did not feel that learning about a word in the context of related words was less desirable than looking it up quickly out of context. It was used by Dante and Petrarch.³⁶ It was also to be found chained in the choirs of late medieval English secular colleges for regular consultation, and it was cited frequently in Italian school exercises of the Quattrocento.³⁷

The thirteenth-century Italian revival of grammatical studies initiated by Hugutio reached its high point with the *Catholicon* of Giovanni Balbi, which was completed on 7 March 1286.³⁸ This work is a grammatical compendium in five parts, of which the first four are conventional – orthography, prosody, derivation and syntax, and figures of speech – and the fifth, and by far the largest, is a Latin dictionary, running to roughly 15,000 main entries, many of them lengthy, with another 15,000 subentries.³⁹ From the fourteenth century

³¹ Uguccione, *Derivationes* (2004), is an edition.

³² Bursill-Hall, *Census*, 308; Black, *Humanism and Education*, 54–5.

³³ Uguccione, *Derivationes* (2004), I.xxv.

³⁴ My counts, based on the index and entry numbering of Uguccione, *Derivationes* (2004).

³⁵ Daly and Daly, 'Some techniques', 235; Weijers, 'Lexicography in the Middle Ages', 143.

³⁶ Hunt, *Teaching and Learning Latin*, 1383.

³⁷ Willoughby, 'Provision of books', 176–7; Black, *Humanism and Education*, 246, 261, 273.

³⁸ There is no modern edition, but the *editio princeps* of 1460 appears to be a good representative of a fairly stable manuscript tradition: see Powitz, 'Le *Catholicon*', 301–3, for this point and for biographical details. The date is in Balbi, *Catholicon* (1460), fo. 372r, 'In M cc lxxxvi anno d[omi]ni nonis marcij'.

³⁹ Balbi, *Catholicon* (1460), fos. 1–64 (grammar), 65–372 (dictionary); entry count mine (a similar count is reported by Shaw, *Contributions*, sect. 1.3.5); overview in Moss, *Renaissance Truth*, 15–19.

onwards, the dictionary might be transmitted without the grammatical sections, and vice versa.⁴⁰ Balbi explained that his work might be called *Catholicon* because of its catholicity, being *communis et universalis*, applicable to all the learned disciplines of his age.⁴¹

Balbi drew on various sources, including Papias, but his most important single source appears to have been the *Derivationes* of Hugutio, transformed by rearrangement. Although he was prepared to group derived forms together in a single entry (so, *audio* includes *auditor*, *exaudio*, *subaudio*, and others), he broke up the long, complex entries of the *Derivationes* into much more manageable pieces (so, *auris* and *obedio* have their own entries). The alphabetical order of lemmata was absolute: as Balbi explained in the prologue to the dictionary, not only was *amo* before *bibo*, and *abeo* before *adeo*, but *polisintheton* was before *polissenus*.⁴² This was not an advance on the carefully devised macrostructure of the *Derivationes*, but an alternative to it: a reader with access to both dictionaries could now choose the coherence of derivational arrangement or the convenience of alphabetical arrangement. Balbi was aware that a large dictionary must be ready to exclude material – ‘it is madness to learn extra, irrelevant material when time is so short’ – but he did add encyclopedic material to what he took over from the *Derivationes*.⁴³ He oriented the *Catholicon* towards the study of ancient literature, bringing a wide range of original quotations into it. Because the *Catholicon* was such a large dictionary, abbreviated versions were prepared. These might actually have a higher entry count than that of the dictionary of 1286, but the text of their entries was shortened.⁴⁴

There are about 190 manuscripts of the *Catholicon*, of which about 120 belong to the fifteenth century; there are more manuscripts of the *Derivationes* in total, but only about 60 of them are from the fifteenth century.⁴⁵ Likewise, of the large Latin dictionaries with vernacular interpretamenta to which we shall turn below, the only one derived from the

⁴⁰ Powitz, ‘Le *Catholicon*’, 316–17.

⁴¹ Balbi, *Catholicon* (1460), fo. 1r, ‘liber iste uocetur catholicon, eo q[uod] sit co[m]munis et vniu[er]sal[is]. Valet siquide[m] ad om[n]e[s] ferme scientias.’

⁴² The prologue (Balbi, *Catholicon* (1460), fo. 65r) is edited in Daly and Daly, ‘Some techniques’, 237, and translated in Daly, *Contributions*, 73.

⁴³ Balbi, *Catholicon* (1460), fo. 1, ‘Dementia ecia[m] est sup[er]uacua addisce[re] in tanta temp[or]is egestate’ (quoting Seneca, *Epist.* 48.12); for Balbi’s encyclopedism, see Powitz, ‘Le *Catholicon*’, 309–10.

⁴⁴ Shaw, *Contributions*, sect. 1.3.5.

⁴⁵ Powitz, ‘Le *Catholicon*’, 321–36; Bursill-Hall, *Census*, 308 and *passim*. In both cases, the count of fifteenth-century manuscripts includes those which may be of the fourteenth or fifteenth century.

Derivationes, the *Declarus* of Angelo Senisio, is a work of the mid fourteenth century; comparable dictionaries of the late fourteenth and fifteenth centuries are derived from the *Catholicon* or other sources.⁴⁶ And whereas the *Catholicon* was printed as early as 1460, with about thirty further editions, there was never a printed edition of the *Derivationes*.⁴⁷ Manuscripts of the latter were undoubtedly available to printers, but the declining popularity of the book must have been so clear that no printer was prepared to take the risk of setting it up in type. By the end of the fifteenth century, the circulation in manuscript and print of the *Catholicon* had eclipsed that of the *Derivationes*, let alone the *Elementarium*, and it had become the most important large dictionary in Latin Christendom. Its successors belonged to the new age of humanism.

An account of the Greek lexicography of Latin Christendom is inevitably only a footnote to the story of Latin lexicography.⁴⁸ The Latin–Greek dictionary called ‘Glossarium Philoxeni’ was known to a few Carolingian scholars but, since they wanted to read Greek rather than to write it, it cannot have been very useful to them; the Greek–Latin dictionary called ‘Glossarium Cyrilli’ survives in three Carolingian manuscripts, and began to circulate again in fifteenth-century Italy.⁴⁹ As I noted above, the *Hermeneumata* were available in the medieval West; two related Greek–Latin glossaries containing material from this source are extant in a total of seven manuscripts, written between the ninth and thirteenth centuries, with the Greek lemmata in Roman letters.⁵⁰ A Greek–Latin dictionary of some 16,000 entries was copied in England in the thirteenth century; the original may have been compiled in southern Italy, a meeting place of the Greek and Latin worlds.⁵¹ (A few Western scholars knew enough Greek to read monolingual Greek dictionaries: the thirteenth-century English bishop Robert Grosseteste is one example, and another is his younger contemporary Willem van Moerbeke, the translator of Aristotle, who spent some time in Greece at a time when the Crusader kingdom of Constantinople had extended the reach of Latin Christendom across much of the eastern Mediterranean.)⁵² There were

⁴⁶ For the source of the *Declarus*, see Marinoni, *Dal ‘Declarus’ di A. Senisio*, xxv–xxix.

⁴⁷ For early editions of the *Catholicon*, see Considine, ‘Si hoc saeculo natus fuisset’.

⁴⁸ For an overview, see the references indexed at Boulhol, *Grec langaige*, 402 s.vv. *dictionnaires* and *glossaires bilingues*.

⁴⁹ Both are edited in *Corpus glossariorum Latinorum*, vol. II, and the former also in *Glossaria Latina*, II.138–291; for both, see Dionisotti, ‘Greek grammars and dictionaries’, 6–15; for the latter in the fifteenth century, Botley, *Learning Greek*, 63.

⁵⁰ *Colloquia of the Hermeneumata*, 2.86–9; texts in *Corpus glossariorum Latinorum*, 3.487–531.

⁵¹ Dionisotti, ‘On the Greek studies’, 24–6.

⁵² For the former, see Dionisotti, ‘On the Greek studies’, 22–4.

some shorter wordlists. But it was only some decades after the revival of Greek studies in fourteenth-century Italy that Greek wordlists would have a significant circulation in the Latin West.⁵³ As for Hebrew–Latin lexicography, its most substantial expression was in the *Interpretationes nominum Hebraicorum*, a tradition of explications of Hebrew proper names from the Bible, which went back to the work of St Jerome in late antiquity, and flowered in the twelfth and thirteenth centuries.⁵⁴

The Celtic Languages

The Latin–Latin glosses on which so much of the monolingual Latin lexicographical tradition described above was founded were soon followed by Latin–vernacular glosses. Until the tenth century, the speakers of language varieties descended from ancient Latin – the ancestors of the modern Romance languages – saw their spoken vernaculars as close enough to written Latin that they did not need to explain Latin words with vernacular words. However, for speakers of the Germanic and Celtic languages, Latin was altogether foreign and called for vernacular glossing.

Christianity and literacy in Latin both came to Ireland from the former Roman empire in the fifth century. The practice of making and collecting glosses doubtless did so as well. The oldest isolated gloss in Old Irish on a Latin text appears to belong to the early seventh century, and about thirty manuscripts of the eighth and ninth centuries are glossed in a mixture of Irish and Latin, some of them extensively.⁵⁵ The making of these Irish and Latin glosses was followed by the compilation of glossaries.⁵⁶ These are early: the oldest of the strata in the text called O'Mulconry's Glossary has been dated to the seventh or early eighth century, although the extant manuscripts are much later.⁵⁷ Another, Cormac's Glossary, has been attributed to Cormac mac Cuilennáin, king of Munster and bishop, who died at the beginning of the tenth century, but is based on earlier material.⁵⁸ A third, called 'Dúil Droma Cetta', for which these two glossaries were sources, can be dated to the tenth century.⁵⁹ Remarkably, although the obvious development from the glossing of Latin texts in Irish would be the making of Latin–Irish

⁵³ Botley, *Learning Greek*, 55–70. ⁵⁴ Dahan, 'Interpretationes'.

⁵⁵ Ó Cróinín, 'Earliest Old Irish glosses', 9–10, 24–7.

⁵⁶ The best texts are now those in the *Early Irish Glossaries Database*.

⁵⁷ Date from MacNeill, 'De origine Scoticae linguae', 113, accepted by Russell, 'Sounds of a silence', 5; Moran, 'Greek in early medieval Ireland', 178; and (cautiously) Mahon, 'Contributions', 3–4, 34–5.

⁵⁸ Russell, 'Sounds of a silence', 28. ⁵⁹ Mahon, 'Contributions', 38.

glossaries, these very early lexicographical texts gloss Irish lemmata in Irish. They are the oldest monolingual wordlists of a vernacular language of Latin Christendom. The linguistic self-confidence to which they witness can be compared to that of the Irish grammatical text called *Auraicept na n-Éces*, the oldest parts of which may also go back to the seventh century, which suggests, as no text in Latin Christendom had done before, that its own language is superior to Latin.⁶⁰ The Irish glossaries contain much etymological material, some of which explains Irish words with reference to words from Latin, Greek, Hebrew, and other languages. This certainly shows the influence of Isidore (and the *Hermeneumata*) and can also be related to a story in the *Auraicept*, according to which the Irish language was devised immediately after the confusion of tongues which ended the building of the tower of Babel, by wise men who took the best from every other language to create Irish.⁶¹

The early medieval Irish glossaries were copied and added to throughout the medieval period. Indeed, it might be said that Gaelic learning continued unbroken well after the Middle Ages had ended in England and in continental Europe, so that O'Davoren's Glossary of 1564, which is rich in legal material, and even Míchél Ó Cléirigh's dictionary of 1643, represent a living tradition.⁶² A later medieval tradition of metrical glossaries may originate in mnemonic verses learned by students at the bardic schools.⁶³

The glossarial traditions of Welsh, Cornish, and Breton are significantly later than that of Irish.

Latin texts were probably being glossed in Old Welsh and in Old Breton by the ninth century, and in Old Cornish by the late ninth or tenth.⁶⁴ A collection of 961 glosses made in Cornwall around AD 1100 includes material from all three languages, as well as Old English and Old French, and is the only sustained medieval attempt at a Cornish wordlist.⁶⁵ After these early glosses, lexicographical activity in Cornish came to an end, and activity in Welsh and Breton paused for a long time. In the fifteenth century, Welsh was recorded in monolingual wordlists, notably collections of words for the use of poets practising in the rich bardic tradition. One bard of the second half of the fifteenth century, Gwilym Tew, collected and commented on

⁶⁰ Ahlqvist, *Early Irish Linguist*, 19, 40.

⁶¹ Moran, 'Greek in early medieval Ireland', 178–89; Ahlqvist, *Early Irish Linguist*, 40, 47.

⁶² For both, see Russell, 'Sounds of a silence', 6, 8–9, and for the former, see also Mahon, 'Contributions', 51–3.

⁶³ Mahon, 'Contributions', 44–8. ⁶⁴ Jackson, *Language and History*, 49–67.

⁶⁵ See Mills, 'Vocabularium Cornicum'.

vocabulary from a thirteenth-century manuscript which preserves much older Welsh poetry, called the Book of Aneirin.⁶⁶ The compilation of manuscript wordlists by bards working in the medieval tradition, such as the poet and herald Gruffudd Hiraethog and his pupil Simwnt Fychan, continued into the sixteenth century.⁶⁷ No Middle Breton lexicography is recorded until the Breton–French–Latin *Catholicon en troys langaiges* of Jehan Lagadeuc (so called because it was an alphabetical dictionary like the *Catholicon* of Giovanni Balbi, not because it was actually an adaptation of this dictionary), of which there is a manuscript dated 16 August 1464 and three printed editions. The first of these, of 1499, was the first printed specimen of the Breton language.⁶⁸

The Germanic Languages: Early Glossaries

Anglo-Saxon rulers became Christian in the sixth and seventh centuries, and their language began to be written down in the seventh. A rich glossarial and lexicographical tradition had begun before the end of this century, the earliest for any Germanic language.⁶⁹ A corpus of three million words of Old English survives, and a quarter of this corpus consists of interlinear glosses to Latin texts, especially psalters, which were studied by young men without much Latin.⁷⁰ Only 1 per cent of the corpus takes the form of actual glossaries but, even so, there are 143 Latin glossaries which include as few as one, and as many as several thousand, Old English interpretamenta, together with others in Latin (or in Old High German, for Anglo-Saxon glossaries circulated in continental Europe and might be added to there); the lemmata are always Latin.⁷¹

The wellspring of Old English lexicography was the school of Theodore and his colleague the abbot Hadrian at Canterbury, established soon after Theodore's arrival there as archbishop in 669.⁷² Theodore was a native speaker of Greek from Asia Minor, and Hadrian was a fluent speaker of Greek and Latin, born in Africa and subsequently resident in Italy; the two

⁶⁶ Burdett-Jones, 'Early Welsh dictionaries', 75; Jarvis, 'Welsh humanist learning', 135.

⁶⁷ Their wordlists are National Library of Wales, MSS Peniarth 230 and Peniarth 189, pp. 13–84, described in Evans, *Report*, 1058–9 and 1015 respectively.

⁶⁸ Lagadeuc, *Catholicon* (1975), is a facsimile of the 1499 edition with a substantial introduction; see also Lindemann, *Die französischen Wörterbücher*, 278–84; the manuscript is Paris, Bibliothèque Nationale, ms lat 7656.

⁶⁹ Franzen, *Old English*, xv–lxxxvi, is an excellent overview, and the whole volume is valuable, as is Derolez, *Anglo-Saxon Glossography*.

⁷⁰ Franzen, *Old English*, xxxvi–xxxvii. ⁷¹ Franzen, *Old English*, xv.

⁷² Lapidge, 'School of Theodore and Hadrian'; Lapidge and Bischoff, *Biblical Commentaries*, 173–9.

brought the learned culture of the Roman Mediterranean to England. Lists of *glossae collectae* which evidently originate in their teaching, and which include material in Old English, occur with more or less rearrangement in some twenty continental European manuscripts (the best-known, now called the Leiden Glossary, which includes some 250 Old English glosses, was written at St Gallen in what is now Switzerland, around 800) and in some major glossaries compiled in England. These include the Épinal-Erfurt Glossary, which is perhaps of the last quarter of the seventh century, in which case it is one of the earliest texts in English. It may originally have run to some 3,800 entries, a third or a quarter of which had Old English interpretamenta.⁷³ Aldhelm (whom we last saw enjoying the word *Molossus*) contributed to it – making him the first identifiable Englishman to have undertaken lexicographical work of any sort – and drew on it in his own writings, a reminder of the two-way relationship between lexicography and literature.⁷⁴ The manuscripts in which the Épinal-Erfurt Glossary is transmitted are large (up to 360 mm tall) and formally written; the ninth-century manuscript of the related Corpus glossary is particularly beautiful. New glossaries continued to be compiled into the first half of the eleventh century. An impressive late example in ABC- and sometimes even ABCD- order, known as the Harley glossary, runs to some 5,500 entries for the surviving ranges Ab–Ad and B–F, about a third of them with Old English interpretamenta.⁷⁵ But the tradition which produced this fine work was cut short by the invasion of England in 1066 by the French-speaking William of Normandy. In the two centuries after the invasion, the English language changed considerably, and the surviving glossaries with Old English interpretamenta gradually ceased to be consulted until their rediscovery by sixteenth-century antiquarians.⁷⁶

Meanwhile, the speakers of the Germanic language varieties east of the Rhine had been evangelized in the seventh and eighth centuries, a process in which Irish and English missionaries played a part. This process brought about the beginnings of German lexicography. The first German glosses date from the eighth century, the very first being perhaps a group which may belong to the early part of the century, accompanied by Old English glosses

⁷³ Lapidge and Bischoff, *Biblical Commentaries*, 179; Franzen, *Old English*, xxvii; edited in *Corpus glossariorum Latinorum*, 5.337–401 and, selectively, in *Old English Glosses to the Épinal-Erfurt Glossary* (1974) (reconstructed entry count at xxi); there is a fine facsimile of both manuscripts in *The Épinal, Erfurt, Werden, and Corpus Glossaries* (1988).

⁷⁴ Lapidge, 'Career of Aldhelm', 31–48.

⁷⁵ Franzen, *Old English*, lvii; there is an unsatisfactory edition as *The Harley Latin–Old English Glossary* (1966).

⁷⁶ Franzen, *Middle English*, xvi–xx.

(and some dry-point glosses which have been uncertainly identified as Old Irish) in a manuscript written at Echternach, a monastery founded by Anglo-Saxon missionaries.⁷⁷ An enormous amount of further glossarial activity took place in German monasteries from the eighth century to the twelfth, and more than a thousand surviving manuscripts bear Old High German or Old Saxon glosses.⁷⁸ This was carried out side by side with the making of glossaries. The first Latin–German glossary, *Abrogans* (formerly called the Keronian glossary from the name of a supposed compiler), a version of a Latin glossary with some 3,670 German interpretamenta, belongs to the late eighth century.⁷⁹ It is much more substantial than any earlier German-language document and has therefore been called ‘the first book in the German language’.⁸⁰ There are other wordlists of a similar age, for instance the classed Latin–German glossary called the ‘*Vocabularius Sancti Galli*’, the Latin lemmata of which appear to be from the *Hermeneumata*.⁸¹ A German–Latin list of words and phrases to do with travel, known as the ‘*Pariser Gespräche*’ or the ‘*Altdeutsche Gespräche*’, was written in the tenth century.⁸² Some of the early manuscripts in which these texts are transmitted, for instance the St Gallen *Abrogans* manuscript and that of the ‘*Vocabularius Sancti Galli*’, are much smaller and less formally written than the finest manuscripts of the Old English glossaries.

The Germanic Languages: Later Dictionaries

The early German glossarial tradition died out in the thirteenth century, no doubt because the lists of Latin lemmata gathered in Latin–German glossaries were less extensive than those gathered in Latin dictionaries from the *Elementarium* of Papias onwards.⁸³ It was succeeded by lexicographical traditions in which German (and, in so far as the two were distinguishable in the fifteenth century, Dutch) interpretamenta were added to Latin dictionaries. One was the *Vocabularius brevilogus* (also known by the slightly later title *Vocabularius brevilogus*), represented by forty-three manuscripts,

⁷⁷ Seebold, ‘Textgeschichte’, 35; Bronner, ‘Dry-point glosses’, 3.

⁷⁸ Bergmann, *Verzeichnis*, presents entries for 1,023 manuscripts; an overview is in Bostock, *Handbook*, 90–103.

⁷⁹ *Das älteste deutsche Buch* (1977) is a facsimile, with commentary volume, of one of the manuscripts, St Gallen, Cod. Sang. 911. Word count and Latin source from Splett, ‘*Abrogans deutsch*’.

⁸⁰ Bostock, *Handbook*, 96. ⁸¹ St Gallen, Cod. Sang. 913; see Bostock, *Handbook*, 100.

⁸² Paris, BN ms lat 7641; Bostock, *Handbook*, 101–3.

⁸³ Grubmüller, *Vocabularius ex quo*, 46–8.

predominantly from northern Germany, and by twenty-two printed editions from 1475–6 onwards, the first of them prepared for the press by the future humanist Johannes Reuchlin.⁸⁴ The Latin lemmata of the dictionaries in this tradition appear to derive from Balbi's *Catholicon*, though they have been rearranged into three sections, namely nouns, verbs, and *indeclinabilia*. The majority of entries also have Latin interpretamenta, and perhaps one in twenty adds German interpretamenta. Another dictionary from the second half of the fourteenth century, the *Vocabularius Lucianus*, so called because it brings lexical obscurities into the light (*in lucem*), is extant in fifty-five manuscripts, predominantly from Bavaria and Austria.⁸⁵ It too has some German interpretamenta, and a note to the reader expresses the hope that they will not be found objectionable, suggesting that their presence in a dictionary of Latin is a novelty. The most widely circulated Latin–German dictionary of the later Middle Ages was the so-called *Vocabularius ex quo*, extant in 289 known manuscripts from 1410 onwards – more, as far as I know, than any other medieval dictionary – and in 48 early printed editions. It registered some 20,000 words.⁸⁶ The largest German–Latin wordlists of the fifteenth century do not seem to have been quite so ample: the *Vocabularius incipiens Teutonicum* of c. 1483–4 had about 13,100 entries.⁸⁷

As well as general alphabetical dictionaries, there were subject-ordered Latin–German wordlists such as the *Liber ordinis rerum*, probably compiled at the end of the fourteenth century, which presents more than 10,000 entries in 250 classes, and is extant in 70 manuscripts.⁸⁸ There were also technical wordlist traditions such as the philosophical and theological *Abstractum*, compiled around 1300 (and indebted to the *Expositiones* of Brito), which is extant in about a hundred manuscripts.⁸⁹ In these cases and that of the general dictionaries alike, the exceptional liveliness of medieval German lexicography can be seen; in no language other than German were so many dictionaries produced and copied.

The lexicography of Middle English developed slowly.⁹⁰ At first, Latin texts and glossaries were glossed in English, Anglo-Norman, or

⁸⁴ Grubmüller, *Vocabularius ex quo*, 31–9; Benzing, *Bibliographie*, items 1–22.

⁸⁵ Grubmüller, *Vocabularius ex quo*, 39–44.

⁸⁶ *Vocabularius ex quo* (1988–2001) is an edition; manuscript count from *Handschriftencensus*.

⁸⁷ Entry count mine.

⁸⁸ *Liber ordinis rerum* (1983) is an edition: the manuscripts are listed I.xiii–xv, and the date is discussed I.xciv; entry count my own from the indexes in vol. II. See also the subject-ordered Latin–Low German or Latin–Middle Dutch wordlists discussed and edited in De Man, *Middeleeuwse systematische glossaria*.

⁸⁹ Overview in Illing, 'Abstractum-Glossar'.

⁹⁰ Overview in Franzen, *Middle English*, xv–lxviii; the whole volume is valuable.

both.⁹¹ Material in both of these vernaculars which occurs in one manuscript of Osbern Pinnock's *Panormia* may be authorial, and therefore of the twelfth century.⁹² The trilingual English society of the thirteenth and fourteenth centuries was served by trilingual wordlists, and by other texts such as the thirteenth-century *Tretiz* of Walter de Bibbesworth, a rhyming Anglo-Norman vocabulary with some English glosses.⁹³ As the use of Anglo-Norman declined, the focus of English lexicography became the making of Latin–English wordlists and the addition of English glosses to monolingual Latin dictionaries.⁹⁴

This led to the development of new traditions of dictionaries in fourteenth- and fifteenth-century England, comparable to the new traditions of fourteenth-century German-speaking Europe (however, the circulation of the English dictionaries, both in manuscript and in print, was much more restricted).⁹⁵ An early example, a dictionary in the *Catholicon* tradition known from a manuscript copied in the early fifteenth century for the library of Battlefield College near Shrewsbury, has English interpretamenta in a minority – some 2,600 in all – of its entries. Members of the so-called *Medulla Grammaticae* tradition of the late fourteenth and fifteenth centuries, which like the Battlefield College dictionary were related to the *Catholicon*, might have English interpretamenta in a majority of their 13,000 to 20,000 entries. A wordlist in this tradition was printed in 1500 with the title *Ortus vocabulorum*. Two traditions which reverse *Medulla* entries so that their lemmata are in English and the interpretamenta in Latin are known by the titles *Promptorium parvulorum* and *Catholicon anglicum*; the former was printed in 1499 (the *editio princeps* has 10,502 entries).⁹⁶ These late medieval dictionaries were superseded in the second and third quarters of the sixteenth century.

Christianity came later to the speakers of the north Germanic languages of Scandinavia than to the speakers of the west Germanic languages of England and of continental Europe south of Jutland. The lexicography of the north Germanic languages was correspondingly belated.⁹⁷ The first extant Latin–Icelandic glosses date from about 1200, and there are Latin–Swedish glosses

⁹¹ Hunt, *Teaching and Learning Latin*, is a detailed study in which glosses are collected and indexed.

⁹² Hunt, *Teaching and Learning Latin*, I.376–8.

⁹³ Franzen, *Middle English*, xxx–xxxvii; Bibbesworth, *Le Tretiz* (1990), is an edition.

⁹⁴ Franzen, *Middle English*, xxvii–xlii.

⁹⁵ Overview of scholarship to date in Franzen, *Middle English*, xlii–xliii and xlvii–liv.

⁹⁶ Entry count from *Lexicons of Early Modern English*.

⁹⁷ Overview in Raschellà, 'Vernacular gloss writing'.

from the thirteenth century and Latin–Danish glosses from the fourteenth. An abridgement of the *Catholicon* with Swedish interpretamenta added to many of the entries, as in the *Catholicon*-derived dictionaries of England and of German-speaking Europe, is extant in a single manuscript of the second half of the fifteenth century.⁹⁸

The Romance Languages

As I noted above, written Latin only gradually became a foreign language for the speakers of the language varieties which turned into French, Spanish, Italian, and the other Romance languages. Two eighth-century glossaries from northern France give equivalents for a total of 4,877 of the hard words in the Vulgate Bible, and a couple of hundred of these gloss an older Latin word with a newer Latin form which developed into a French word, as when Latin *caseum* ‘cheese’ is glossed *formaticum* (compare modern French *fromage*).⁹⁹ These interpretamenta can hardly be described as being in a Romance language. When the lexicography of Italian, Spanish, and French did begin in the tenth and eleventh centuries, it was at first in the form of bilingual glosses and wordlists, in which the second language was often not Latin.

The sixty-six or so lemmata in the Monza Vocabulary, a tenth-century glossary with Demotic Greek interpretamenta, are sometimes called Italian rather than Latin, in which case they would stand at the beginning of Italian lexicography.¹⁰⁰ The first Italian interpretamenta may be a group of 150 glosses made in the eleventh century in the dialect of Otranto on a Hebrew text. Similar glosses continued to be made into the sixteenth century. From the thirteenth century onwards, a short Hebrew–Italian wordlist of philosophical terms was copied, and in the fifteenth century a trilingual Biblical Hebrew–Italian–Arabic wordlist called *Maqre Dardeqe* was compiled by a French resident of Italy called Perez Trabot (see Chapter 9). In all of these cases, the Italian was written in Hebrew characters.¹⁰¹

Glosses in Italian on Latin texts appear from the first half of the thirteenth century onwards, and are particularly associated with the less advanced school texts, reflecting the confidence with which the highly literate

⁹⁸ *Latinskt–Svenskt glossarium* (1918–94) is an edition.

⁹⁹ *Reichenauer Glossen* (1968); *formaticum* is at 1.158 (line 285).

¹⁰⁰ Aerts, ‘Monza vocabulary’, provides text and commentary; Schweickard, ‘Italian’, 673, writes that they ‘cannot be considered as Italian’.

¹⁰¹ Rubin, ‘Judeo-Italian’, 308–9 (glosses) and 304–5 (wordlists).

Italians who had progressed beyond these texts handled Latin.¹⁰² A Latin dictionary (based, as noted above, on the *Derivationes* of Hugutio rather than on the *Catholicon*) with about 1,400 Sicilian forms added to its Latin interpretamenta, compiled by the Benedictine abbot Angelo Senisio, is dated 1348.¹⁰³ Latin–Italian wordlists were compiled by teachers of grammar from the fourteenth-century Goro d’Arezzo onwards, a number of them being, like Goro’s, arranged by subject. Although Goro was a Tuscan, not all the compilers were, and so the Italian interpretamenta were in a number of language varieties, for instance the Bergamask of the fifteenth-century teacher Gasparino Barzizza.¹⁰⁴ From the fourteenth century onwards, wordlists of Italian and other vernaculars, especially German, were compiled for use in commerce; the first Italian–Latin wordlists appeared in the fifteenth century, and short monolingual wordlists were compiled in the same century.¹⁰⁵ So by the end of the fifteenth century, Italian lexicography was flourishing, although there was not yet a large Latin–Italian dictionary.

The language of the ‘Glosas Emilianenses’, which date to the late tenth or early eleventh century, is an Iberian Romance variety – Aragonese rather than Castilian, but still sometimes described loosely as the earliest written Spanish. The model for these glosses, which amount to some 145 words, was probably insular and perhaps Irish rather than English.¹⁰⁶ As was true of the Germanic languages, there was a discontinuity between the earliest Spanish glossarial activity and the lexicography of the later Middle Ages, which began with two wordlists of the late fourteenth century and continued with at least two more before 1490.¹⁰⁷ The first printed Latin–Spanish dictionary appeared in that year: the *Universal vocabulario en latin y en romance* of the chronicler Alfonso Fernández de Palencia. Rather than drawing on a manuscript tradition, it presented a text of the dictionary of Papias, taken from the *editio princeps* of 1476, in parallel with a version in which the interpretamenta were translated into Spanish.¹⁰⁸ It was not reprinted because, only two years later, the

¹⁰² Black, *Humanism and Education*, 275–7.

¹⁰³ Marinoni, *Dal ‘Declarus’ di A. Senisio*, is an edition of the Sicilian interpretamenta; for the date, see *ibid.*, ix; the entry count is mine, from the index, *ibid.*, 265–79.

¹⁰⁴ Black, *Humanism and Education*, 107–9; Schweickard, ‘Italian’, 672–3; Olivieri, ‘Primi vocabolari’, 75–83.

¹⁰⁵ Schweickard, ‘Italian’, 673–4; Pausch, *Das älteste italienisch–deutsche Sprachbuch*, is an edition of an Italian–German wordlist of 1424, for which see also Hüllen, *English Dictionaries*, 326–31.

¹⁰⁶ *Glosas Emilianenses* (1996) is an edition; for the insular model, see *ibid.*, 43, 63.

¹⁰⁷ Bibliographical overview in Niederehe, *Bibliografia*, 7–23.

¹⁰⁸ Niederehe, ‘Universal vocabulario’, sets out the relationship.

Latin–Spanish *Dictionarium ex sermone Latino in Hispaniensem* of Elio Antonio de Nebrija was published (see Chapter 14).

The earliest Old French glosses are of the eleventh century, and these are never in a purely Latin–French context. A very large early group by a named glossator, Rashi (for whose Hebrew–French lexicography see Chapter 9), are in Hebrew characters and annotate Hebrew texts.¹⁰⁹ Two eleventh-century manuscripts with Old French glosses on Latin texts associate them in one case with glosses in Old English, and in another with glosses in Old High German.¹¹⁰ From the early twelfth century onwards, Latin texts began to be glossed in Anglo-Norman and, as we have seen, a considerable tradition of glosses and glossaries developed from these beginnings.¹¹¹ Likewise, Old French glosses began to be gathered into Latin–French glossaries as early as the twelfth century, and into Hebrew–French glossaries from the beginning of the thirteenth.¹¹²

By the end of the thirteenth century, the lexicography of French was beginning to draw on the Latin dictionary tradition. *Abavus*, a Latin wordlist derived from the *Elementarium* of Papias with French interpretamenta, is extant in six manuscripts, of which the oldest, perhaps an abridgement of a lost predecessor, runs to 2,662 entries. Two fourteenth-century manuscripts in the tradition run to more than 9,000 entries, and draw on Balbi's *Catholicon* as well as on Papias.¹¹³ Another Latin–French tradition, based primarily on the *Catholicon*, is called *Aalma*; the oldest of the fifteen extant manuscripts, which belongs to the second half of the fourteenth century, has been printed and runs to 13,680 entries, but an unprinted member of the same tradition is nearly twice as long.¹¹⁴ Dictionaries in this tradition were printed under titles such as *Catholicon abbreviatum* from the early 1480s to the 1520s.¹¹⁵ Other large Latin–French dictionaries of the fifteenth century were likewise derived from the *Catholicon*, for instance the manuscript *Dictionarius* of Firmin Le Ver and the closely related *Vocabularius familiaris et compendiosus* of Guillaume Le

¹⁰⁹ Kiwitt and Dörr, 'Judeo-French', 141–4, 154–6, 164–5; Lindemann, *Die französischen Wörterbücher*, 120–2.

¹¹⁰ Hunt, *Teaching and Learning Latin*, 1.9; De Cesare, 'Su di un gruppo di glosse antiche francesi'.

¹¹¹ Early examples are in Hunt, *Teaching and Learning Latin*, 1.19–27.

¹¹² Lindemann, *Die französischen Wörterbücher*, 124–35.

¹¹³ Four versions are edited, with numbered entries, in Roques, *Recueil général*, vol. I; overview in Shaw, *Contributions*, sect. 2.5.1, and discussion in Lindemann, *Die französischen Wörterbücher*, 136–49 and 175–201.

¹¹⁴ Roques, *Recueil général*, vol. II; see also Merrilees, 'Deux versions de l'*Aalma*', Shaw, *Contributions*, sect. 2.5.3, and Lindemann, *Die französischen Wörterbücher*, 201–21.

¹¹⁵ Lindemann, *Die französischen Wörterbücher*, 222–40.

Talleur (of which there were editions of 1490 and 1500), of 12,874 and 17,813 lemmata respectively.¹¹⁶ A Latin dictionary of 4,825 entries, about a quarter of them having interpretamenta in the Picard dialect of French, was compiled around 1335–40 on the basis of material from Brito's *Expositiones* and of French glosses from a copy of Evrard of Béthune's grammatical text *Graecismus*; a Picard–Latin wordlist in A- order is written in its margins. A French–Latin dictionary of some 9,500 entries, conventionally known as *Glossarium Gallico–Latinum*, closely related to the work of Firmin Le Ver and therefore to the *Catholicon* tradition, was compiled in or around the second quarter of the fifteenth century.¹¹⁷

Other Languages

Small collections of words from other languages were made here and there in medieval Latin Christendom. An etymological list of eighteen place-name elements, some of them Gaulish, is preserved with other glossaries in a manuscript of about AD 800, and may be one or two centuries older than the manuscript.¹¹⁸ A twelfth-century guide for pilgrims to Santiago de Compostela has a Latin–Basque wordlist of a dozen items for the use of those who would pass through the Basque country – though the first glimmer of Basque lexicographical activity is a century earlier, for two of the 'Glosas Emilianenses' are in Basque.¹¹⁹ Fifteenth-century guides have brief wordlists of Turkish, modern Greek, and colloquial Arabic and Hebrew for the use of pilgrims in the eastern Mediterranean; and a fifteenth-century manuscript, probably also compiled for practical use, preserves forty words of Yassic, an Iranian language formerly spoken in Hungary.¹²⁰

The most substantial wordlists of languages of Latin Christendom other than those of the Celtic, Germanic, and Romance families are those of the Slavic languages that were spoken by Roman Catholics and written in the Roman alphabet. The first glosses in Czech were made between the tenth and thirteenth centuries and, like some of the earliest glossarial activity in

¹¹⁶ Entry counts from the *Database of Latin Dictionaries*, which provides searchable versions of the editions published as Le Ver, *Dictionarius* (1994), and Le Talleur, *Dictionarius familiaris et compendiosus* (2002); discussion in Lindemann, *Die französischen Wörterbücher*, 240–7.

¹¹⁷ *Duo glossaria* (1998) is an edition of all three texts; entry counts from the *Database of Latin Dictionaries* and from Shaw, *Contributions*, sect. 2.5.5.

¹¹⁸ Blom, 'Endlicher's glossary'.

¹¹⁹ Considine, *Small Dictionaries*, 73; *Glosas Emilianenses*, 112–13.

¹²⁰ Considine, *Small Dictionaries*, 21 (eastern Mediterranean), 18 (Yassic).

French and Italian, are on Hebrew texts, and written in Hebrew characters.¹²¹ Three Latin–Czech lists of technical vocabulary survive from the fourteenth century, all in verse so as to be memorized by students, the longest running to 7,000 lemmata. A subject-ordered Latin–Czech dictionary of 3,485 lemmata has been edited from a manuscript of 1409. All of these wordlists are associated with the name of the fourteenth-century Bohemian teacher Claretus de Solentia (Bartoloměj z Chlumce).¹²² Besides this Czech material, a Latin–Polish wordlist of some 400 lemmata was compiled around 1420, and there is at least one Latin–Croatian wordlist of the latter part of the fifteenth century (for these fifteenth-century wordlists in the context of Slavic lexicographical traditions, see Chapter 21).¹²³

Conclusion

The story which has been told in this chapter can be summed up in a few points. First, the lexicography of Latin Christendom in what we call the Middle Ages was overwhelmingly Latin. The central tradition is that of Latin glossaries and dictionaries, from the early *glossae collectae* to the *Liber glossarum*, and thence to the *Elementarium* of Papias, the *Derivationes* of Hugutio, and the *Catholicon* of Giovanni Balbi with its numerous descendants. Many of the large Latin–vernacular dictionaries of the fourteenth and fifteenth centuries derive from this tradition. Secondly, a number of vernaculars were treated by lexicographers: significant wordlists documented the vocabularies of Czech, English, French, German, Irish, Italian, Spanish, Swedish, and Welsh. Although most bilingual dictionaries were Latin–vernacular, there were some vernacular–Latin dictionaries, and a few in which lemmata in one vernacular were answered by interpretamenta in another. However, with the exception of the monolingual wordlists of Irish throughout the period and of Welsh and Italian in the fifteenth century, very little of the work we have surveyed had a vernacular language as its principal object of inquiry.

Two further points about the lexicographical vision of the period may be added. The first of these arises from the observation, with which I agree, that ‘the Middle Ages did not start out with a clear example of what a dictionary could be’.¹²⁴ I have argued elsewhere that the concept ‘dictionary’ scarcely

¹²¹ Hill, ‘Judeo-Slavic’, 599–603.

¹²² Stankiewicz, *Grammars and Dictionaries*, 16; Hüllen, *English Dictionaries*, 362–5.

¹²³ Stankiewicz, *Grammars and Dictionaries*, 45 (Polish), 84 (Croatian).

¹²⁴ Weijers, ‘Lexicography in the Middle Ages’, 139.

existed in 1500, let alone earlier.¹²⁵ But if the dictionary as we know it is an inventory of all the words in normal use in a given language, made up of discrete entries in each of which a different headword is followed by information about its meaning and use, and structured in such a way that entries can be consulted individually, then the dictionary as we know it is rooted in the lexicographical traditions of the languages of medieval Latin Christendom. It was this tradition which, especially as handled in the alphabetical dictionaries of Papias in the eleventh century and Giovanni Balbi in the thirteenth, provided a formal model for the printed dictionaries of the Renaissance and their successors.

A second point, which balances this one, is that the lexicographical traditions surveyed in this chapter must not simply be read teleologically. They deserve attention in their own right, as a rich and varied body of work – and a surprisingly cosmopolitan one. This last quality came in part from their debt to ancient lexicography: any indebtedness to the *Hermeneumata*, for instance, ultimately reflected the extension of the Roman empire across parts of three continents. The roots of Anglo-Saxon lexicography were in the teachings of Theodore and Hadrian, and hence in the intellectual life of late antique Africa and Asia Minor. Early glossarial activity in French, Italian, and Czech was a product of the Jewish diaspora. And, as we saw in Chapter 11, the vision of the makers of the Codex Comanicus wordlists extended into Turkic and Persianate Asia. The story of the lexicography of Latin Christendom is part of a story of far-reaching cultural transmission and reception.

¹²⁵ Considine, 'History of the concept of lexicography'.

Early Modern Western and Central Europe

JOHN CONSIDINE

The drama of the lexicography of medieval Latin Christendom takes place in the theatre of Latin lexicography, as we see the completeness of the transformation from the modest glossographical traditions of late antiquity to the conceptual grandeur and the ample extent of Latin dictionaries such as the *Panormia* of Osbern Pinnock or the *Derivationes* of Hugutio, or to the equal extent and elegant alphabetized information management of Latin–Latin and Latin–vernacular dictionaries such as the *Catholicon* of Giovanni Balbi and its successors. The drama of the lexicography of early modern western Europe lies partly in the theatre of Latin and Greek lexicography. The set of changes to high culture which we call the Renaissance led the most literate western Europeans into a new relationship to the Latinity of the ancient world, entailing a rejection of that of the Middle Ages, and this demanded a new generation of Latin dictionaries. The rediscovery of ancient Greek which was a salient feature of the Renaissance led to the development of Greek–Latin dictionaries; as we shall see, one of these was the largest and most complex of all the early modern dictionaries of European languages. But even more dramatic was the broad and rapid development of the lexicography of a number of the Romance and Germanic languages: a Latin dictionary of the 1690s does not look completely unlike the *Catholicon*, and a Greek dictionary of the 1690s does not look completely unlike some late antique Greek–Latin dictionaries, whereas the *Dictionnaire de l'Académie française*, also published in the 1690s, does look completely unlike any French wordlist of the age of the *Catholicon*.

As we saw in Chapter 13, several of the Romance and Germanic vernaculars had appeared in medieval wordlists of increasing size. By 1500, a vernacular–Latin dictionary in one of these traditions might list some 10,000 words of the vernacular – the French–Latin *Glossarium Gallico–Latinum* had about 9,500 entries, the *editio princeps* of the English–Latin *Promptorium parvulorum* had 10,502, and the German–Latin *Vocabularius*

incipiens Teutonicum had about 13,100 – but what the dictionary said about the structure and use of vernacular words or phrases would be primarily a means to find their Latin equivalents. Over the next two centuries, the status of a number of the vernaculars, which had already been shifting before 1500, rose considerably, and they developed formal literary varieties, which were codified in dictionaries and grammars.¹ (One might say that these formal varieties were vernaculars with respect to Latin but that, in a different perspective, they were classical varieties with respect to the spoken vernacular varieties used by most people in everyday life.) By 1700, formal varieties of three languages – Italian, French, and Spanish – were treated in large, sophisticated, printed monolingual dictionaries, and a fourth, English, was treated in smaller but still wide-ranging monolingual dictionaries. Even when the major dictionaries of a vernacular used Latin as a metalanguage, as in the cases of German, Dutch, and (in an unpublished dictionary) Danish, they might explain the meanings of tens of thousands of words, sometimes with attention to etymology, word-formation processes, and literary use: the largest dictionaries of these three languages were more elaborate than the monolingual dictionaries of English.

The story of early modern European lexicography is not, to be sure, only a story of dictionaries of Latin (and the other learned languages) and of well-documented vernaculars. Many languages were spoken in Europe in the period, some documented very well and others hardly at all, a point to which I shall return at the end of this chapter. Moreover, many languages spoken beyond Europe came to the attention of European traders, conquerors, and most importantly missionaries between 1500 and 1700, and the wordlists in which they were documented will not be discussed here (for them, see Chapters 26–32).

The great majority of the dictionaries discussed in this chapter were, in contrast to the great majority of those discussed in Chapter 13, printed books. Indeed, whereas the most significant background to the story of medieval lexicography in what is now central and western Europe is the practice of writing, the most significant background to the story of early modern lexicography in the same area is the practice of printing from movable type. The ancient technology of writing, inextricably associated with the late antique survival of the Christianity of the western Roman empire, spread slowly with the evangelization, or sometimes re-evangelization, of neighbouring lands. By contrast, the technology of printing from movable type,

¹ See Burke, *Languages and Communities*.

which emerged in German-speaking Europe in the middle of the fifteenth century (its first datable results were published in 1454), spread very quickly through much of Christendom, though its adoption in the Islamic world was long delayed. Its applicability to the production of dictionaries was seen from its earliest years: the first of a series of incunabular editions of the *Catholicon* is dated 1460.² The advantages of printing for reproducing large collections of technical information are evident, and the printed dictionary as a genre flourished with great vigour in early modern Europe.³ Indeed, the fully alphabetized dictionary, in which specific information about words is separated from general information about the grammar of a language, is very much an artefact of print: the *Catholicon* is an important forerunner of the genre, but its lexicographical portion is only the fifth part of a grammatical compendium. The key event in the differentiation of the genres of dictionary and grammar took place in the year 1502, to which we now turn.

Latin

In 1502, the *Dictionarium* of the Italian religious Ambrogio Calepino was printed for the first time. This was a fully alphabetized monolingual wordlist of Latin, with minimal preliminaries and no separate grammatical section: its subject was the vocabulary of Latin, and nothing else. The Latin to which it attended was, as far as possible, that of classical antiquity rather than that of the age immediately before Calepino's own.

So it was that rather than revising and adapting the *Catholicon* of Giovanni Balbi, a process which had proved fruitful for many later medieval lexicographers, Calepino drew much of his material from the *Cornucopiae* of Niccolò Perotti, compiled by 1478 and first printed in 1489.⁴ This was an account of the vocabulary of classical Latin, with a brilliant and rather surprising structure: rather than starting with a medieval wordlist, Perotti started with the epigrams of Martial, commenting on them word by word and in great lexical detail, thereby keeping his discussion of words in a close relationship with unquestionably classical usage. Calepino fragmented, alphabetized, and augmented Perotti's work just as Giovanni Balbi had treated the *Derivationes* of Hugutio two centuries earlier (see Chapter 13). Like Balbi's, his dictionary was hugely influential, at first in a long series of

² For references to the controversy about its date, see references in Coates et al., *Catalogue of Books Printed in the Fifteenth Century*, II.361 (item B-010).

³ On printing and reference books, see, e.g., Blair, *Too Much to Know*, 13–14, 46–55.

⁴ Pade, 'Perotti's *Cornu copiae*'.

editions which remained close to its original form, and then in an even longer series of adaptations which introduced vernacular languages; I shall return to the polyglot Calepino tradition later in this chapter.⁵ In fact, Balbi and Calepino might be called the two fathers of the kind of free-standing alphabetized book of words which is still, for most readers in the Western world, the prototypical dictionary. Balbi was the pioneer of the full alphabetization of headwords as the most convenient macrostructure for a large wordlist, and Calepino was the pioneer of the alphabetized wordlist published as a free-standing book – and, moreover, as a free-standing book with a title which is a variant of *Dictionarium*.⁶

There was not the slightest possibility that Calepino would get everything in the *Dictionarium* right. Perotti's achievement was wonderful considering the resources available to him in the 1470s, but the *Cornucopiae* was neither a comprehensive treatment of the vocabulary of Latin nor an invariably trustworthy one, and Calepino's own reading in the editions of classical texts available to him in the late fifteenth century could only take him a certain distance beyond his principal source. When the French printer, humanist, and biblical scholar Robert Estienne the elder considered the possibility of publishing a new edition of Calepino in the 1520s, he rejected it, deciding that he would rather start afresh and make a new dictionary of his own.⁷

The result was published in 1531 as the *Linguae Latinae thesaurus* (¹1531, ²1536, ³1543; 'Treasury of the Latin language'). The *Thesaurus* did draw on the *Dictionarium*, but excluded post-classical Latin more rigorously than its predecessor had done. It was also much more richly illustrated with quotations from ancient authors than the *Dictionarium*, and they were more accurately referenced, and arranged with more elaborate attention to the phrases in which given words occurred. Whereas the structure of the *Dictionarium* had been alphabetical, the structure of the *Thesaurus* was derivational: headwords in alphabetical order (*lux* 'light' is before *luxo* 'dislocate, displace') were followed immediately by derivatives, even if these broke the alphabetical sequence (*lucerna* 'lamp' is after *lux* and before *luxo*). The first edition offered some equivalents in French, although the principal metalanguage was Latin, but some of these were discarded in the second edition and the remainder in the third; we shall see how Estienne went on to handle French and Latin after 1536 in the discussion of French lexicography below.

⁵ For the whole Calepino tradition, see Labarre, *Bibliographie*.

⁶ Considine, 'History of the concept of lexicography', 34–5.

⁷ Considine, *Dictionaries in Early Modern Europe*, 40–53.

The development which took place between the lexicographical work of Perotti and that of Robert Estienne can be seen from the example of the word *elegantia* 'elegance'. Perotti's main concern is to explain that *elegantia* was originally used pejoratively to describe excessive choosiness, but that by the time of Cicero *elegantia* could, in moderation, be seen as a good thing; much of this material is paraphrased without acknowledgement from the ancient author Aulus Gellius. Calepino gives his own definitions of *elegantia*, both in general and as applied to rhetoric, drawing on Perotti and hence indirectly on Gellius to enlarge on the first sense, and adapting an ancient text, the *Rhetorica ad Herennium*, to enlarge on the second. So, passages from two ancient authors are buried without acknowledgement in Calepino's entry. In 1531, Estienne cuts Calepino's entry to the bone, extracting the two key definitions and discarding the rest, and adds a clearly referenced passage from the ancient dramatist Plautus to illustrate the word *elegantia* in use; he also retrieves the passage from Gellius which was buried in Perotti's text, presenting it, clearly referenced, s.v. *elegans* 'elegant'. In 1536, he cuts editorial material again: the first of Calepino's definitions is adapted and two new quotations from classical Latin are added, while the second of Calepino's definitions is replaced by two quotations from the orator and philosopher Marcus Tullius Cicero, which make the point that the word can be applied to rhetoric. In 1543, Estienne adds ten more quotations and references, including the passage from the *Rhetorica ad Herennium* which was buried in Calepino's entry. Calepino's first definition still survives, altered but recognizable, at the beginning of Estienne's entry of 1543, but the entry as a whole has been transformed, from an explanation of the ancient word in which no ancient text is cited, to a structured presentation of fifteen ancient texts in which the word is used.

The 1531 edition of Estienne's *Thesaurus* was a folio of 940 leaves, and the 1543 edition ran to 1,584 leaves: in effect, two or three large volumes. No new classical Latin dictionary on this scale was published until the eighteenth century, for a great dictionary, like a great tree, casts a shadow in which it is difficult for rivals to grow. Large single-volume monolingual Latin dictionaries were, however, produced for various purposes by scholars across Europe. A particular publishing success, running to some seventy editions between 1535 and 1630, with scattered successors until 1820, was the dictionary, first published as *Observationes in M. T. Ciceronem*, of the Italian teacher and philosopher Marius Nizolius (Mario Nizzoli).⁸ This registered only the

⁸ Breen, 'Observationes'.

extensive vocabulary of Cicero, whose usage was regarded as irreproachable by many sixteenth-century readers seeking to identify the best classical Latin. The *Thesaurus eruditionis scholasticae* (1572) of the German pedagogue Basilius Faber (Schmidt), an attractively laid-out Latin dictionary (with some German equivalents) for the use of students, was on a similar scale and was also reprinted into the eighteenth century. A significant specialized dictionary was the posthumously published *Etymologicon linguae Latinae* (1662) of the Dutch humanist Gerardus Joannes Vossius. The most ample Latin dictionary of the seventeenth century, however, occupied new territory: this was the *Glossarium ad scriptores mediae et infimae Latinitatis* (1678) of the French medievalist and Byzantinist Charles du Fresne, sieur du Cange, a survey of the vocabulary of post-classical Latin in three folio volumes, rich in material drawn directly from manuscript sources. It was unsurpassed until the twentieth century, and still rewards consultation in the twenty-first.⁹

The central position of Latin in educational curricula across Europe required a huge output of Latin dictionaries for the classroom. A few of them will be mentioned below, in discussions of the lexicography of the Romance and Germanic languages; another, which was published in Cologne in 1513 as part of the school textbook *Pappia puerorum* of Joannes Murmellius of Deventer, will be mentioned in Chapter 21 because a Latin–German–Polish adaptation of 1528 was one of the first printed dictionaries to include Polish.

One series of Latin pedagogical dictionaries which calls for discussion here rather than in association with the lexicography of any one speech community is that produced by the Moravian theologian and writer on education Johannes Amos Comenius (Jan Amos Komenský).¹⁰ Its forerunner was the *Janua linguarum* (1611; ‘Gateway to languages’) published in Salamanca by the exiled Irish Jesuit William Bathe. This was, in its first edition, an introduction to Latin for speakers of Spanish, presenting more than 1,100 Latin phrases with Spanish equivalents, followed by an alphabetical Latin–Spanish vocabulary; editions in other languages followed. In 1631, Comenius published a *Janua linguarum reserata* (‘Gateway to languages opened’), consisting only of introductory matter and 1,000 numbered Latin sentences. These, unlike Bathe’s, were deployed in tight, informative groups: after the introduction came ‘On the beginning of the world’, ‘On the elements’, ‘On the firmament’, ‘On fire’, and so on, so that the *Janua* was a way into encyclopedic

⁹ Considine, *Dictionaries in Early Modern Europe*, 267–84.

¹⁰ Blekastad, *Comenius*, 170–6 and 200–3 (*Janua*), 534–6 (*Orbis*).

knowledge as well as Latinity. Within four months, an edition with English and French equivalents printed in parallel to the Latin sentences had appeared as *Porta linguarum trilinguis reserata*. A second monolingual Latin edition, published in the following year, already had an alphabetical index of the Latin words in the *Janua*, each keyed to the sentence in which it occurred, and before long, bilingual and multilingual editions were being published with Latin–vernacular wordlists after the sentences; there was a Latin–English one in 1636. A smaller text followed under the title of *Vestibulum* (‘Antechamber’ to the gate of learning), editions of which might also end with Latin–vernacular wordlists.¹¹ Finally, an illustrated adaptation, the *Orbis sensualium pictus*, appeared in 1658, a forerunner to the modern classed dictionaries which present vocabulary as captions to a series of pictures; the Latin–German first edition had many successors in German and other languages. By the end of the seventeenth century, Europe was flooded with Comenius-derived Latin dictionaries – limiting the count to those which included German, at least 142 editions had been published by 1700 – and the *Orbis sensualium pictus* continued to be republished as a schoolbook in central Europe until the end of the nineteenth century.¹²

Greek and the Other Learned Languages

The Greek–Latin lexicography of the Renaissance began with manuscript wordlists compiled in the fifteenth century by the first generations of Latin-using Europeans since the classical world who made a concerted attempt to learn ancient Greek.¹³ By 1478, an alphabetized Greek–Latin wordlist had been edited by the Italian monk Giovanni Crastoni and was available in print. Crastoni also prepared a Latin–Greek dictionary. A series of increasingly bulky alphabetized Greek–Latin dictionaries followed the first, and sold well; for an alternative to this tradition, the monolingual Greek *Magnum ac perutile dictionarium* (1523) of Favorino, based on Byzantine dictionaries, see Chapter 12. Successors to Crastoni’s Latin–Greek dictionary were less popular, for more people wanted to read Greek than to write it – moreover, by the

¹¹ See, for example, the tiny Latin–Hungarian wordlist in Comenius, *Januae linguae Latinae vestibulum* (1676), sig. G4r–v.

¹² For the German-language editions, see Jones, *German Lexicography*, items 309–451; for editions in all languages from 1700 onwards (excluding antiquarian and scholarly reprints), see Urbánková, *Soupis*, items 341–54 (*Janua*); 585–657 and 659–60 (*Orbis pictus*, items 657, 659, and 660 being editions of 1883, 1889, and 1896 from Hradec Králové (Königgratz), Warsaw, and Prague respectively); 831–45 (*Vestibulum*).

¹³ For these and their printed successors to the 1520s, see Botley, *Learning Greek*, 61–70.

1530s, new editions of Calepino gave the Greek equivalents of many of the Latin headwords, so they could be used as Latin–Greek dictionaries when one was wanted. A great mass of learned remarks on the Greek language was published as *Commentarii linguae Graecae* by the learned humanist Guillaume Budé in 1529, and material from this collection and perhaps from Budé's manuscript *Nachlass* entered the Greek dictionary tradition. A *Lexicon Graecolatinum* of 1554 advertising the use of unpublished notes of Budé's on its title page was published as a folio of nearly 1,500 pages, and was further enlarged in 1562 by the French physician Robert Constantin, whose name continued to appear on the title pages of Greek dictionaries into the seventeenth century.¹⁴

By this time, a new Greek–Latin dictionary was in preparation.¹⁵ The project had been begun in the 1550s by Robert Estienne, and was taken up after his death by his son Henri. It was published as *Thesaurus Graecae linguae* in 1572, with a volume of supplementary material in the following year. The *Thesaurus* was the product of an extraordinary depth of knowledge of ancient Greek. It was, as I have written elsewhere, 'the most comprehensive and sophisticated lexical record of any European language that had ever been published'.¹⁶ It would remain unsurpassed in this respect throughout the period discussed in this chapter, its only competitor being Robert Estienne's *Thesaurus* of 1543 (but much more classical Greek is extant than classical Latin): none of the living languages was the subject of a dictionary so wide-ranging and so richly documented. It was correspondingly huge: the main text ran to 4,208 folio pages, documenting some 64,000 words, and the alphabetical index ran to 864 pages in the supplementary volume. This index was necessary because the *Thesaurus Graecae linguae* had the same derivational structure as the *Linguae Latinae thesaurus*, so that a root word like *kalos* 'beautiful' would be followed by all the words which Estienne regarded as being derived from it, such as *philokalos* 'loving beauty' or *aikallō* 'I flatter'. Updated editions of the *Thesaurus* were produced in the nineteenth century, and these are still consulted.

The great bulk of the *Thesaurus* made it expensive and unwieldy, and there was an evident need for a one-volume abridgement: a dictionary on the same sort of scale as recent editions in the *Lexicon Graecolatinum* tradition, but enriched by Estienne's stupendous learning. This abridgement appeared as

¹⁴ Considine, *Dictionaries in Early Modern Europe*, 33–8; Reverdin, 'Figures d'hellénisme', 85–9.

¹⁵ Considine, *Dictionaries in Early Modern Europe*, 67–100.

¹⁶ Considine, *Dictionaries in Early Modern Europe*, 82.

a single folio volume, under the familiar title of *Lexicon Graecolatinum*, in 1580. Its macrostructure was, like that of the *Thesaurus*, derivational rather than fully alphabetical. It was compiled by one Johannes Scapula, who did not fully acknowledge his debt to the *Thesaurus*. Much worse, Estienne got none of the profits from Scapula's *Lexicon* to offset the very considerable expenses of publishing the larger work. A series of editions of the *Lexicon* was published over the next century, with isolated successors into the nineteenth century. An edition of 1652 was corrected by Cornelius Schrevelius, headmaster of the Gymnasium at Leiden, and two years later Schrevelius produced his own Greek–Latin and Latin–Greek *Lexicon manuale* on the basis of Scapula's, a tall octavo printed in double columns, with headwords in full alphabetical order. Many editions followed. It was generally understood throughout the period that Greek was taught to students who had already mastered Latin, which was therefore the metalanguage of Greek dictionaries, but some Greek–vernacular dictionaries were produced for people who lacked a formal education but wished to read the New Testament in the original Greek.¹⁷

Two other Greek–Latin dictionaries of the seventeenth century call for notice here. The first dictionary of New Testament Greek was published in 1619 by Georg Pasor, at the time headmaster of the Protestant academy in Herborn where Comenius had recently been a student (another teacher at Herborn, Zacharias Rosenbach, would produce a dictionary of the important Greek translation of the Old Testament a few years later).¹⁸ Since many Protestants in the seventeenth and eighteenth centuries were very interested in reading the New Testament in the original, and were much less interested in reading pagan classical Greek literature, this work was much reprinted, as was an abridgement of 1624. Indeed, Schrevelius advertised in the preface to the *Lexicon manuale* that he had drawn on Pasor's work: although his dictionary was primarily of the classical language, he wanted it to be clear that it could be used to read Scripture as well. A greater work of scholarship, but much less widely circulated, was the *Glossarium ad scriptores mediae et infimae Graecitatis* (1688) of Charles du Cange, which did for Byzantine Greek what his earlier dictionary had done for medieval Latin.¹⁹

The other learned languages studied in early modern Europe were those of the eastern Mediterranean, primarily Biblical Hebrew. Christian scholars had not attempted the lexicography of Hebrew before the sixteenth century, and

¹⁷ For Greek–English examples, see Considine, 'Lexicography of the learned languages', 352–3.

¹⁸ Delling, 'Erste griechisch–lateinisch Wörterbuch'.

¹⁹ Considine, *Dictionaries in Early Modern Europe*, 265–9.

they continued to benefit from the Jewish lexicographical tradition (for which see Chapter 9) thereafter. However, a wordlist which formed part of Johannes Reuchlin's *De rudimentis Hebraicis* ('On the rudiments of Hebrew', 1506) initiated a Christian tradition of Hebrew lexicography, in which the most important early modern figure was the German Protestant Johannes Buxtorf the elder, who had been educated at Herborn, and served for more than thirty years as professor of Hebrew at Basel.²⁰ He produced a series of dictionaries documenting Biblical Hebrew and Aramaic from 1600 onwards, among which his *Lexicon Hebraicum et Chaldaicum* of 1615 had the longest life, being reprinted more than twenty times in a series of editions ending in 1845. His masterpiece, *Lexicon Chaldaicum, talmudicum, et rabbinicum*, which handled post-Biblical Hebrew – Buxtorf was among the first Christian Hebraists to study rabbinic literature – was completed by his son Johannes Buxtorf the younger and published in 1639–40.²¹ Apart from Hebrew, substantial dictionaries of literary Arabic (by Franciscus Raphelengius, 1613, followed by Jacobus Golius, 1653); Ge'ez (by Jacobus Wemmers, 1638, followed by Hiob Ludolf, 1661); Coptic (1643, by Athanasius Kircher); and literary Persian (by Golius, 1669, followed by Ange de Saint Joseph, 1684; see Chapter 11) were produced, all with Latin as the metalanguage.²²

Polyglot Dictionaries

Bilingual wordlists for practical use had circulated in manuscript in late medieval Europe, and in 1477 the first of these was printed. It was an Italian–German classed vocabulary, published in Venice under the title *Introito e porta* ('Entrance and gate') by a German entrepreneur called Adam von Rottwil.²³ The place of publication was appropriate: German-speaking merchants had traded with Venice for a long time and had their own quarters on the Grand Canal, in a building on the site of the present Fondaco dei Tedeschi. The first nine editions, to 1501, were all Italian–German, and a few more bilingual editions appeared until the middle of the sixteenth century. From 1510 onwards, tetraglot editions adding French and Latin were published in Rome and Augsburg, and from 1532 tetraglot editions presenting

²⁰ Burnett, *From Christian Hebraism to Jewish Studies*, 120–33.

²¹ I owe material in this sentence to Aharon Maman.

²² For them, see respectively Toomer, *Eastern Wisdom*, 42–50; Smitskamp, *Philologia Orientalis*, item 219, and Fumagalli, *Bibliografia etiopica*, items 1184–8; Hamilton, *Copts and the West*, 195–239; Smitskamp, *Philologia Orientalis*, item 345.

²³ Rossebastiano Bart, *Antichi vocabolari*, 41–4 (this volume), 45–309 (subsequent tradition).

German, Latin, and Polish with Italian or French were published in Kraków. Pentaglot editions followed, the first of which, published in Venice in 1513, added Spanish to the languages of the 1510 tetraglot. A pentaglot published in the multilingual city of Antwerp in 1534 offered Italian, French, Spanish, Dutch, and Latin; a hexaglot tradition which ran from 1537 to 1636 added English to these five languages; heptaglots published in Antwerp from 1540 added German as well as English, and octoglots published in Paris and Lyon from 1546 to 1580 added ancient Greek. Other hexaglots published in Vienna shifted their linguistic focus eastwards, substituting Czech and Hungarian for Spanish and English. In all, eighty-nine editions were published and, although some of them were typeset by compositors who did not know all the languages which they were asked to handle, and were therefore decidedly inaccurate, these little dictionaries must have given thousands of early modern readers some acquaintance with vernaculars other than their own.

Another polyglot tradition began with a Latin–French–Dutch triglot published in Antwerp in 1495, with subsequent editions to 1530.²⁴ In 1527, it was substantially reworked by the schoolmaster Noel van Barlement of Antwerp, and was published as part of a French and Dutch bilingual text, preceded by some well-devised dialogues. Barlement's work ran to about 188 editions, many of them issued under a title beginning *Colloquia et dictionariolum*.²⁵ As in the *Introito e porta* tradition, the number of languages represented increased steadily: by 1551 Latin and Spanish had been added and, by 1575, Italian and German. Editions including English – hexaglot, heptaglot, and octoglot – are recorded from 1576 to 1692.²⁶

The dictionaries in the *Introito e porta* and Barlement traditions were generally oblong octavos, which could present up to eight narrow columns of text in parallel on a two-page opening. There were much bigger polyglot dictionaries as well, descended from the *Dictionarium* of Calepino.²⁷ The first of these was published in Antwerp in 1545, and added German, Dutch, and French equivalents to Calepino's original Latin and the Greek which had been added to it, to make a pentaglot. Almost simultaneously, a triglot Venice edition of 1545–6 added Italian to the earlier Latin and Greek. Just as further languages were added to new editions of the small polyglot dictionaries, so they were added to new polyglot dictionaries in the Calepino tradition. But whereas eight languages were really the most that could be set out in parallel on a single opening of an octavo, the polyglot Calepinos

²⁴ Claes, *Lijst van Nederlandse woordenlijsten*, items 37, 40, 81, 91, and 93.

²⁵ Pablo Núñez, *Arte de las palabras*, II.202–311. ²⁶ Alston, *Bibliography*, II, items 27–68.

²⁷ Overview, sorted by language, in Labarre, *Bibliographie*, 116.

were folios, and new equivalents were added to the text of each entry rather than being marshalled in parallel columns. As many as eleven languages were presented in a series of five editions from 1590 to 1627: Latin, Hebrew, Dutch, Greek, Spanish, French, Polish, Italian, Hungarian, English, and German. These huge books (the 1627 edition is a folio of 1,896 pages) have survived well in libraries.²⁸ Few readers can have used them to their full potential, but their compendiousness must always have been impressive. The last polyglot Calepinos, a series of seven-language editions published in Padua, appeared between 1718 and 1778.

The date range of the eleven-language Calepinos, 1590 to 1627, coincides roughly with the height of the early modern European polyglot dictionary traditions. As the seventeenth century went on, good compact bilingual dictionaries were produced in increasing numbers, offering more information than could the polyglots to the user who needed to manage one foreign language at a time. But in their day, the polyglots must have been in everyone's hands: if each edition of the Barlement *Colloquia et dictionarium* was published in a modest print run of 750, then there existed, at one time or another, more than 130,000 copies of this polyglot dictionary alone.

The Romance Languages

The first Renaissance dictionary of a European vernacular was a dictionary of Spanish, the Spanish–Latin *Dictionarium ex Hispaniensi in Latinum sermonem* (1495) of Elio Antonio de Nebrija, conventionally called his *Vocabulario*, to distinguish it from his Latin–Spanish *Lexicon: hoc est dictionarium ex sermone Latino in Hispaniense[m]* of 1492.²⁹ Both 'Renaissance dictionary' and 'dictionary of Spanish' must of course be qualified. Nebrija's *Vocabulario* was a Renaissance dictionary in so far as it was not in an unbroken line of development from the lexicography of the Middle Ages, but the difference between 'medieval' and 'Renaissance' in this respect was to some extent a matter of the lexicographer's self-perception and its consequences, not of a dramatic shift in the contents and methodology of the dictionary. And it was a dictionary of Spanish in so far as its headwords were Spanish, or more precisely Castilian; but its primary function was not so much to describe and analyse the vocabulary of Spanish as to make the vocabulary of Latin accessible to speakers of Spanish as they wrote and even spoke Latin.

²⁸ See Labarre, *Bibliographie*, item 186.

²⁹ Hamann, *Translations of Nebrija*, 11–15; the long titles given here are those which stand at the head of the first column of dictionary entries in each book.

But these qualifications do not make the *Vocabulario* any less of a landmark in European lexicography. Its 19,393 entries made it the most ample vernacular–Latin dictionary of fifteenth-century Europe. Taken in conjunction with the Latin–Spanish *Lexicon*, moreover, it was half of a bidirectional dictionary and, as such, it was exemplary. The Spanish-born humanist Juan Luis Vives wrote that such a dictionary should be compiled, after the example of Nebrija’s, for every vernacular, and his words were prefixed to the first bidirectional dictionary of Latin and German.³⁰ Many editions followed, throughout the sixteenth century and into the beginning of the seventeenth, and the *Vocabulario* was used as the basis of bilingual dictionaries of Spanish with South American and Mesoamerican languages, and with Tagalog (see Chapters 26, 27, and 29).³¹ The very first of these translations of Nebrija, Pedro de Alcalá’s Spanish–Arabic *Vocabulista Arauigo* of 1505, was the first substantial printed dictionary of two European vernaculars. Similarly, the first major dictionary of Portuguese, by Jerónimo Cardoso, which appeared in editions from 1562 onwards, was drawn on in the making of the first extant dictionary of Chinese and a European language, the Portuguese–Chinese dictionary of Matteo Ricci and Michele Ruggieri (see Chapters 15 and 29).

Only one major monolingual dictionary of Spanish was published in the two centuries after Nebrija, Sebastián de Covarrubias Horozco’s *Tesoro della lengua castellana o española* (1611). The *Tesoro* is a folio of some 700 pages (a supplement remained in manuscript until the twentieth century), presenting some 11,000 entries and 6,000 subentries, rich in encyclopedic and etymological material.³² As its bulk suggests more plainly than its entry count, it was, at the time of its publication, an unusually elaborate treatment of a European vernacular.

Many bilingual and multilingual dictionaries included Spanish.³³ In 1516, the Habsburg duke of Burgundy succeeded to the Spanish throne as Carlos I, making the Low Countries a Spanish territory, in which trilingual dictionaries of Spanish, French, and Dutch were called for; as we have seen, these three languages appeared together in polyglot dictionaries. A French–Spanish vocabulary appeared in 1565 as part of a French grammar for speakers of Spanish, and was succeeded by more elaborate dictionaries, notably César

³⁰ Dasypodius, *Dictionary* (edn 2), sig. π2r; the passage was already in edn 1, sig. π3v.

³¹ Hamann, *Translations of Nebrija*, 6–7, 43–84.

³² Overview in Spanish in Reyre, ‘Prólogo’, and in English in Considine, *Academy Dictionaries*, 111–12.

³³ Overview in Alvar Ezquerro, ‘Diccionarios del español’, 176–86; fuller treatment in Pablo Núñez, *Arte de las palabras*.

Oudin's *Tesoro de las dos lenguas francesca y española* of 1607, the standard dictionary in its field for the rest of the seventeenth century (an adaptation of 1609 added Italian and was also reprinted). An Italian–Spanish and Spanish–Italian dictionary, the *Vocabulario de las dos lenguas toscana y castellana* of Cristóbal de las Casas, ran to ten editions between 1570 and 1608; its successor, the *Vocabolario italiano e spagnolo* (1620) of Lorenzo Franciosini, was reprinted into the eighteenth century. Dictionaries of Spanish and English were produced in the mid sixteenth century, during the marriage of Felipe II of Spain and Mary I of England, and again from the end of the century onwards. The first Spanish–German dictionary, a derivative of Oudin, appeared only in 1670.

The lexicography of French at the beginning of the sixteenth century was largely a matter of Latin–French dictionaries in a late medieval tradition. As we have seen, Robert Estienne provided French equivalents in the first two editions of his *Latinae linguae thesaurus*. These provided a foundation for, and were succeeded by, a Latin–French *Dictionarium Latinogallicum* (1538 and subsequent editions), from which a French–Latin *Dictionnaire françoislatin* was produced in the following year.³⁴ Like other vernacular–Latin dictionaries created by reversing Latin–vernacular entries, it presented a number of multi-word vernacular lemmata which originated as the interpretamenta of single Latin headwords: for instance, at the very beginning of the dictionary, ‘Aage, AEtas’ (‘age’) is followed by subentries such as ‘Petit ou bas aage, AEtatula’ (‘early age’) and ‘L’aage de quatre ans, Quadrimatus’ (‘age of four years’). A long series of sixteenth- and early seventeenth-century dictionaries were descended from the *Dictionnaire françoislatin*, a notable member being the *Thresor de la langue francoyse* (¹1606, ²1621) compiled by the French diplomat Jean Nicot, the eponym of the word *nicotine*. The headwords of these dictionaries were of course French; Latin equivalents continued to be presented as late as the *Thresor*, but an increasing amount of French-language material was added to them. An alternative to their alphabetical order was offered by the subject-ordered *Indiculus universalis/L’univers en abregé* (1667) of François Pomey, much reprinted in several languages.

Meanwhile, the bilingual and multilingual lexicography of French developed much like that of Spanish.³⁵ There were small early dictionaries bringing French together with two Germanic languages, Dutch and German, spoken,

³⁴ For this and the ensuing traditions, see Wooldridge, *Les débuts*, esp. 18 (a diagrammatic overview).

³⁵ Overviews in Quemada, *Les dictionnaires*, 48–50, 567–78; for Italian, see Tancke, *Die italienischen Wörterbücher*, 78–82.

respectively, with French in the Low Countries and with French and Italian in the Swiss Confederacy (where a tradition of bidirectional dictionaries of French and Italian was initiated in 1584). As we have seen, there were dictionaries of French and Spanish from 1565 onwards. France was the nearest continental European neighbour of England (indeed, they shared a land frontier until the English loss of the Pale of Calais in 1558), and French wordlists were made for speakers of English from the fifteenth century onwards. Two of the most important, those of John Palsgrave (1530) and Randle Cotgrave (1611), are rich sources for the French language of their respective periods; they will be mentioned again below in their English context.

Early in the seventeenth century, a new monolingual dictionary of Italian would emerge as the most sophisticated vernacular dictionary in Europe, and its impact on French lexicography was so significant that we must turn to it at this point. There had been Latin–Italian and Italian–Latin dictionaries in the sixteenth century.³⁶ There had also been monolingual Italian dictionaries, which documented the vocabulary of Dante, Petrarch, and Boccaccio, the *Tre corone* ('Three crowns'): the three fourteenth-century Tuscan authors whose writings were seen not only as pre-eminent in Italian literature but also as providing a model for the literary Italian language. These two traditions were brought together in the *Vocabolario degli Accademici della Crusca* of 1612, an alphabetical dictionary of the Tuscan variety of Italian, produced by the members of one of the numerous academies of early modern Italy, the Accademia della Crusca of Florence.³⁷ It was a folio of 960 pages, presenting 24,595 lemmata with careful attention to the ways in which each word had been used in literature in Tuscan, which was quoted extensively, just as Latin literature had been quoted extensively in the *Latinae linguae thesaurus* of Robert Estienne and in subsequent dictionaries. The metalanguage was Italian, supplemented with Latin equivalents, which were presented briefly and unobtrusively. Although the writings of the *Tre corone* were quoted more than any others, the *Vocabolario* was not simply a dictionary of their usage, and indeed the tendency of the third edition (1691) was to represent modern authors, and technical vocabulary, more fully than the first edition had done. The only monolingual dictionary of a European language to compare with the *Vocabolario* at the time of its publication was the *Tesoro* of Covarrubias,

³⁶ See Tancke, *Die italienischen Wörterbücher*, 26–35 (monolingual) and 49–68 (bilingual with Latin).

³⁷ Considine, *Academy Dictionaries*, 9–27.

and the *Vocabolario* not only had a greater entry count but also, on account of its clear focus on the best literary usage, a much greater normative potential.

There was nothing like it in France. A French–Latin dictionary on the same physical scale as the *Vocabolario* was published in 1635, the *Invantaire des deus langues* (*sic*) of Philibert Monet, but its bilingual structure limited its ability to display the qualities of the French language. In the 1630s, a French language academy was constituted, and its members agreed to work co-operatively on the making of a new French dictionary.³⁸ Earlier dictionaries were of course used – the unseemly moment when one academician threw a copy of Nicot at another, who retaliated with a copy of Monet, shows that both books were on the table at their meetings.³⁹ But the *Dictionnaire de l'Académie française* which was eventually published in two folio volumes in 1694 was quite largely a new creation. Like the *Vocabolario degli Accademici della Crusca*, it presented a specialized, high-status language variety: but that of a metropolitan elite rather than of canonical literary authors. It was therefore illustrated with invented examples of good usage rather than with quotations from literary texts. These supported carefully worded definitions in French. The structure of the dictionary was derivational, like that of the largest Estienne dictionaries of Latin and Greek, and although this structure made the *Dictionnaire de l'Académie* more difficult to use (and was abandoned in later editions), it was part of a commitment to making the dictionary tell its readers as much as possible about how the vocabulary of French worked. The wordlist was quite restricted: there were 5,492 main entries and 13,269 subentries. However, these entries were supplemented by those of a *Dictionnaire des arts et des sciences*, edited by Thomas Corneille, a member of the academy, and published in two folio volumes later in 1694, which presented technical vocabulary, or technical senses of common words.

Two other major dictionaries of French appeared in the last quarter of the seventeenth century.⁴⁰ The first was the alphabetically ordered *Dictionnaire françois* of Pierre Richelet (1679–80), the first fully monolingual dictionary of French and the first dictionary of any European vernacular whose first edition was published in two volumes rather than one, a gesture of confidence in the dignity of French and the desirability of a good monoglot dictionary of the language. Richelet acknowledged that the *Vocabolario degli Accademici della Crusca* had been an inspiration for his work, which was illustrated with quotations from French literature. The second, also alphabetically ordered,

³⁸ Considine, *Academy Dictionaries*, 28–41, 51–61. ³⁹ Considine, *Academy Dictionaries*, 38.

⁴⁰ Considine, *Academy Dictionaries*, 41–50.

and registering some 40,000 entries to Richelet's 25,000, with noteworthy encyclopedic content, was the three-volume *Dictionnaire universel* (1690) of the renegade academicien Antoine Furetière. A publishing market in which three multi-volume monolingual dictionaries of the same language could compete was something new in the history of European lexicography.

The Germanic Languages

None of the Germanic languages were registered in dictionaries as innovative as Nebrija's until the 1530s. German and Dutch were served by a strong production of printed vernacular–Latin dictionaries from the late fifteenth century onwards, but these formed an unbroken tradition with the vernacular–Latin lexicography of the late fourteenth and early fifteenth centuries.⁴¹ English was likewise served by one printed vernacular–Latin dictionary, the *Promptorium parvulorum*, from 1499 into the 1530s, but this was essentially a work of the first half of the fifteenth century. From the 1530s onwards, however, dictionaries for readers of English were produced in considerable numbers.⁴² Their story will be told next, and then those of the dictionaries of German; Dutch; and the Scandinavian languages.

The wordlist of the *Promptorium* was updated and supplemented in the English–French *Lesclarcissement de la langue francoyse* of John Palsgrave (1530), a large grammatical compendium which included several wordlists classified by part of speech, the list of verbs being particularly long and sophisticated. Much more influential was the Latin–English *Dictionary* of Sir Thomas Elyot (1538 and further editions), basically an adaptation of Calepino's *Dictionarium*, which was revised and enlarged several times in the next three decades, with attention to dictionaries in the Estienne tradition, a process resulting in the Latin–English *Thesaurus linguae Romanae et Britannicae* of Thomas Cooper (1565).⁴³ Neither Elyot nor Cooper reversed their Latin–vernacular dictionaries as Nebrija had done, and English–Latin dictionaries indebted to the Latin–English tradition became available only from the 1550s onwards. The late sixteenth and seventeenth centuries saw no major new dictionaries of Latin and English, but a robust production of Latin–English and English–

⁴¹ Bibliographical overviews in Claes, *Bibliographisches Verzeichnis* (for German), and Claes, *Lijst van Nederlandse woordenlijsten* (for Dutch); pre-1501 dictionaries are items 1–147 and 1–45 respectively.

⁴² The best overviews are now Ashgate *Critical Essays*, vols. III (sixteenth century) and IV (seventeenth century); Stein, *English Dictionary Before Cawdrey* is now dated, but useful for the sixteenth century.

⁴³ Starnes, *Renaissance Dictionaries*, 45–110.

Latin dictionaries of different sizes, with some particularly interesting work being done in the area of phraseology.⁴⁴

Palsgrave's *Lesclarcissement* was something of an anomaly in the story of the early modern bilingual lexicography of English and the vernaculars: it was very large, and its lexicographical material was presented in several separate wordlists interspersed with grammatical information (which comprised the majority of the book). The first free-standing printed dictionary of English and another vernacular was the *Dictionary in Englyshe and Welshe* (1547) of the Bible translator William Salesbury. The other vernaculars to which printed sixteenth-century bilingual dictionaries gave English-speakers access were the three prestigious Romance languages, namely Spanish, Italian, and French. Russian–English wordlists totalling some 6,000 entries were made by Mark Ridley, chief physician to the tsar of Russia until 1598, and appear to be the most extensive of all lexicographical records of sixteenth-century Russian, but they remained in manuscript.⁴⁵ By far the most distinguished of the sixteenth-century bilingual dictionaries made for English-speakers was the Italian–English *Worlde of Wordes* (1598) by the teacher and translator John Florio, in which the English equivalents show a strong, witty, poetic grasp of the language at a time when its literary potential was being developed and expanded by Shakespeare and others.

Florio's dictionary was published in a greatly expanded second edition in 1611, and this year was an *annus mirabilis* for English lexicography because it also saw the publication of Randle Cotgrave's French–English *Dictionarie of the French and English Tongues*, the English equivalents in which show a similar inventiveness to Florio's. The bilingual English dictionaries of the rest of the century are less spectacular, though an ambitious polyglot, the *Ductor in linguas* (¹1617, ²1625) of John Minsheu, who had previously compiled a Spanish–English dictionary, deserves notice, as do editions of Florio and Cotgrave with, respectively, English–Italian and English–French second parts; further manuscript wordlists of Russian; the beginnings of a tradition of bilingual dictionaries with Dutch; and, towards the end of the century, a new generation of French dictionaries succeeding the Cotgrave tradition.

The first wordlists in which English headwords were explained in English were glossaries of technical vocabulary in longer books.⁴⁶ John Rastell's

⁴⁴ Starnes, *Renaissance Dictionaries*, 111–324; for the phraseological works, see Knappe, *Idioms*, 163–82.

⁴⁵ Ridley, *Dictionarie* (1996), is an edition.

⁴⁶ Schäfer, *Early Modern English Lexicography*, vol. I, covers the period to 1640; thereafter, overview in *Ashgate Critical Essays*, IV.xxx–xxxix.

Exposiciones terminorum legum Anglorum (c. 1524), a dictionary of legal terms, has been called 'the first printed stand-alone monolingual English lexicon'.⁴⁷ In fact, it is obviously not monolingual, since it presents Law French and English texts in parallel columns, making it clear that the lexical items which it registers (including a number of Latin words and phrases) belong equally to either language: 'Chemyn est le haut voy ou chescun home passa . . . Chemyn is ye hye way where euery man goth.' There certainly was a tradition of legal lexicography in sixteenth- and seventeenth-century England, leading to several large seventeenth-century dictionaries, and there were also specialized dictionaries of, for instance, heraldry and medical terminology. Medieval English was documented in glossaries (notably to the works of Chaucer) and in free-standing wordlists of Old English, the first to be printed being the *Dictionarium Saxonico-Latino-Anglicum* of William Somner (1659). A first English etymological dictionary, Stephen Skinner's *Etymologicum linguae Anglicanae*, was published in Latin (1671) and in an abridged English translation by one Richard Hogarth (*Gazophylacium Anglicanum*, 1689).

The seventeenth century saw a tradition of six free-standing monolingual dictionaries of English.⁴⁸ The first, Robert Cawdrey's *Table Alphabeticall* (1604), is conventionally called the first English dictionary, and in this case convention is right: there were already more English headwords in the *Promptorium parvulorum*; some headwords in the *Exposiciones terminorum legum anglorum* were in use as English words and were explained in English in Rastell's *Exposiciones*; but Cawdrey's was the first book devoted to presenting a list of English words with English explanations. These were, as Cawdrey's title page put it, 'hard vsuall English wordes, borrowed from the Hebrew, Greeke, Latine, or French. &c. With the interpretation thereof by plaine English words'. The next four dictionaries in the hard-word tradition showed steadily increasing headword counts, and also increased in physical bulk: Cawdrey's *Table* is a pocket-sized octavo, whereas the fifth dictionary in the tradition, Edward Phillips' *New World of English Words* (1658), is a folio, designed to join a shelf of reference books. By the end of the seventeenth century, the tradition was losing its focus on hard words, and its eighteenth-century successors (for which see Chapter 23) would register the general vocabulary of English.

⁴⁷ Lancashire, 'Perils of firsts', 243.

⁴⁸ Overview in *Ashgate Critical Essays*, IV.xxi–xxix; classic but dated account in Starnes and Noyes, *English Dictionary from Cawdrey to Johnson*, 13–63, with correctives in Miyoshi, *First Century*, 1–102.

The most extraordinary English dictionary of the seventeenth century did not belong to the hard-word tradition. It was the 'Alphabetical dictionary' compiled by William Lloyd as an appendix to John Wilkins' *Essay Towards a Real Character and a Philosophical Language* (1668), which gave equivalents for English words in the terms of an artificial language devised to correspond to a wonderfully methodical semantic analysis of all human knowledge.⁴⁹ For instance, the verb *read* was defined as 'AC. III. 4', meaning that its place in the analysis was as a corporeal action, of the third class ('actions peculiar to men') and the fourth subclass (reading and spelling), while the adjective *ready* was defined by cross-references to words whose places in the analysis had been stated at their own entries in the dictionary, 'Present; Soon, Near; Easie; Willing, Inclined', and so on. This goes far beyond the encyclopedism of the *Janua* of Comenius. Never before, perhaps, had a dictionary of any language been founded on such a systematic inquiry into the meanings of words.

The strong late medieval tradition of Latin–German and (to a lesser extent) German–Latin lexicography may account for the slow development of a new generation of humanistic dictionaries in German-speaking Europe.⁵⁰ The first was the Latin–German *Dictionarium* (1535) of Petrus Dasypodius, a teacher in Strasburg. It was based on Calepino's *Dictionarium*, with derivational ordering inspired by that of Robert Estienne's *Thesaurus*, but stricter: for instance, Estienne treats *cognatus* 'a relative' between *cogito* and *cognomen*, whereas Dasypodius moves it to his entry for *nascor*, of which *natus* is the past participle.⁵¹ A second edition, with a German–Latin section reversing the Latin–German wordlist, appeared in 1536, a stout quarto of nearly a thousand pages. Nearly fifty editions followed the first, and it was an influence on later dictionaries of German, and also of Dutch, Czech, and Hungarian.⁵² Another Latin–German dictionary, the *Dictionarium Latino–Germanicum* (1541) of Petrus Cholinus and Johannes Frisius, was based on the *Dictionarium Latinogallicum* of Robert Estienne (Frisius also produced a very successful bidirectional dictionary of German and Latin based on a smaller work in the Estienne tradition); its entries were reversed to make the first free-standing German–Latin dictionary, *Die Teütsch Spraach* of Josua Maaler (1561). A foreword to *Die Teütsch Spraach* by the naturalist and collector of languages

⁴⁹ Knappe, 'Theory meets empiricism', is a good introduction to the 'Alphabetical dictionary', with references to earlier work.

⁵⁰ For an overview in English from Dasypodius to the end of the seventeenth century, see Jones, 'Lingua Teutonum victrix?'; bibliographies are Claes, *Bibliographisches Verzeichnis* (sixteenth century), and Jones, *German Lexicography* (seventeenth century).

⁵¹ For Dasypodius, Calepino, and Estienne, see Müller, *Deutsche Lexikographie*, 66.

⁵² Müller, *Deutsche Lexikographie*, 69–70.

Conrad Gessner remarked that the French, Italians, and English had good dictionaries, and that speakers of German should have one as well; Dasypodius had made a similar remark, not mentioning English, back in 1535.⁵³

Dutch and Low German were rightly seen as forming a continuum at the beginning of the sixteenth century. So, for example, a printed dictionary of 1477, the *Vocabularius qui intitulatur Teuthonista vulgariter dicendo der Duytschlender* ('Vocabulary which is called *Teutonista* or, in the vernacular, *The Deutschländer*') of Gerard van der Schueren, presents the Germanic language variety of the Duchy of Cleves on the lower Rhine, and appears in the standard bibliographies of pre-1600 Dutch and German dictionaries alike.⁵⁴ Likewise, a Dutch adaptation of Dasypodius' dictionary was published in Antwerp in 1542 with its title unaltered: *Dictionarium Latinogermanicum* could, as far as the publisher was concerned, apply to what we would call a Dutch dictionary just as well as to what we would call a German one.⁵⁵

But in 1546 a first wordlist treated Dutch as distinctive; this was the Dutch–French *Naembouck van allen naturelicken, ende ongheschuumden vlaemschen worden* ('Nomenclator of all the native, and not borrowed, Dutch words') published by the printer and typefounder Joos Lambrecht of Gent.⁵⁶ From 1562 onwards, the great printer Christophe Plantin of Antwerp published a series of impressive dictionaries, on all of which his employee Cornelis Kiliaan (Kilianus, van Kiel) worked: a Latin–Greek–French–Dutch *Dictionarium tetraglotton* based ultimately on Robert Estienne's *Dictionarium Latinogallicum*; a Dutch–French–Latin *Thesaurus Theutonicae linguae* (1573); and a Dutch–Latin dictionary with Kiliaan named as editor, the *Dictionarium Teutonico–Latinum* (¹1574, ²1588). The third edition of Kiliaan's dictionary, with 40,000 entries, was published in 1599 as *Etymologicum Teutonicae linguae*. The *Etymologicum* was, when it was published, the most sophisticated dictionary of a Germanic language, rich in information about the Dutch dialects and indeed other Germanic language varieties, and responsive to the sixteenth-century growth of the vocabulary of Dutch. It became the standard Dutch dictionary of the seventeenth century.

⁵³ For *Die Teütsch Sprach* and its preface, see Considine, *Dictionaries in Early Modern Europe*, 130–5; Dasypodius' remark is in *Dictionarium* (edn 1), sig. π2v.

⁵⁴ Claes, *Lijst van Nederlandse woordenlijsten*, item 1; Claes, *Bibliographisches Verzeichnis*, item 8.

⁵⁵ Müller, *Deutsche Lexikographie*, 70.

⁵⁶ Overview of Dutch lexicography from the Middle Ages to Kiliaan in Van Rossem, *Portret van een woordenaar*, 158–76, 180, 186–91.

From Gessner's day to the end of the seventeenth century, the lexicography of German developed slowly. Georg Henisch, born in a German-speaking town in what is now Slovakia, and for forty years a teacher in Augsburg, published the first volume of a splendid German–Latin dictionary, rich in phraseological and etymological information and in subtle discriminations of sense, under the title *Teütsche Sprach vnd Weißheit* ('German language and wisdom') in 1616, and this would have taken German lexicography to a new level if it had been completed, but Henisch died two years later, leaving the dictionary at the letter G.⁵⁷

Meanwhile, a German-language academy, the Fruchtbringende Gesellschaft, had been founded under the auspices of Prince Ludwig of Anhalt, inspired by the Accademia della Crusca, of which Ludwig was a member.⁵⁸ Dictionary projects were undertaken by members of the Gesellschaft, but they never resulted in a full-scale German dictionary. A characteristic of these projects was a strong interest in the derivational processes of German – still a concern for lexicographers of German (see Chapter 22) – leading to an interest in the isolation of the basic forms, *Stammwörter*, from which the rest of the vocabulary of German was derived. Hence, a list of 4,884 *Stammwörter* forms part of a major publication by a member of the Gesellschaft, Justus Georgius Schottelius' *Ausführliche Arbeit von der teutschen Hauptsprache* (1663; 'Comprehensive work on the German cardinal language').⁵⁹ The one major dictionary to be closely associated with the Fruchtbringende Gesellschaft, and the last major German-language dictionary to be completed in the seventeenth century, was Kaspar Stieler's *Der Teutschen Sprache Stammbaum und Fortwachs, oder Teutscher Sprachschatz* ('The stock and ramifications of the German language, or German language treasury') of 1691, which registers 68,000 German words in 2,000 quarto pages. Although the metalanguage is Latin, the object of the dictionary is to present the vocabulary of German in a derivational structure, with an extensive index and a grammar. It was, by the end of the century, the most extensive printed dictionary of any Germanic language.

Printing in the vernacular came to the two principal Scandinavian kingdoms, Denmark and Sweden, in the 1490s, and their production of printed dictionaries was likewise a little belated. Short classed vocabularies and dictionaries of synonyms emerged early in the sixteenth century – Latin–Danish from 1510, Latin–Swedish from 1538 – and were followed by more

⁵⁷ Considine, *Dictionaries in Early Modern Europe*, 135–8.

⁵⁸ Overview of the Gesellschaft and lexicography in Considine, *Academy Dictionaries*, 73–9.

⁵⁹ See McLelland, J. G. *Schottelius's Ausführliche Arbeit*; for *Hauptsprache*, see esp. 35–8.

substantial dictionaries for school use, including a Swedish version of Comenius' *Janua linguarum* and a Danish version of his *Orbis sensualium pictus*.⁶⁰

Two interesting new developments took place in the Kingdom of Sweden in the seventeenth century. The first was the production from 1637 onwards of small dictionaries of Latin, Swedish, and Finnish. The last of these languages was of practical interest to Swedish lexicographers because what is now Finland was under Swedish rule at the time, but it was also of theoretical interest because Finnish was so different from Swedish (in modern terms, it is not an Indo-European language), and material from a Latin–Swedish–Finnish dictionary was used by the seventeenth-century polymath Martin Fogelius in his discovery of the relationship between Finnish and Hungarian.⁶¹ The second development in Swedish lexicography was a turn to the past, originally driven by a patriotic interest in Swedish history, which led to the production of dictionaries of early Scandinavian material and of the extinct Gothic language, which was thought to have been spoken by ancestors of the Swedes. This lexicographical activity led the Swedish poet and philologist Georg Stiernhielm to project a visionary etymological dictionary of the languages of the world, in which Swedish would have been given a prominent place.⁶²

The lexicography of early Scandinavian was also carried out in the Kingdom of Denmark, which included Iceland, leading to the first dictionaries of Icelandic, the conservatism of the language being such that these brought together material from medieval literature and from the contemporary spoken language.⁶³ But the most remarkable lexicographical work undertaken in the kingdom before 1700 was directed towards a comprehensive dictionary of modern Danish, of which the final redaction runs to twenty-two volumes in manuscript.⁶⁴ Its editor, Matthias Moth, was reluctant or unable to have it printed, and the project archive passed after his death into the Danish royal library, where it remains. No other early modern dictionary of any European language pays such ample attention to spoken and regional usage. There is some Norwegian material in Moth's archive, Norway being, like Iceland, part of the Kingdom of Denmark; a Norwegian dictionary was also printed in 1646.⁶⁵

⁶⁰ Ottosson, 'Data basis', 44–6; Haugen, 'Introduction', 5–9.

⁶¹ Considine, *Small Dictionaries*, 167–8.

⁶² Considine, *Dictionaries in Early Modern Europe*, 240–9, 307–10.

⁶³ Gunnlaugur Ingólfsson, 'Første trykte islandske ordbøger'.

⁶⁴ Considine, *Academy Dictionaries*, 82–92. ⁶⁵ Haugen, 'Introduction', 6.

Other European Languages

A number of Romance and Germanic languages are missing from the discussion above, because they were the subjects of fewer wordlists or none at all: Catalan and Occitan, Romansh and Romanian; Norn, Frisian, and Scots. The list could go on.

Indeed, there are whole language families which have barely been touched on above. To take one example from one of these families, Irish, which as we saw in Chapter 13 was the subject of one of the oldest lexicographical traditions in Europe, continued to be the subject of manuscript wordlists such as the one called O'Davoren's Glossary, which treats hard words, especially from legal texts, and survives in two sixteenth-century manuscripts.⁶⁶ A fully alphabetized monolingual Irish dictionary, the *Foclóir nó sanasán nua* of the Franciscan Míchél Ó Cléirigh (hence O'Clery's Glossary in English), was printed at Leuven, the compiler's place of exile, in 1643. Other Celtic languages; the languages belonging to the great Slavic and Baltic families; Finnish and Estonian; Hungarian; Basque; Albanian; modern Greek; and other languages of Europe such as Romani must likewise be passed over here.⁶⁷

Conclusion

Selective as this account of early modern European lexicography has been, it gives some sense of the diversity of the dictionaries in question. The largest were published in thousands of folio pages, and the smallest were little pocket-books; some were arranged alphabetically, some derivationally, and some by subject. On the other hand, they had significant features in common: quite apart from their shared format as printed codices, and the shared filiations which linked many of them, they all shared a sense that collecting the words of a given language in an inventory, and explaining them in terms of other words, whether in the same language or another, was a useful activity, which might even raise the prestige of the language in question. Lexicography might be a patriotic activity, as it was for the French Academy or for Stieler or Moth; this possibility would be further explored in some of the dictionaries of the eighteenth century.

⁶⁶ Breatnach, *Companion to the Corpus iuris Hibernici*, 100–8.

⁶⁷ All of these languages are touched on, some of them briefly, in Considine, *Academy Dictionaries*.

On a more pragmatic note, the making of dictionaries which described words might be difficult to distinguish from the making of encyclopedias which described things, as we see in cases like those of Comenius or Covarrubias or Furetière, and this too would be a question for lexicographers in the eighteenth century and beyond.

PART III



THE MODERN WORLD:
CONTINUING TRADITIONS

China from c. 1700

HENNING KLÖTER

Long before the seventeenth century, two major types of lexicographical macro-arrangement had evolved in China (see Chapters 3 and 6). One type was based on the graphical components of individual characters, regardless of the character reading. Another type was according to the reading of characters, regardless of graphical features. This complementary division into sound-based and shape-based arrangements was continued far into the twentieth century, and became obsolete only with the recent advent of digital lexicography.

Despite the continuation of long-standing traditions after the seventeenth century, some major changes in Chinese lexicography can be identified for the period analysed in this chapter. First, beginning in the seventeenth century, monolingual Chinese dictionaries played an increasingly important role in the dissemination of linguistic standards. Secondly, beginning in the late sixteenth century, Chinese–Western contacts (for which see also Chapter 29) induced new lexicographic practices. Western missionaries introduced not only alphabetic writing, but also new ideas concerning the selection of source and target languages. As a consequence, lexicography gradually developed towards more diversity with regard to the Chinese and Western languages documented in dictionaries. At the same time, alphabetically arranged dictionaries dominated the newly evolving field of bilingual Western–Chinese lexicography. It was not until the twentieth century that Western and Chinese modes of lexicographic arrangement merged into a new hybrid. Thirdly, beginning in the nineteenth century, the Chinese language began to undergo major changes. New words were coined, many of which were Chinese translations of Western scientific and technical terms. In the early twentieth century, a northern Chinese vernacular language became the new standard of literary composition. These lexical and sociolinguistic changes naturally posed new challenges to Chinese lexicographers. After the late nineteenth century, however, there was, at least to some degree,

a mismatch between rapid linguistic changes and the continuation of conservative traditions in lexicography. It was only after the foundation of the People's Republic of China in 1949 that different strands of lexicographic practice converged into a new quasi-standard of dictionary arrangement. Since the second half of the twentieth century, standard dictionaries have been dictionaries of the modern language (spoken and written) and have followed an alphabetical macro-arrangement.

Preliminaries: Which Chinese?

In Chinese terminology, most modern Chinese dictionaries are dictionaries of Hànyǔ, in other words, the language of China's biggest ethnic group, known as Hànn Chinese or Hànn people. From a linguistic perspective, however, the Chinese label *Hànyǔ* in the title of a dictionary is as ambiguous as the English word *Chinese*, since both can refer to distinct Sinitic languages which need to be distinguished carefully. As we shall see, one of the major lexicographic developments after the seventeenth century is the development towards diversity in terms of the documented language. A Chinese dictionary of this period can be the dictionary of a standard language of a particular period, a dictionary of the spoken vernacular of a certain region, or a dictionary of the language of a particular social group.

Many pre-modern Chinese dictionaries document the classical literary language known as Classical Chinese or Literary Chinese (*wényán wén*). In imperial times, this language was purely used as a written medium, and it remains controversial whether it was ever based on a spoken language.¹ Classical Chinese is the language of the canonical works of the Confucian tradition; it was also used as a language of administration in written documents of the imperial governments. As a written literary language, Classical Chinese needs to be distinguished from Mandarin as a spoken language – and *Mandarin* itself can refer to at least four distinct linguistic entities: (1) the *lingua franca* of imperial officials before the twentieth century, in Chinese known as *guānhuà* ('officials' language'); (2) the national language (*guóyǔ*) of the Republic of China which was founded in 1912; (3) the standard or 'common language' (*pǔtōnghuà*) of the People's Republic of China founded in 1949; (4) the Mandarin dialect group (*guānhuà fāngyán*), including local Sinitic varieties spoken in the north-east, the north-west, and the south-west.²

¹ See Mair, 'Buddhism and the rise of the written vernacular', 708; Norman, *Chinese*, 83.

² Simmons, 'Mandarin, varieties of', 672.

Aside from the group of Mandarin dialects, there are also various regional vernaculars belonging to distinct dialect groups. In Chinese terminology, these vernaculars are defined as topolects (*fāngyán*) of the Hànn language. It has to be emphasized that these topolects are mutually unintelligible regional vernaculars, which could justifiably be treated as individual languages. As S. R. Ramsey has pointed out, the Chinese emphasis on linguistic unity is largely due to cultural considerations:

Unlike the peoples who speak Romance, the Chinese are not divided into a number of national units corresponding roughly to the several groups of closely allied dialects. Rather, the Chinese language is spoken by a single group of people with a common cultural heritage. China is not only the most populous country on earth, it is also the oldest social institution, and the Chinese people belong to and follow cultural and national traditions that have continued since the days of the Han empire and before.³

Against this backdrop, it is not surprising that the rise of modern Chinese dialect lexicography in the nineteenth century was largely due to Western influence. Until the middle of the nineteenth century, China had been largely closed to outsiders. As a consequence of the first Opium War and the Treaty of Nanking (1842), Hong Kong Island became a British crown colony, and the eastern and south-eastern port cities of Shànghǎi, Níngbō, Fúzhōu, Xiàmén, and Guǎngzhōu were opened to foreign traders, diplomats, and missionaries. Since these cities are located in four of the seven major dialect areas, the enforced opening of China led to a new interest in the regional languages of the south-east, most importantly those of the port cities of Shànghǎi, Xiàmén (Amoy), and Guǎngzhōu (Canton), as well as their surrounding areas.⁴ In addition to these vernaculars, known as Mǐn (Fúzhōu, Xiàmén), Wú (Shànghǎi, Níngbō), and Yuè (Guǎngzhōu), missionary topolect lexicography also included Hakka (Kèjiā) dialects.⁵ In contrast to the aforementioned varieties, Hakka is not associated with one major urban centre.

Western missionaries and diplomats thus made significant contributions to the lexicographic documentation of these vernaculars. From a sociolinguistic perspective, the significance of this contribution cannot be overestimated. In the light of the aforementioned ambiguity of terms such as ‘Han language’ or ‘Chinese’, they laid the foundation for a thorough disambiguation in regional terms. It must be emphasized that the recorded vernaculars were not minor

³ Ramsey, *Languages of China*, 16–17.

⁴ For an overview of the dialect areas, see Norman, ‘Dialect classification’.

⁵ For details, see Kwok, T’sou, and Chin, ‘Yuè 粵 dialects’; Lau, ‘Hakka dialects’; Lien, ‘Mǐn 閩 dialects’; You, ‘Wú 吳 dialects’.

village dialects, but rather regional languages with a substantial number of speakers and, at least in the cases of Mǐn and Cantonese, with distinct histories of vernacular literature.

Monolingual Lexicography Before 1949

The publication of the *Kāngxī zìdiǎn* 康熙字典 in 1716 marks a milestone in the history of Chinese lexicography. Ironically, despite its indisputable importance, the dictionary itself did not produce any significant lexicographical innovations. Its authority derives from both its unprecedented voluminous size and its direct support by the imperial court, notably the Kāngxī emperor. A graphic arrangement of dictionaries according to components known as ‘classifiers’ (*bùshǒu*) had been invented some 1,600 years before (see Chapters 3 and 6). The *Kāngxī zìdiǎn*’s specific list of 214 classifiers – known as ‘Kangxi classifiers’ – remained authoritative into the twentieth century. Yet it is often neglected that this list had first been introduced in a dictionary some 100 years earlier.⁶ The entries in the *Kāngxī zìdiǎn* comprise 47,043 single characters, complemented by semantic definitions, sound glosses, character variants, and textual evidence. There is significant intertextual overlap between the *Kāngxī zìdiǎn* and earlier dictionaries, and Marc Winter describes it as ‘entirely a compilation of entries in earlier dictionaries rather than a newly written lexicographical work’.⁷

Whereas the *Kāngxī zìdiǎn* produced no lexicographic innovations in a narrow sense, its overall significance lies in the prescriptive role ascribed to it by official authorities. From the publication of the *Kāngxī zìdiǎn* until today, Chinese monolingual dictionaries have played a crucial role in standardization processes: they have identified areas of standardization such as pronunciation and graphic arrangement of characters, sources legitimizing linguistic correctness, and the linguistic norms themselves. In the case of the *Kāngxī zìdiǎn*, this role is neatly reflected in its full title, *Yùdìng Kāngxī zìdiǎn*, ‘Imperially commissioned compendium of standard characters from the Kāngxī period’.⁸ As Endymion Wilkinson writes, the label *zìdiǎn* in the title ‘was chosen by the emperor to suggest that the new work contained correct and authoritative forms, pronunciations, and definitions of the characters’; it was only during the nineteenth century that the word *zìdiǎn* came to be used in the modern sense of ‘dictionary’.⁹ It is noteworthy that the *Kāngxī zìdiǎn*

⁶ See Winter, ‘Kāngxī zìdiǎn 康熙字典’. ⁷ Winter, ‘Kāngxī zìdiǎn 康熙字典’, 481.

⁸ The translation follows Winter, ‘Kāngxī zìdiǎn 康熙字典’.

⁹ Wilkinson, *Chinese History*, 80.

defines correctness with regard to individual written graphs, that is, the number and arrangement of strokes constituting a graph. Of the 47,043 characters listed in the *Kāngxī zìdiǎn*, some 20,000 are graphic variants.¹⁰ Entry information has a strict historical orientation: sound glosses cite historical rhyme dictionaries dating back to the eighth century, and textual evidence almost exclusively cites phrases from canonical Confucian texts.¹¹ From this we can also infer what kind of linguistic information is not documented in the *Kāngxī zìdiǎn*: contemporary spoken language is entirely missing, in terms of both (standard) pronunciation and lexicon.

Political and social changes at the beginning of the twentieth century paved the way for new lexicographical developments. Dictionaries played a major role in language planning and the implementation of new language standards. In addition, lexical changes were followed, albeit with some delay, by new forms of lexicographic arrangement and entry information.

Before the end of the last imperial dynasty in 1911, no systematic attempts had been made to standardize Mandarin (*guānhuà*) in terms of pronunciation, lexicon, and grammar. This changed after the foundation of the republic in 1912, when systematic efforts were made to codify a new national language (*guóyǔ*).¹² These efforts came along with the promulgation of new devices of phonetic notation, notably new systems of Romanization and a non-Latin system of phonetic symbols known as *zhùyīn zìmǔ* (Mandarin Phonetic Symbols, now known as *zhùyīn fúhào*).¹³ Mandarin Phonetic Symbols are based on components of Chinese characters representing sounds, the first four being ㄅ for [b], ㄆ for [p^h], ㄇ for [m], and ㄈ for [f]. The appropriation of sound-based scripts paved the way for new modes of lexicographic arrangement, as will be explained presently. These developments took place at a time when language in general had become a core target of nationalist and leftist ideologies. The young republic witnessed the rise of major independent publishing houses which were to play important roles in the development of modern lexicography in China, notably Commercial Press (*Shāngwù Yīnshūguǎn*), established in 1904, and Zhonghua Book Company (*Zhōnghuá shūjù*), founded in 1912.¹⁴

During this period of sociolinguistic transformation, monolingual lexicography developed towards diversification in terms of documented language,

¹⁰ Wilkinson, *Chinese History*, 80. ¹¹ Winter, 'Kāngxī zìdiǎn 康熙字典', 483.

¹² For details, see Kaske, 'National Language Movement'.

¹³ For details, see Simmons, 'Transcription systems: Gwoyeu Romatzyh 國語羅馬字'; Wippermann, 'Transcription systems: Zhùyīn fúhào 注音符號'.

¹⁴ Reed, *Gutenberg in Shanghai*, 212–40.

lexicographic arrangement, and entry information. As a matter of fact, two dictionaries played the leading role in the early history of the modern language, namely the *Guóyīn zìdiǎn* 國音字典 ('Dictionary of national pronunciation') and the *Guóyīn chángyòng zìhuì* 國音常用字彙 ('Glossary of frequently used characters in national pronunciation'). The former is the codification of the first version of the national standard pronunciation, known as 'Old national pronunciation' (*lǎo guóyīn*), which was an artificial mixture of northern and southern features. Released in 1919, the *Guóyīn zìdiǎn* indicates the standard pronunciation for some 6,500 characters with phonetic symbols. The dictionary served as a norm for teaching manuals to be used in the new primary school subject, National Language. It was only in 1932 that this linguistic hybrid was replaced by the new national pronunciation (*xīn guóyīn*), based on the Běijīng dialect, which in turn was the basis of the *Guóyīn chángyòng zìhuì*.¹⁵

Although these two dictionaries mark milestones in the early history of the modern standard language, their lexicographic significance is rather limited. Of more lasting influence as Chinese monolingual dictionaries per se are two multi-volume dictionaries of the republican period: the *Cíyuán* 辭源 ('Origin of words'), edited by Lù Ěrkuí 陸爾奎, and the *Cíhǎi* 辭海 ('Ocean of words'), edited by Shū Xīnchéng 舒新城. Comprising close to 100,000 entries arranged according to the system of 214 Kangxi classifiers, the *Cíyuán* focuses on the development of the language approximately up to the Opium War (c. 1840), but it also contains some lexical innovations of the second half of the nineteenth century.¹⁶ The *Cíhǎi* covers ancient and modern words, including scientific and technical terms as well as personal and geographical names.¹⁷ Although neither dictionary is as much oriented towards language history in terms of entry information as the *Kāngxī zìdiǎn*, neither can be considered a dictionary of the contemporary language in a strict sense. Even if a lexicographer of the time had endeavoured to document contemporary language use, this undertaking would have faced serious challenges. Most importantly, Mandarin in the early twentieth century possessed anything but a standardized vocabulary. Instead, the standardization of pronunciation was accompanied by lexical and grammatical diversification, experimentation, and continuous reinvention. As Edward Gunn writes, 'the period following 1918 . . . brought on a major, if not unique, revolution in prose style, and one

¹⁵ Kaske, 'National Language Movement', 142; Chen, *Modern Chinese*, 16–22; DeFrancis, *Nationalism and Language Reform*, 55–66.

¹⁶ Kholkina, 'Lexicography, modern'. ¹⁷ Kholkina, 'Lexicography, modern', 594.

that probably surpassed the Japanese experience in the number and scope of innovative constructions'.¹⁸

The releases of the *Cíyuán* and the *Cíhǎi* can be considered milestones in the proliferation of so-called *cí*-type dictionaries, which need to be distinguished from *zì*-type dictionaries, as represented by *Kāngxī zìdiǎn*. An appraisal of the typological distinction discussed here requires some explanatory remarks on the relation between Chinese written graphs and language. As pointed out above, entries in the *Kāngxī zìdiǎn* are single characters (*zì*), and entry information pertains to single characters only. In contrast, entries in dictionaries of the *cí* type are lexical items and other expressions of any length, including single- and multiple-character expressions. It has become customary to translate these two types respectively as 'character dictionary' (*zì* type) and 'word dictionary' (*cí* type). It is important to emphasize, however, that these translations suggest a conceptual complementarity that is not fully applicable to the dictionaries in question: a *zì*-type dictionary contains characters standing for both free and bound morphemes. Since the former have word properties, it is evident that many entries in 'character dictionaries' also qualify as words. Vice versa, many 'words' documented in *cí*-type dictionaries are monosyllabic expressions written with one character.

Since the *Kāngxī zìdiǎn* is based on classical texts of the Confucian canon, the exclusion of polysyllabic entries comes as no surprise. In Classical Chinese, the majority of characters correspond to one monosyllabic morpheme, for example 王 *wáng* 'king', 問 *wèn* 'ask', 皆 *jiē* 'all, every', 知 *zhī* 'know', and 必 *bì* 'certainly, must'.¹⁹ This straightforward relational triangle character–morpheme–syllable has inspired many myths about the Chinese language, one of these being the idea that Chinese only has monosyllabic words, which in turn led to the characterization of Chinese as a 'monosyllabic language'.²⁰ This generalization arguably owes its existence to the fact that early linguistic analyses of Chinese were based on Classical Chinese texts of the Confucian tradition only. For some 2,000 years, however, Classical Chinese has existed side by side with vernacular texts that possess a considerably higher portion of polysyllabic expressions.²¹ After the middle of the nineteenth century, China's increasing exposure to the Western world fostered a new interest in Western knowledge, which again found expression in various translation activities. Since many words in Western texts lacked

¹⁸ Gunn, *Rewriting Chinese*, 40. ¹⁹ See Zádrapa, 'Word and wordhood, premodern', 550.

²⁰ For linguistically sound demystifications, I refer to Kennedy, 'Monosyllabic myth', and DeFrancis, *Chinese Language*, 177–88.

²¹ See Masini, *Formation of Modern Chinese Lexicon*, 121ff.

Chinese equivalents, these translations introduced large-scale lexical innovations, leading to an even stronger tendency towards polysyllabism in the formation of the modern lexicon.²² Examples of modern neologisms which are attested in texts of the second half of the nineteenth century and which are still used today include 版權 *bǎnquán* ‘copyright’, 保險 *bǎoxiǎn* ‘insurance’, 博物館 *bówùguǎn* ‘museum’, 傳播 *chuánbō* ‘diffuse’, 大腦 *dànǎo* ‘brain’, 電池 *diànchí* ‘battery’, 法院 *fǎyuàn* ‘tribunal, court’, 鋼筆 *gāngbǐ* ‘fountain pen’, 公司 *gōngsī* ‘company’, and 公園 *gōngyuán* ‘public park’.²³

Such polysyllabic expressions cannot be found in a traditional *zì*-type dictionary, since they are not part of the lexicographical concept underlying the compilation of these dictionaries. Hence, with their systematic inclusion of polysyllabic entries, the *Cíyuán* and the *Cíhǎi* effectively broadened the lexical scope of Chinese lexicography. Although both dictionaries were only to a limited extent devoted to lexical innovations, they nonetheless contributed to a proliferation of a lexicographic arrangement that allowed the incorporation of the polysyllabic modern lexicon. To be sure, polysyllabic expressions can already be found in earlier dictionaries. A well-known example is the *Pèiwén yùnfǔ* (‘Treasury of rhymes from the Hall for Honouring Literature’), commissioned in 1704 and completed in 1711.²⁴ This collection of compounds and phrases, however, was an aid for poetic composition and therefore does not qualify as a dictionary of contemporary modern vocabulary.

Despite this lexicographical innovation, single graphs continued to occupy the top position within the lexicographical hierarchy. In the *Cíyuán* and the *Cíhǎi* as well as in modern dictionaries, multi-character expressions are subentries appearing under single head characters. Here is an example of the traditional arrangement of single- and multi-character entries:

Head character	Subentries (first level)	Subentries (second level)	Gloss
人	人工 人物	人工湖 人物畫	man, person man-made, artificial personality artificial lake portrait painting

²² Mair, ‘Buddhism and the rise of the written vernacular’, 122.

²³ Examples selected from Masini, *Formation of Modern Chinese Lexicon*, 159–75.

²⁴ Wilkinson, *Chinese History*, 81.

The head character is 人, a two-stroke character with the modern reading *rén* ‘man, person’. Under the main entry, subentries appear on different levels: on the first level are expressions written with two characters, followed by subentries consisting of three characters, and so on. Within each level, the sequence follows the number of strokes of the second character: 人工 (composed of 人 and 工 *gōng* ‘work, labour’, three strokes) precedes 人物 (composed of 人 and 物 *wù* ‘thing, matter’, eight strokes). The same principle underlies the sequence of 人工湖 *réngōng hú* ‘artificial lake’ and 人物畫 *rénwù huà* ‘portrait painting’. From these examples it is evident that traditional dictionaries of the *cí*-type maintain a strict orientation towards written graphs and their graphic properties. These basic principles of lexicographic arrangement are still applied today, albeit sometimes with modifications.

The publication of the first volumes of the *Guóyǔ cídiǎn* 國語辭典 (‘Dictionary of the national language’) marked another crucial step in the modernization of lexicography in China. Editorial work started in 1928 and involved some prominent linguists of the younger generation, namely Wāng Yí 汪怡, Lǐ Jǐnxī 黎錦熙, and Yuen Ren Chao 趙元任.²⁵ In contrast to the *Cíyuán* and the *Cíhǎi*, the *Guóyǔ cídiǎn* has a much stronger focus on Mandarin as the new standard language. Whereas the first editions of the former do not indicate pronunciation, the latter has a sound-based arrangement according to Mandarin Phonetic Symbols. In addition, the pronunciation of each entry is indicated in Mandarin Phonetic Symbols and the transcription system known as Gwoyeu Romatzyh. The *Guóyǔ cídiǎn* thus for the first time combines essential features of modern Chinese lexicography: sound-based arrangement, inclusion of single and multiple-character expressions, coverage of the modern lexicon, and indication of standard pronunciation.

Early Bilingual Lexicography

Early bilingual lexicography involving Chinese and European languages is associated with two non-obvious differences when compared to (native) monolingual lexicography: the selection of Chinese source languages, and a digraphic arrangement of entries involving Chinese characters and transcriptions using the Roman alphabet. Both differences are attested in the earliest period of Chinese–Western bilingual lexicography, that of the contributions of Western missionaries after the late sixteenth century (for which,

²⁵ Kholkina, ‘Lexicography, modern’, 595; Wilkinson, *Chinese History*, 86.

see also Chapter 29). During this period, direct cultural contact with China led to an immense interest in Europe in the language (or languages) of China and the Chinese script; conversely, from the Chinese point of view, the use of alphabetic writing provided new possibilities for lexicographic arrangement. Previous scholarship has focused on the contributions by Jesuit missionaries, notably Michele Ruggieri, Matteo Ricci, and Nicolas Trigault. It has been claimed that the former two were the authors of a handwritten Portuguese–Mandarin dictionary compiled in the late sixteenth century, although there remain doubts concerning its authorship.²⁶ Matters are less controversial in the case of Trigault’s voluminous *Xīrú Ěrmùzī* 西儒耳目資 (‘Aid to the ears and eyes of Western scholars’), published in co-operation with Chinese literati in 1626. Although this dictionary does not qualify as a bilingual dictionary, it is nonetheless a milestone in the history of Chinese–Western lexicography, since it marks the first systematic attempt to integrate Romanized transcriptions of character readings into a Chinese dictionary.

The Jesuits in China had a clear policy of favouring the language of the political elite for social intercourse. As a consequence, transcriptions of Chinese words and phrases in documents compiled by the Jesuits are in Mandarin (*guānhuà*) only. The only Sinitic vernacular that is documented in bilingual dictionaries pre-dating the eighteenth century is Southern Mǐn. These dictionaries likewise resulted from missionary work but, unlike those compiled by missionaries in China proper, Southern Mǐn dictionaries were written by missionaries based among the Chinese settlers in the Philippines. Four manuscript dictionaries dating back to the first half of the seventeenth century are still extant, but it seems certain that many other works have been lost.²⁷ The extant sources differ considerably in terms of the integration of Chinese and alphabetic writing. Pedro Chirino in his handwritten *Dictionarium Sino Hispanicum* (1604) attaches Romanized transcriptions to Chinese characters. The anonymous *Bocabulario de la lengua sangleya* (c. 1620) uses Romanized transcriptions exclusively, for Chinese (Southern Mǐn) entries and also for examples. Ricci’s and Ruggieri’s aforementioned Portuguese–Mandarin dictionary has a spatial separation of Romanized transcriptions and Chinese characters into two different columns. Unfortunately, these early handwritten dictionaries are not accompanied by any kind of textual evidence explaining the principles of entry arrangement; it

²⁶ Ruggieri and Ricci (attrib.), *Dicionário Português-Chinês* (2001), is a facsimile edition, with introduction by J. W. Witek at 155–67 and a useful essay by P. F.-M. Yang at 171–209; Barreto, ‘Reseña’, offers counterarguments against Ruggieri and Ricci as authors.

²⁷ Klöter, *Language of the Sangleys*, 51–82; Van der Loon, ‘Manila incunabula’, 95–108.

is thus not evident whether the use of different scripts was a matter of individual lexicographical principles or whether differences resulted from ad hoc solutions without in-depth reflection.

It is, however, evident that cultural contact between China and Europe had led to the compilation of a significant number of bilingual dictionaries long before the nineteenth century. However, due to a lack of reliable bibliographical information, the dimensions of early European–Chinese lexicography cannot be indicated with certainty. Huiling Yang estimates that between the sixteenth and the early nineteenth centuries more than sixty manuscript dictionaries were written.²⁸ According to different bibliographies, at least eighteen dictionaries, wordlists, and grammatical treatises of Southern Mǐn alone must have been compiled between the fifteenth and the eighteenth centuries, but the current whereabouts of most of the manuscripts remains unknown.²⁹ Important works were written not only by missionaries based in China, but also by European philologists who had studied the missionary documents. The earliest Chinese dictionary printed in Europe was a 44-page Chinese–French dictionary published in 1670. It appeared as an appendix to the French translation of Athanasius Kircher's *China illustrata*, of which the Latin original had been published three years before.³⁰ Some 150 years later, in 1813, the French diplomat and proto-Sinologist C. L. Joseph de Guignes published the *Dictionnaire chinois, français, et latin* in Paris. Since de Guignes failed to mention that the dictionary was based on a manuscript by the Italian Franciscan missionary Basilio Brollo, the publication stirred up a large scandal among European scholars.³¹

After the beginning of the nineteenth century, bilingual European–Chinese dictionary compilation came to full blossoming. This development can mainly be attributed to two reasons. First, the development of new printing techniques allowed for the production of a substantial number of bilingual and digraphic editions. Secondly, the enforced opening of China's major port cities for missionary work and China's growing involvement in international affairs created various new needs for high-quality dictionaries. As a matter of fact, however, the first lexicographic milestones were set even before the opening of five port cities following the Treaty of Nanking in 1842, the best-known example being Robert Morrison's Chinese–English *Dictionary*

²⁸ Yang, 'Making of the first Chinese–English dictionary', 304.

²⁹ Klöter, *Language of the Sangleys*, 53–6. ³⁰ Masini, 'Notes on the first Chinese dictionary'.

³¹ For details, see Masini, 'Chinese dictionaries prepared by Western missionaries', 191–2; Lundbæk, 'Establishment of European sinology', 29.

of the Chinese Language in Three Parts. Printed in a run of 750 copies, Morrison's dictionary was not only of obvious benefit for missionaries preparing for service in China, but was also widely studied by the first cohort of European Sinologists.³²

The magnificence of Morrison's lexicographic contribution cannot be overestimated. According to W. South Coblin, he first expressed his intention to compile a dictionary in 1808.³³ Starting more or less from scratch, the production of 4,597 dictionary pages within fifteen years must be considered a major accomplishment in itself. It is especially impressive in the light of Morrison's various other tasks and duties as a Bible translator, missionary, educator, and language student.³⁴ In addition, the compilation of the dictionary was hampered by both technical constraints and legal restrictions. Morrison arrived in China in 1807, at a time when it was strictly forbidden for Chinese to teach their language to foreigners. Prior to his arrival in China, Morrison had already studied Chinese in England. The printing of digraphic entries was part of his original plan, since the British printer Peter Perring Thoms was sent to China to assist Morrison with solving the technical problems. Morrison and Thoms sought co-operation with professional Chinese punch cutters, but this undertaking was likewise subject to legal restrictions. As a consequence, the preparation of the dictionary resembled an undercover activity.³⁵

In sociolinguistic terms, Morrison's dictionary is a continuation of the Jesuit tradition of documenting the languages of China's cultural and political elite, namely Classical Chinese and Mandarin. Although he heavily relied on the *Kāngxī zidiǎn* for lemma selection, he added various example sentences from other sources, including Classical Chinese literature and phrases belonging to spoken Mandarin and the northern written vernacular.³⁶ The use of transcriptions based on the pronunciation of Nánjīng for Mandarin phrases is evident.³⁷ As regards the quotations from the classics, however, it needs to be emphasized that there is no intrinsic relation between Classical Chinese texts and their recitation in Mandarin. Classical Chinese can be recited in any Sinitic variety, and there are also established recitation traditions in Japanese, Korean, and Vietnamese. The issue becomes clearer when we compare

³² For the print run, see Yang, 'Making of the first Chinese–English dictionary', 316.

³³ Coblin, 'Robert Morrison and the phonology of mid-Qīng Mandarin', 341.

³⁴ Yang, 'Making of the first Chinese–English dictionary', 300–3.

³⁵ Starr, 'Legacy of Robert Morrison', and Lehner, *Druck chinesischer Zeichen in Europa*, 38.

³⁶ Yang, 'Making of the first Chinese–English dictionary', 315.

³⁷ Coblin, 'Robert Morrison and the phonology of mid-Qīng Mandarin', 341.

Morrison's dictionary with the Hokkien (Southern Mǐn) dictionary compiled by the English Congregationalist Walter Henry Medhurst in 1832. We note that Medhurst prepared his dictionary when he was based among the overseas Chinese community of Malacca after 1816, which explains both the early date of its compilation and the choice of a regional vernacular.³⁸ Nonetheless, although the source language of Medhurst's dictionary is a regional vernacular, the bulk of expressions and example phrases are, as in Morrison's, selected from standard written sources, appended by Romanized transcriptions indicating the Southern Mǐn recitation form. This explains why identical entries and pieces of textual evidence can be found in dictionaries of two different Sinitic varieties. So, for instance, for the expression 儒士, Medhurst has 'Jê soō "a learned man"' and Morrison has 'Heō sze "a learned man"'; for the expression 雍也可使南面, Medhurst has 'Yŭng yěá, k'ho soo lâm bēēn "Yŭng can be set to face the south"' and Morrison has 'Yung yay k'ho she nan meen "Yung is worthy of a throne"'; and for the expression 同, Medhurst has 'tōng "the same, alike"' and Morrison has 't'hung "together, the same"'.

As bilingual lexicography came into full swing after the opening of China in the middle of the nineteenth century, dictionary compilation developed in different directions with regard to the Chinese and the Western languages documented in the dictionaries. On the Chinese side, during the hundred years following the beginning of Protestant missionary work in China, an unprecedentedly rich variety of Sinitic regional vernaculars were recognized as either source or target languages of bilingual dictionaries. As a matter of fact, Western missionaries attached more importance to regional vernaculars than Chinese dictionary-makers. To be sure, topolects had caught the attention of Chinese lexicographers much earlier. For example, the lexicographic tradition of compiling rhyme books had been applied to dialects at least by the Yuán dynasty (1271–1368; see Chapter 6).³⁹ In addition, during the Míng (1368–1644) and Qīng (1644–1911) dynasties, Chinese scholars continued the tradition of the *Fāngyán* ('regional words') lexicography initiated by Yáng Xióng nearly two thousand years earlier (see Chapters 3 and 6).⁴⁰ However, although the history of Chinese topolect dictionaries has yet to be written, it seems safe to claim that the proliferation of topolect lexicography is largely linked to the Western missionary tradition. So it was that during the second half of the nineteenth century, *Chinese* in the title of dictionaries became an umbrella term for a wide range of distinct vernaculars. Mandarin was

³⁸ Klöter, *Written Taiwanese*, 108–9. ³⁹ See also Klöter, 'Dialect dictionaries', 50.

⁴⁰ See also Yong and Peng, *Chinese Lexicography*, 93–4.

documented in Medhurst's *Chinese–English Dictionary* (1842–3) and his *English–Chinese Dictionary* (1847–8); Cantonese in Samuel Wells Williams' *Tonic Dictionary of the Chinese Language* (1856) and Wilhelm Lobscheid's *English and Chinese Dictionary* (with Mandarin and Cantonese transcriptions) (1866); Wú in Joseph Edkins' *Vocabulary of the Shanghai Dialect* (1869) and William Morrison's *Anglo–Chinese Vocabulary of the Ningpo Dialect* (1876); Northern Mǐn in R. S. Maclay and C. C. Baldwin's *Alphabetic Dictionary of the Chinese Language in the Foochow Dialect* (1870); Southern Mǐn in Carstairs Douglas' *Chinese–English Dictionary of the Vernacular or Spoken Language of Amoy* (1873) and John MacGowan's *English and Chinese Dictionary of the Amoy Dialect* (1883); Hakka in Donald MacIver's *English–Chinese Dictionary in the Vernacular of the Hakka People* (1905). Most of these bilingual dictionaries were unidirectional.

As regards the Western language in a bilingual dictionary, English dominated other European languages. As pointed out by David Prager Branner, 'the majority of the early Protestant missionaries were British and American, and the surge in printing of Western-language dictionaries was in the main limited to the English language until near the end of the [nineteenth] century . . . It was only in the case of Mandarin dictionaries that English lagged behind other European languages.'⁴¹

Dictionaries compiled by missionaries or members of foreign consular services in China dominated bilingual lexicography during the second half of the nineteenth and throughout the first half of the twentieth centuries. Although these dictionaries were in the first instance compiled for the purposes of foreigners in China, they also gained increasing popularity among Chinese who had interactions with Westerners. Parts of Morrison's dictionary were reprinted by Chinese publishing houses and used as teaching manuals for Chinese students.⁴² Although some reference works on Chinese and other languages were also compiled by Chinese authors, these do not qualify as bilingual dictionaries in a narrow sense.⁴³ Outside China, with the opening of new Chinese studies departments at several European universities during the nineteenth century, Chinese–English dictionaries started to play an important role in academia. New cohorts of Sinology students relied on dictionaries for the study and translation of the Chinese classics. Towards the end of the century, most European Sinologists relied on Mandarin pronunciation when it came to the recitation of texts and the indication of character readings. Some important works in this respect are the *Chinese–English*

⁴¹ Branner, 'Notes on the beginnings of systematic dialect description', 236.

⁴² Si, 'Reprinting Robert Morrison's dictionary'.

⁴³ A few examples are mentioned in Yong and Peng, *Chinese Lexicography*, 270–1.

Dictionary (1892) by the British consular official Herbert Giles and a dictionary of the same title by the Australian Congregationalist missionary Robert H. Mathews (1931).⁴⁴

Up to today, it has remained a contentious issue among lexicographers whether and how Chinese and alphabetic writing should be integrated. A basic question is whether Chinese characters should be used at all. In the first half of the nineteenth century, the answer was clear. According to John DeFrancis, 'at the time of Morrison's death in 1834 there was still unanimous agreement among experts and amateurs alike that the Chinese language should be characterized as exceedingly difficult, uniquely monosyllabic, and most important, necessarily ideographic in its written form'.⁴⁵ As pointed out above, Morrison had made great efforts to realize the digraphic edition of his dictionary. His argument was that, 'Any thing in Chinese, beyond common place topics of colloquial intercourse, is quite unintelligible when expressed in Letters of the Roman Alphabet. The Character must be present to the eye, or to the mind, in order to be understood'.⁴⁶

Following the linguistic diversification of bilingual lexicography after the middle of the nineteenth century, however, arguments for and against the use of Chinese characters changed. One reason was that most of the south-eastern vernaculars, in contrast to Classical Chinese or the northern vernacular, lacked an established tradition of writing. As a consequence, missionaries made use of the Roman alphabet to a much higher degree when it came to the writing of vernacular texts in general and to the compilation of dictionaries in particular. It was in the province of Fújiàn during the late nineteenth century that alphabetic writing came to be most widely used as the main script of Southern Mǐn translations of Christian texts and in teaching manuals for local Chinese children and women. Through its increasingly wide use, the Roman alphabet gradually gained new legitimacy as an independent script of the Southern Mǐn vernacular. The Scottish missionary-cum-lexicographer Carstairs Douglas expressed a more radical independence from Chinese writing than anyone before him.⁴⁷ His Chinese–English dictionary uses the Roman alphabet only, as he explained:

The most serious defect is the want of the Chinese character. This is due to two causes: (1) There are a very large number of the words for which we have not been able to find the corresponding character at all, perhaps

⁴⁴ See also Wilkinson, *Chinese History*, 85.

⁴⁵ DeFrancis, *Nationalism and Language Reform*, 18.

⁴⁶ Morrison, *Dictionary of the Chinese Language*, I.ix.

⁴⁷ For details, see Alsford and Fuehrer, 'Carstairs Douglas'.

a quarter or a third of the whole . . . (2) Even if the characters had been found, it would have been very difficult or impossible for me to use the Chinese character in printing at home [in England] . . . Meantime, while I greatly regret that the Chinese character does not appear in the book, I am in one sense glad that it is absent. For it may serve to make manifest the fact that the Vernacular of Amoy is an independent language, which is able to stand alone without the help of the written character.⁴⁸

Douglas' example received both positive and negative responses. An example of the latter is that of William Campbell, a Scottish Presbyterian missionary to Taiwan, who, in the preface of his *Dictionary of the Amoy Vernacular*, opposed the growing disregard of Chinese writing:

Following the trend of recent events in China, it will thus be seen that no sympathy is shown here for the action of those missionary brethren who push forward Roman letters with the avowed intention of thrusting Chinese methods of writing and printing into the scrap-heap. No: seeing that native periodicals are now increasing by the hundred, so the humble contribution herewith submitted also comes forward, not as a Supplanter, but as a cheap convenient little Handbook for helping those who use it to a fuller and more accurate knowledge of the written language of China.⁴⁹

Ironically, even the strongest opponents of character use maintained an invisible character-based lexicographic arrangement. For example, lemmata in Douglas' dictionary are exclusively monosyllabic morphemes, followed by polysyllabic expressions presented as subentries, as shown in the following sequence:

káu a dog	siáu-káu a mad dog. sē-káu a wild animal like a small wolf
káu nine	jī-káu twenty-nine . . . káu-kàng-hong heavy gales about the ninth month
káu the mouth	é-káu dumb, ill-developed . . . toā-chih-káu speaking with difficulty.

Obviously, the practice of grouping all subentries containing one particular monosyllabic morpheme under one lemma is not an alphabetical arrangement in a stricter sense. Instead, it follows the Chinese tradition of *zì*-type lexicography, in which single characters representing monosyllabic morphemes take centre stage. Until today, most alphabetically arranged dictionaries maintain this practice of arranging subentries under single characters. In

⁴⁸ Douglas, *Chinese-English Dictionary*, viii–ix.

⁴⁹ Campbell, *Dictionary of the Amoy Vernacular*, i.

other words, the growing importance of alphabetic arrangements in the second half of the nineteenth century did not pave the way for a radical alphabetic arrangement of Chinese dictionaries.

Lexicography After 1949

The establishment of the People's Republic of China in 1949 marks an important political moment in China's modern history. Political and social changes following the regime change also had a great impact on the development of modern lexicography. In the paragraphs that follow, four aspects will be discussed as major developments of post-1949 lexicography: the development of standard dictionaries and the successful spread of the standard language; the rise of bilingual lexicography 'made in China'; the development of large-scale projects in monolingual lexicography; and functional and typological diversification.

In the early twentieth century, dictionaries such as the *Guóyīn zìdiǎn* and the *Guóyǔ cídiǎn* were assigned important roles in the spread of the new national language. However, language planning before 1949 remained largely unsuccessful due to bureaucratic failures and inefficient government institutions. Despite the existence of a new standard and dictionaries codifying the national language, few people in China were able to speak it. At the same time, literacy rates remained low. After 1949, the new government placed much emphasis on national language planning and language and literacy education. The first measures were taken during the 1950s: the national language was renamed as *pǔtōnghuà* 'common language', the first lists of simplified Chinese characters were released, and a new transcription system, known as *Hànyǔ Pīnyīn*, was devised. It is beyond doubt that China's modern language planning project was extremely successful. According to UNESCO figures, in 2010, the literacy rate of the age group 25–64 was 97 per cent.⁵⁰ In 2017, sixty years after the release of the first guidelines for its promotion, some 70 per cent of the population were able to speak *pǔtōnghuà*.⁵¹ *Hànyǔ Pīnyīn* is now the undisputed national and international standard for the transcription of Mandarin.

Successful language planning is usually the result of several factors, dictionary compilation being one of them. In the context of language planning in China after 1949, two major works can be identified, the *Xīnhuá zìdiǎn* 新華字典 ('New China dictionary') and the *Xiàndài Hànyǔ cídiǎn* 現代漢語詞典

⁵⁰ UNESCO, 'Education: Literacy rate'. ⁵¹ Lǐ, 'Tuīgǎng pǔtōnghuà', 3.

(‘Modern Chinese dictionary’). With its narrow scope, simple lexicographic metalanguage, and modest selling price, the *Xīnhuá zìdiǎn* is a typical example of a pocket-sized dictionary for general use.⁵² The first edition (1953) contains 6,520 entries, most of which are single characters.⁵³ Since then, new editions have been released roughly every six years, the eleventh in 2010. The different editions have followed the most recent developments in language standardization. An example is the second edition, which introduced an alphabetic arrangement based on Hànyǔ Pīnyīn. In 2016, with 567 million copies sold worldwide, the *Xīnhuá zìdiǎn* set a Guinness World Record for ‘the most popular dictionary and the best-selling regularly updated book’.⁵⁴

The *Xiàndài Hànyǔ cídiǎn* is the *cí*-type counterpart of the *Xīnhuá zìdiǎn*.⁵⁵ Its main lemmata are single characters, followed by multiple-character subentries. The latest edition, published in 2016, contains some 10,000 entries (single characters) and around 70,000 multiple-character subentries. Preparation of the *Xiàndài Hànyǔ cídiǎn* started in the late 1950s, and trial editions were released in 1960 and 1965.⁵⁶ Due to the political turmoil during the Cultural Revolution, formal publication of the first edition was delayed until 1978. Preparatory work involved some of China’s most prominent linguists, including Dīng Shēngshù 丁声树, Lí Jīnxī, Lǐ Róng 李荣, Lǚ Shūxiāng 吕叔湘, and Wáng Lì 王力.

The prescriptive approach of the *Xiàndài Hànyǔ cídiǎn* is obvious. As pointed out by Siu-Yau Lee, from the outset, the preparation of the *Xiàndài Hànyǔ cídiǎn* was declared an integral part of promoting *pǔtōnghuà* and simplified characters.⁵⁷ It has regularly claimed authority for defining correctness of the shape of individual characters, pronunciation, entry selection, usage, definitions, and so on. Correctness in terms of selected entries and their definitions has often had strong ideological dimensions. On the other hand, however, the compilers have also made obvious efforts to document actual language use. The descriptive approach is evident in the sixth edition, for which some 3,000 new entries were selected. Many of the newly selected entries are colloquial expressions reflecting social trends and lifestyles of the

⁵² The editor-in-chief of the first edition of the *Xīnhuá zìdiǎn* was Wèi Jiàngōng 魏建功 (see Kholkina, ‘Lexicography, modern’, 596). It was published under the name of the *Xīnhuá Cǐshū Shè* (‘New China Dictionary Publisher’).

⁵³ Kholkina, ‘Lexicography, modern’, 596.

⁵⁴ ‘Guinness World Records Announces Xinhua Dictionary’.

⁵⁵ The editorial board of the first unpublished trial version of the *Xiàndài Hànyǔ cídiǎn* was headed by the linguist Lǚ Shūxiāng. The first published version was released in 1978, the editor-in-chief being Dīng Shēngshù. This and the later editions were published under the name of the Chinese Academy of Social Sciences.

⁵⁶ Tang, ‘Xiàndài Hànyǔ cídiǎn 現代漢語詞典’, 608–9. ⁵⁷ Lee, ‘Defining correctness’, 430.

reform era, such as 闪婚 *shǎnhūn* ‘whirlwind wedding’, 闪离 *shǎnlí* ‘whirlwind divorce’, 宅男 *zháinán* ‘homebody, indoorsy (male)’, and 宅女 *zháinǚ* ‘homebody, indoorsy (female)’.⁵⁸

After the 1980s, the typological distinction of *zì*-type and *cí*-type dictionaries culminated in two major multi-volume works: the eight-volume *Hànyǔ dà zìdiǎn* 汉语大字典 (‘Unabridged Chinese character dictionary’) and the thirteen-volume *Hànyǔ dà cídiǎn* 汉语大词典 (‘Unabridged Chinese dictionary’). The publication of these multi-volume works followed in the footsteps of works published outside the People’s Republic of China, the most important example being Morohashi Tetsuji’s 諸橋轍次 thirteen-volume *Dai Kan-Wa jiten* 大漢和辭典 (‘Unabridged Chinese–Japanese Dictionary’), publication of which began in 1955.⁵⁹ Lexicographic information in the *Hànyǔ dà zìdiǎn* focuses on single characters only; the *Hànyǔ dà cídiǎn*, by contrast, has comprehensive sections containing multi-character subentries. Covering 54,678 head characters, the *Hànyǔ dà zìdiǎn* broke the record of lexicographically attested characters originally set by the *Kāngxī zìdiǎn* by some 7,500 items. This record has in the meantime been broken by the *Zhōnghuá zìhǎi* 中华字海 (‘Chinese ocean of characters’), edited by Lěng Yùlóng 冷玉龙 and Wéi Yīxīn 韦一心, with its 85,586 different characters. Entry information in the *Hànyǔ dà zìdiǎn* is diachronic and centres on historical character forms, references in traditional dictionaries, modern pronunciations and historical sound glosses, and meanings. Although the number of head characters in the *Hànyǔ dà cídiǎn* is ‘only’ 23,000, with its almost 350,000 subentries, it is by far the most comprehensive and authoritative dictionary of the Chinese language. Citing quotations from literary sources covering a period of 3,000 years, it likewise follows a diachronic approach.⁶⁰

The compilation of monolingual standard dictionaries was followed by new developments in bilingual lexicography. Whereas before 1949 the major works had been by lexicographers from the West, after the beginning of the reform era in 1978, China’s state language agencies and publishing houses successfully established themselves in the field of bilingual lexicography. Especially with China’s growing involvement in international trade and diplomacy, general and specialized Chinese–English dictionaries started to dominate the market. It has been estimated that between 1978 and 2007 some 1,500 bilingual dictionaries including Chinese were published.⁶¹ It is very

⁵⁸ See Klöter, ‘“What is correct Chinese?” revisited’, 65–7.

⁵⁹ Wilkinson, *Chinese History*, 91.

⁶⁰ For details, see Klöter, ‘Chinese lexicography’, 887; Wilkinson, *Chinese History*, 88–91.

⁶¹ Yong and Peng, *Bilingual Lexicography*, 14.

likely that the number has risen sharply over the past ten years. To give a rough indication: a search query for 汉英词典 ('Chinese–English dictionary') on the website of one of China's biggest online book retailers yields more than 4,500 results, compared to a little more than 100 titles for the Chinese–French equivalent.

A pioneering contribution to modern bilingual lexicography was the *Hàn–Yīng cídiǎn* 汉英词典 published in 1978. Later, its editor-in-chief Wú Jǐngróng 吴景荣 combined his plea for a modern dictionary based on scientific compilation with harsh condemnations of earlier works.⁶² Most importantly, he imputed 'superficial acquaintance with Chinese language and culture' to previous lexicographers such as Giles or Mathews.⁶³ Wú's *Hàn–Yīng cídiǎn* was largely based on the *Xiàndài Hànyǔ cídiǎn*. The *Hàn–Yīng cídiǎn* in turn became a model dictionary for other language editions, which followed Wú's work in terms of entry selection, arrangement, examples, and use of information labels. A Chinese–German edition was published in 1985, followed by a Chinese–French edition in 1991 and a Chinese–Spanish edition in 1999.⁶⁴ In other words, the first set of bilingual dictionaries published in China were all unidirectional Chinese–L2 dictionaries. The first major English–Chinese dictionary published by Chinese lexicographers is the *English–Chinese Dictionary* in two volumes, first compiled by a group of Shànghǎi-based specialists under the leadership of Lù Gǔsūn 陆谷孙.

Modern Chinese–English lexicography has not remained restricted to works published in the People's Republic of China. Well-known earlier examples of Chinese–English dictionaries published outside China are Liang Shih-chiu's 梁實秋 *Practical Chinese–English Dictionary* (revised editions appeared under the title *Far East Chinese English Dictionary*) and Lin Yutang's 林語堂 *Chinese–English Dictionary of Modern Usage*, both published in the early 1970s, in Taipei and Hong Kong respectively. A more recent example that deserves special mention is the *ABC Chinese–English Comprehensive Dictionary* by John DeFrancis, published in 2003 in Hawaii. These titles nicely exemplify some of the differences between bilingual Chinese lexicography inside and outside China. Inside China, as pointed out above, the arrangement of bilingual dictionaries by and large followed in the footsteps of the *Xiàndài Hànyǔ cídiǎn*, in a hybrid alphabetic–graphic arrangement. By contrast, the arrangement of Liang's dictionary is very traditional, being based on the list

⁶² Branner and Meng, 'Curious lexicographic relic', 559–63.

⁶³ Wú, 'Chinese–English dictionaries', 519.

⁶⁴ For details, see Wu, 'Chinese–English dictionaries', 521; Klöter, 'Chinese lexicography', 888.

of 214 Kangxi classifiers. Lin's dictionary, on the other hand, arranges the sequence of entries according to a now widely forgotten numerical 'Instant Index System'. DeFrancis' *ABC* dictionary is the most prominent example of a strict letter-by-letter alphabetical order without interference of written graphs, irrespective of single characters representing superordinate morphemes. From a philological perspective, the major bilingual multi-volume dictionary is the seven-volume Chinese–French *Grand dictionnaire Ricci de la langue chinoise*. It covers 13,500 head characters and 300,000 meticulously annotated subentries.

Chinese–English dictionaries published inside and outside China differ not only in terms of arrangement but also in terms of entry selection and definition. A defining feature of dictionaries published in the People's Republic of China (PRC) after the late 1970s is the use of lexical innovations reflecting communist ideology and social changes of the post-1949 period. Lin's and Liang's dictionaries, on the other hand, follow different approaches when selecting and defining words associated with communism. In contrast to PRC dictionaries, they both avoid political terminology reflecting specific PRC usage, for example expressions like 毛澤東思想 *Máo Zédōng sīxiǎng* 'Mao Zedong Thought', 三反運動 *Sānfǎn Yùndòng* 'the movement against the three evils', and 延安精神 *Yán'ān jīngshen* 'Yan'an spirit' (the city of Yan'an has heroic associations in the story of Chinese communism). Examples of communist terminology in a broader sense – words and expressions also used outside the PRC – appear as entries, albeit with English translations different from those found in PRC dictionaries. For example, in Wu's Chinese–English dictionary, 紅旗 *hóngqí* is translated and explained as 'red flag or banner (often as a symbol of the proletarian revolution or of an advanced unit)'. Liang Shih-chiu's translations of the same entry are '1. A red flag or banner (as a symbol of revolution); 2. A red flag used as a symbol of danger'. Lin Yutang offers the second explanation only: 'red flag denoting danger'.

In the past decades, dictionaries have developed into best-selling products of China's liberalized and diversifying book market, and many new types of dictionaries have emerged. One is a category of learners' dictionaries, targeted at the needs of the rapidly growing number of students learning Mandarin as a foreign language. As exemplified by Lǐ Lùxīng's 李禄兴 *Dictionary of 5000 Graded Words for New HSK*, entry selection in many learners' dictionaries follows the standard set by the official Hànyǔ Shuǐpíng Kǎoshì (HSK) language examination. Some learners' dictionaries, like Tēng Shǒu-hsin's 鄧守信 *Far East 3000 Chinese Character Dictionary*, indicate stroke orders

and frequencies of 3,000 individual characters. Another typological innovation is a group of dictionaries containing bilingual definitions (*shuāngjiě cídiǎn*), in other words, bilingual L2–Chinese dictionaries having L2 and Chinese definitions. These are typically based on monolingual dictionaries of Western languages of well-known international publishing houses, for example the *Longman Dictionary of Contemporary English*.

With the advent of electronic dictionaries at the end of the twentieth century, the traditional printed dictionary has a new competitor. Some of the major printed dictionaries now have online versions, one of the prominent examples being the *Hànyǔ dà cídiǎn*. Since the beginning of the twenty-first century, e-dictionaries that are not based on a previous printed edition have quickly conquered the field of general and learners' dictionaries. Before the advent of digital lexicography, looking up an entry in a Chinese dictionary often involved the identification of classifiers and thumbing through classifier and character indexes. This practice was based on lexicographical principles of ordering which have had a thorough influence on philological conceptualizations of the Chinese language and script in general. The use of digital dictionaries involves speaking to, typing, or finger writing on the relevant hardware. Due to their multi-mode access to entries, electronic dictionaries have effectively evaded the necessity of deciding whether a dictionary should be sound-based or shape-based. It is foreseeable that this innovation will fundamentally affect traditional conceptualizations of the Chinese language.

Conclusion

Dictionary-making for the Chinese language is not only a matter of collecting words according to particular principles; lexicographers also have to come to grips with the graphical elements representing the words. When Chinese lexicography produced one of the most prestigious character dictionaries in history, the *Kāngxī zìdiǎn*, alphabetic writing had already made its first modest forays into the fringes of China's literary circles. Some 200 years would pass before alphabetic writing in the form of Romanized transcriptions first appeared in an official standard dictionary. From there, it was only a short step to the establishment of alphabetically arranged standard dictionaries. To state the obvious, the proliferation of alphabetic writing in Chinese lexicography does not imply that the character script was replaced. Instead, as has been pointed out in this chapter, the convergence of Western and Chinese traditions in modern Chinese lexicography has led to hybridity and

not to expulsion. The use of different scripts in a digraphically arranged dictionary is but one area of lexicographic practice discussed in this chapter. Closely related and yet distinct fields are the nature of entries (single characters or single plus multiple characters), and the oppositions of literary and vernacular language; old and new terms; and prescription and description.

Japanese, Korean, and Vietnamese from c. 1800

DAVID LURIE, HEOKSEUNG KWON, AND JOHN D. PHAN

The story of Japanese and Korean lexicography from the beginnings to the eighteenth century was told in Chapter 10 above as part of the story of the lexicography of the Chinese periphery, and the story of the earliest lexicography of Vietnamese, which was undertaken by European missionaries, is told in Chapter 29 below. The story which this chapter will tell is a complex one, tracing the continuing development of the traditions of Japanese, Korean, and Vietnamese lexicography, in contact with each other and with Western and Chinese traditions.

Japanese

DAVID LURIE

Since the start of the nineteenth century, Japanese lexicography has been transformed by interlocking social, linguistic, and technological changes. During the Tokugawa period (seventeenth to mid nineteenth centuries), a feedback loop between expanding literacy and woodblock printing led to a boom in commercial publishing, including many dictionaries and encyclopedias. At the same time, new intellectual movements produced both demand for innovative reference works and scholars dedicated to compiling them. From the latter half of the nineteenth century, an influx of Western words, texts, and ideas contributed to rapid linguistic change, and the rise of universal primary education expanded markets for dictionaries; elites inside and outside the government also promoted a standardized national spoken and written language. (One result was a new discipline of Japanese linguistics or national language studies (*kokugogaku*), which would be the field of most modern lexicographers.) Throughout the twentieth century, social and technological change generated further neologisms, and in the post-war period, script and orthographic reforms and rapid expansion of secondary and

tertiary education fomented new demand for dictionaries in homes, classrooms, and offices. Publishers in headlong competition participated in a thriving and diverse market for monolingual dictionaries; in the latter half of the twentieth century, the more popular ones sold in the millions. Space limitations preclude coverage here, but there is an equally rich modern history of bilingual dictionaries, references devoted to particular fields or speech communities, and encyclopedias.¹

Linguistic and Orthographic Parameters

From the beginnings of widespread literacy in the Japanese archipelago, Chinese characters (Japanese *kanji*; hereafter, ‘sinographs’) were logographically associated with native vocabulary and used to read and write complete texts in a range of styles, from orthodox Literary Sinitic (‘Classical Chinese’) to localized vernacular forms. Sinographs were also used as phonographs to spell out native words, exclusively or in conjunction with logographs. These phonographs eventually evolved into graphically distinct sets unique to Japan: the *kana* ‘syllabaries’ (which actually correspond to morae).²

Due to this complex situation it is difficult to distinguish between monolingual Japanese dictionaries and bilingual dictionaries of Chinese and Japanese. Until the nineteenth century, the dominant imperative in Japanese lexicography was to match sinographs with Japanese readings: both localized pronunciation of loanwords (*on’yomi*) and calque-like links between sinographs and their logographic Japanese readings (*kun’yomi*). Thus the vast majority of pre-nineteenth-century wordlists provided Japanese readings for sinographs or listed sinograph ‘spellings’ for particular Japanese words. In many cases, Chinese dictionaries were reprinted in Japanese editions with added

¹ On early modern and modern Japanese linguistic thought, see Kaiser, ‘Linguistic thought in Japan’, and Bedell, ‘Kokugaku grammatical theory’. On nineteenth- and twentieth-century script and language reform, see Seeley, *History of Writing in Japan*, 136–87; Twine, *Language and the Modern State*; and Gottlieb, *Kanji Politics*. On the modern discipline of *kokugogaku*, see Miller, ‘Traditional linguistics’; Ramsey, ‘Japanese language’; Clark, *Kokugo Revolution*; and Lee, *Ideology of Kokugo*. For the history of scholarship on languages other than Japanese, see Doi, *Study of Language in Japan*. Introductory discussions of bilingual dictionaries can be found in Jiten kyōkai, *Nihon no jisho no ayumi*. Yamagiwa, *Bibliography*, provides a dated but still impressive survey of the entire field of modern reference works. On the entire history of Japanese lexicography, see Okimori et al., *Nihon jisho jiten*, and Okimori, *Zusetsu*.

² On these aspects of the history of writing in Japan, see Lurie, *Realms of Literacy*, and Lurie, ‘Development of Japanese writing’. For a broader East Asian perspective, see Kornicki, *Languages, Scripts, and Chinese Texts*, and Kin, *Literary Sinitic and East Asia*.

phonograph glosses and reading marks, which transformed them into Japanese dictionaries of sinographs with local readings.

Wordlists could be ordered semantically – for instance, from celestial phenomena, to the landscape, to plants and animals, and so on. More commonly, they sorted sinographs into groups sharing a graphic element ('radical'; Japanese *bushu*), or sorted Japanese words by initial *kana*. Two equally arbitrary *kana* sorting systems were available: ordering by initial consonant and vowel (the 'fifty-syllable table' (*gojūon-jun*)), and following the order of the 'Iroha-uta', a remarkable mnemonic poem that employed each mora once and only once.³ Either of these was equally suited to organizing a wordlist but, because of associations between the fifty-syllable table and Western alphabetical order, it dominated from the late nineteenth century onwards.

Late Early Modern Dictionaries and the Nineteenth-Century Revolution

As discussed in Chapter 10, abundant manuscript-based dictionaries gave rise to an expansive printed corpus of lexicographical references in the Tokugawa period. Despite its overall diversity, this field was dominated by two works (given their many variants, really subgenres): *Setsuyōshū* and *Wagyokuhen*. Both of these matched sinographs and Japanese readings: the former listed Japanese words by their initial *kana* (subordinately sorted into semantic categories) and provided their sinographs, for writers looking for logographic 'spellings' of words they already knew; the latter sorted sinographs by radical and provided Japanese glosses for them, for readers puzzled by an unfamiliar logograph. In principle, neither provided discursive definitions or other information (such as usage examples, parts of speech, etymologies, and so on), although publishers competing over new selling points produced some exceptions.

Burgeoning new intellectual and literary fields – studies of Chinese vernacular language and literature, Dutch studies, Sinitic poetry, various genres of vernacular Japanese poetry, and so on – gave rise to many innovative reference works, but scholars emphasize three large-scale dictionaries as key departures of the late Tokugawa period.⁴ The first, Tanikawa Kotosuga's *Wakun no shiori*, which covered the full range of Japanese

³ On the *Iroha-uta*, see Seeley, *History of Writing in Japan*, 106.

⁴ In addition to references from Chapter 10, see Yuasa, 'Edo jidai'.

vocabulary and began publication in 1777, is discussed in Chapter 10.⁵ The second, *Gagen shūran* ('Index of elegant words'), was compiled by Ishikawa Masamochi and began publication in 1826; it collected terms from classical *waka* poetry and related genres with abbreviated definitions and exhaustive citation of usages (still valuable today). The third, *Rigen shūran* ('Index of vernacular words'), is a collection of local dialect (largely of Edo), spoken language, and non-literary vocabulary compiled in the early nineteenth century by Ōta Zensai.

Collectively, these three (and other lesser-known contemporary works) indicate new attention to the meaning, usage, and etymology of Japanese words as such, independent of their sinographic orthography. This is in keeping with the overall tenor of the Tokugawa-period *Kokugaku* ('nativist philology') movement.⁶ However, it is important to note that none of the three great dictionaries were published *in toto* until the Meiji Period (1868–1912): the last instalment of *Wakun no shiori* came out in 1887, the first complete edition of *Gagen shūran* was published the same year, and *Rigen shūran* existed only in manuscript until its first printing in 1899–1900. Thus their publication history highlights continuities across the whole nineteenth century and shows that the move away from sinograph glossing as the *sine qua non* of lexicography was not simply a matter of Western influence.

Shortly after the 1868 Meiji Restoration, the new Ministry of Education attempted to sponsor official lexicographic projects.⁷ A team of scholars started work on an early state-sponsored dictionary, *Goi*, but after a handful of volumes (1871–84), the project was abandoned. Partly because of discord within that editorial team, Ōtsuki Fumihiko was ordered in 1875 to compile a separate dictionary on his own. Completed after a decade of work, this was abandoned by the ministry in 1888, and privately published under the title *Genkai* ('Sea of words') (1889–91).⁸ The failure of *Goi* and the termination of government sponsorship of *Genkai* meant an early and conclusive end to public lexicography in Japan. Compilers of dictionaries are often employed by public universities and research institutes, but their lexicographical projects themselves are funded and published by private companies.

For *Genkai*, the principles for headword selection and the entry structure were inspired by Webster's *American Dictionary of the English Language* (some

⁵ Throughout this chapter, Japanese names are in traditional order, family name first.

⁶ On *kokugaku*, see Nosco, *Remembering Paradise*, and Burns, 'Politics of philology in Japan'.

⁷ For a survey of Meiji period dictionaries, see Hida et al., *Meiji-ki kokugo jisho*. For background on the linguistic terrain of this period, see Saitō, *Kanbunmyaku*.

⁸ See Konno, *Genkai o yomu*.

entries are simply Japanese translations from Webster), but Ōtsuki also incorporated original ideas about Japanese grammar, collected his own usage examples, and drew heavily on early modern (and older) Japanese dictionaries. At around 40,000 entries, it was not large by later standards, but *Genkai* attempted to reflect the full vocabulary of the language, and its entries were systematically structured, included pronunciation as well as orthography, and provided comparatively extensive definitions (as well as, often, examples and etymologies). For these reasons it is often described as the first modern Japanese dictionary. It went through many editions in various formats and, after Ōtsuki's death in 1928, his manuscripts and notes were taken over by a team of editors, resulting in the expanded *Daigenkai* ('Great sea of words') (1932–7); at 97,000 entries, it is over twice the size of the original.

Diversification and Expansion in the Twentieth Century

The expansion of *Genkai* into *Daigenkai* symbolizes one of the major developments of twentieth-century lexicography: the rise of large dictionaries containing more than 100,000 words.⁹ The compilation of the first of these, *Dai-Nihon kokugo jiten* ('National language dictionary of the Japanese empire') (1915–19), began in 1903, about a decade after the publication of *Genkai*. Ueda Kazutoshi, founder of modern Japanese linguistics (*kokugogaku*), shared credit but was just a figurehead: this immense dictionary (roughly 204,000 entries) was edited virtually singlehandedly by Matsui Kanji, who for years rose early to write thirty-three entries per day. Based on systematic collection of vocabulary from classical as well as contemporary sources, *Dai-Nihon kokugo jiten* innovated by including usage examples for its modern terms, which had precedent in literary wordlists such as the *Gagen shūran* but was new for a modern dictionary. A decade or so later, the largest of all Japanese dictionaries was published: *Daijiten* ('Great dictionary') (1934–6), in twenty-six volumes with a claimed count of 700,000 entries. Its great size is partly due to the inclusion of many proper nouns, but it also covers vocabulary from medieval and early modern sources that even now can be difficult to find elsewhere.

Dai-Nihon kokugo jiten, *Daijiten*, and *Daigenkai* were costly multi-volume sets, but a new class of large single-volume dictionaries, sized – and priced – more

⁹ For a survey, see Kurashima, *Kokugo to Kokugo jiten*.

accessibly, entered the market at mid-century, shortly after the Second World War. The first of these was the 1955 *Kōjien* ('Wide garden of words'), compiled by the historical linguist Shinmura Izuru (who was involved in the completion of *Daigenkai* among other earlier lexicographical projects) and his son Takeshi, a professor of French literature. Among the reasons for *Kōjien*'s immediate success were the elder Shinmura's reputation, the cachet of the publisher, and the quality of the definitions, but most significant were its size (about 200,000 entries); its coverage of the whole range of Japanese vocabulary, from classical texts to Chinese loanwords to recent coinages and technical terms; and its inclusion of encyclopedic content. All of these features made it suitable as an all-purpose reference.¹⁰ The seventh best-selling book of 1955, *Kōjien* remains the market leader in this category, but it has been joined by similar dictionaries from competing publishers, most prominently two edited by Matsumura Akira, *Daijirin* ('Great forest of words', 1988) and *Daijisen* ('Great fountain-head of words', 1995), both of which emphasize contemporary language and usage in their coverage and definitions.

Another significant mid-century development was the appearance of smaller dictionaries exclusively focused on contemporary language. This meant omitting classical vocabulary and obscure Sinitic loanwords, collecting current usages, and monitoring the emergence of new terminology. The pioneer of this category was the 1943 *Meikai kokugo jiten* ('Clearly explained national language dictionary'), largely the work of Kenbō Hidetoshi, although it was credited to an older and more famous scholar. True to its title, it was distinguished by the clarity of its definitions, but it also ordered words by actual pronunciation and provided pitch accents; most importantly, it included numerous contemporary terms that had never appeared in a dictionary.

Spawning many imitators in the following decades, *Meikai kokugo jiten* established the category of the small (fewer than 100,000 entries), portable dictionary of contemporary Japanese, and also directly gave rise to two important successors. Kenbō himself went on to edit the first four editions (1960–92) of *Sanseidō Kokugo jiten* ('Sanseido's national language dictionary'), which documented the current state of the language with clear and concise definitions, including many new words found through extensive collection of usages from popular media. Kenbō's former associate Yamada Tadao (a

¹⁰ See Shinmura Takeshi, *Kōjien monogatari*, and Shinmura Yasushi, *Kōjien wa naze umareta ka*.

distinguished historian of Japanese lexicography) became the driving force behind the 1972 *Shin meikai kokugo jiten* ('New clearly explained national language dictionary'), an entirely new work known for elaborate, innovative definitions.¹¹

At the pinnacle of post-war lexicography is *Nihon kokugo daijiten* ('Great dictionary of Japanese national language', 1972–6), the compilation of which was directed by Matsui Shigekazu. In 20 volumes covering about 450,000 entries, this was a monumental undertaking built on the foundation provided by the pre-war *Dai-Nihon kokugo jiten* (edited by Matsui's grandfather). The *Nihon kokugo daijiten* pays sustained attention to Tokugawa and Meiji period language, which had been comparatively neglected by modern lexicography, and includes examples from cited sources for all words. The earliest known usages are not always cited, but the unprecedented quantity and quality of these examples make it an unrivalled scholarly resource.

Dictionaries of Classical Japanese and Literary Sinitic

Two categories of modern wordlist overlap with the monolingual dictionary. The first is the classical Japanese dictionary (*kogo jiten*), ranging from portable single volumes to large multi-volume sets.¹² The former are aimed at high school students taking mandatory classical language classes, and cover words commonly found in literary works from the Heian period (794–1185) that are the core of the classical canon. The latter include more obscure vocabulary and have better coverage of the medieval and early modern periods. Wordlists of elegant or poetic terms (such as the *Gagen shūran*) were precursors of the *kogo jiten*, and a few modern exemplars were published before the war, but the genre did not take off until the 1950s, when the consolidation of the modern vernacular style and the segregation of classical language study into specialized secondary school course sequences created a new market for such focused references. Pre-war dictionaries such as *Genkai* had included classical vocabulary because it was still an important part of everyday written language, but by the post-war period this was no longer the case. The single-volume classical language dictionary, pioneered by Sanseidō's 1953 *Meikai kogo jiten* ('Clearly explained classical language dictionary'), emerged in

¹¹ On Kenbō and Yamada and their dictionaries, see Sasaki, *Jisho ni natta otoko*. The latter's work on lexicographical history includes Yamada, *Sandai no jisho*, and Yamada, *Kindai Kokugo jisho no ayumi*.

¹² Yamaguchi, 'Kogo jiten'.

tandem with the modern dictionary of contemporary language inaugurated by the same publisher's *Meikai kokugo jiten* a decade earlier.

A slightly later development in this category was the large-scale dictionary documenting the full breadth of pre-modern and early modern language, epitomized by the *Jidaibetsu Kokugo daijiten* ('Great dictionary of national language divided by historical period'), which was intended to provide a complete historical record of Japanese up to the nineteenth century. Sadly, several projected sections were never published, but those that were – one volume for the ancient period (1967) and five for the late medieval period (1985–2000) – are monuments of Japanese lexicography.

The second type of specialized wordlist is the *kanwa jiten* ('sinograph–Japanese dictionary').¹³ Through the Meiji period, versions of *Wagyokuhen* and similar wordlists satisfied the considerable demand for sinographic references, but the modern genre stems from the 1917 *Daijiten* ('Great character dictionary'). This large single-volume dictionary was primarily the work of Sakaeda Takei, but it appeared under the name of several editors, including Ueda Kazutoshi.¹⁴ The *Daijiten* collected more sinographs than its Japanese predecessors (nearly 15,000) and also stood out for its extensive and systematic treatment of multi-graph compounds and multiple readings of individual graphs. It was sorted in the traditional Chinese radical order culminated by the 1716 *Kāngxī zidiǎn* (for which see Chapter 6), but gave each sinograph a unique identifying number that enabled cross-references and indices by stroke count and Japanese reading. This new format was influential: especially in the post-war period, when publishers were responding to the same educational market that called for compact single-volume dictionaries of classical and contemporary language, numerous sinograph dictionaries modelled on *Daijiten* reached the market.

Among the foremost twentieth-century dictionaries is the *Dai kanwa jiten* ('Great sinograph–Japanese dictionary').¹⁵ This huge reference was compiled starting in 1925 under the direction of a scholar of Chinese classics who was close to Matsui Kanji and intended to emulate his *Dai-Nihon kokugo jiten*: Morohashi Tetsuji, whose surname is used informally to refer to his own dictionary. An abortive first volume was published in 1943, but the war led to the abandonment of the project until it was restarted a decade later; it was completed in twelve volumes plus an index (1955–60). Collecting

¹³ Atsugi, 'Kanwa jiten'.

¹⁴ Ueda, *Ueda's Daijiten*, is an edition of this work published in the United States in 1942 to meet the wartime need for a *kanji* dictionary.

¹⁵ See Kida, *Dai kanwa jiten o yomu*.

approximately 50,000 sinographs, with subentries for approximately 530,000 multi-graph compounds, it included rich citations of Chinese usages and their traditional commentarial glosses. Morohashi's dictionary inherited the *Daijiten*'s searching innovations in numbering graphs and providing multiple indices, and it also systematically incorporated the fruits of early modern Chinese lexicography.

There are practical benefits to a separation of 'classical Japanese' and 'literary Sinitic', but nonetheless this division distorts the history of Japanese language and inscription. Some pre-modern genres did have usage segregated along these lines (for example classical *waka* poetry excluded Sinitic vocabulary), but the mainstream was hybrid forms of writing that combined elements from both realms. Much vocabulary from pre-modern logographic but non-literary Sinitic genres appears in neither class of dictionary, an omission addressed by the more inclusive *Kogo taikan* ('Great survey of classical language', 2011–), with two of a projected four volumes published to date.

Market Contraction and Digitization

The expansion of Japanese lexicography in the modern period was caused by changes in the structure and scale of the education system, the rise of virtually universal literacy, and script and language reform. Similarly, the present and future of the field are shaped by ongoing social, economic, and media transformations. Falling populations of students have shrunk a market for dictionaries that had thrived until the early twenty-first century, and fewer publishers have the resources to produce new dictionaries or sustain continuing editions.

Another development is the decline of the physical paper book. As in other advanced economies, some genres are more resistant to digitization than others but, in Japan at least, dictionaries started this transition earlier and have taken it farther than most other classes of publication. In the latter half of the 1980s, CD-ROM dictionaries began to enter the market; the best known was a 1987 version of the third edition of *Kōjien*. In the 1990s they were joined by media for portable 'electronic book' players, but a more dramatic development was the emergence of 'electronic dictionaries' (*denshi jisho*), lighter and more compact calculator-sized machines with solid state memories, which caught on in the mid 1990s, with sales averaging well over two million units a year throughout the first decade of the twentieth century.¹⁶

¹⁶ On the early history of electronic dictionaries, see Nishikawa, 'Denshi jisho'. For *denshi jisho* sales figures, and more generally on the developments discussed in this section, see Iima, *Kokugo jiten no yukue*.

The next stage of digitization was more disruptive. In 1999 a commercial search site began providing an electronic version of *Daijirin* for free, which undercut the market for CD-ROM dictionaries and then, once the smartphone appeared in the late 2000s, demand for dedicated portable electronic dictionaries also began to dry up. Presently, data from *Daijirin* and *Daijisen* are freely available online, with Google providing abridged entries from a third dictionary (the *Iwanami kokugo jiten*) on its search page. Dictionaries can be purchased as smartphone apps, but free online services put a ceiling on their sales. Digitization has brought easier access and portability, and it enables lexicographers to update continually rather than waiting for new paper editions, but it also makes it difficult for already struggling publishers to stay in this market. The result will be the end of the rich variety of distinctive competing dictionaries that has been a hallmark of modern Japanese lexicography.

Recent years have also seen the emergence of aggregator databases combining many dictionaries and other references with digitized primary sources, allowing searches across all of them. The most prominent is JapanKnowledge, operated by a subsidiary of Shōgakukan (publisher of *Nihon kokugo daijiten* and *Daijisen*) since the early 2000s. Subscription fees make it unaffordable to most readers without university library privileges, and it is unclear whether such a model could fund the production of new dictionaries, but it portends renewed transformation of Japanese lexicography.

Korean

HEOKSEUNG KWON

Korea opened its ports to Japan in 1876 and signed treaties with Western countries during the 1880s and 1890s. With the opening of Korea to the Western world, Western missionaries established educational institutions in Korea. They recognized the relative value of Hangeul, the Korean alphabet, against Hanja, Chinese characters as used to write Korean, so they not only used Hangeul for their printed literature but also taught it to Korean young people. The missionaries also needed to learn the local language to carry out their religious missions in Korea. In the meantime, growing Korean nationalism, and the widespread social reforms of the 1890s known as the Gabo Reform, called for the use of Hangeul. In 1894, 450 years after its invention, Hangeul was adopted in official documents and school education.

Early Bilingual Dictionaries

Korean lexicography in its modern sense starts with the compilation of bilingual dictionaries in the second half of the nineteenth century. They were mostly compiled for the learning of Korean by foreign missionaries and diplomats who faced the need to communicate with Korean people.¹⁷

One early publication which lies outside the main missionary tradition of Korean lexicography was the *Translation of a Comparative Vocabulary of the Chinese, Corean, and Japanese Languages* printed in Batavia in 1835 by W. H. Medhurst (whose dictionary of Hokkien is discussed in Chapter 15), who may have been responsible for the translation. Its original must have been a multilingual glossary in the pre-1800 native tradition. One of the first missionary dictionaries was the *Dictionnaire français–coréen*, completed in 1869 by the French priest Stanislas Férron and never published. It presents 10,328 French headwords with Korean equivalents, and sometimes includes examples and information about register. The original manuscript is lost, but four manuscript copies are extant.¹⁸

In 1874, M. P. Putsillo, a Russian official and historian, published a Russian–Korean dictionary (*Opyt russko–korejskago slovarja*) in St Petersburg. He saw the urgent need for a word book while working for the resettlement of Korean immigrants in the Primorskaja Oblast of Russia, which shared a border with Korea. Although his dictionary was indebted to the *Translation of a Comparative Vocabulary*, it is a milestone in Korean lexicography: the first printed bilingual Korean dictionary.¹⁹ Putsillo's dictionary, 730 pages long, contains 3,369 Russian headwords and about 7,300 Korean equivalents in both Cyrillic transcriptions and Korean letters.

The French missionary Félix-Clair Ridel started work on a Korean–French dictionary in 1864. Upon its completion, one of the priests who served in Korea under Ridel, Eugène-Jean-Georges Coste, was sent to Japan in 1878 to have it printed. It was published in Yokohama in 1880 by Cerf Lévy, publisher of the French newspaper *L'Écho du Japon*, under the title *Dictionnaire coréen–français*, ascribed to 'Les missionnaires de Corée de la Société des missions étrangères de Paris' ('The missionaries in Korea of the Foreign Missionary

¹⁷ Duval, 'Korean lexicography', 904, gives an overview. Recent studies include Cho, 'Seoyang-in pyeonchan-ui gaechwagi daeyeog sajeon-e daehan jonghabjeog geomto'; Lee, 'Palioebangjeongyohoeui *Hanbuljaejon* (1880)gwa *Bulhanjaejon* (1869) bigyoyeongu'; and Song, 'Yeonghansajeon-ui yeogsawa sidaebyeol yeonghansajeon-ui teugjing'.

¹⁸ Kang, '19segi huban Joseon-e pagyeondoan pali oebangjeongyohoe seongyosadeul-ui *Bulhansajeon* yeongu', 190.

¹⁹ Kho, *Korean Studies in Russia*, 9; Hong, 'Ppujillo nohansajeon-e gwanhan ilgo'.

Society of Paris'). The first bilingual dictionary whose source language is Korean, it contains 27,194 Korean headwords, with information on pronunciation, some equivalent Chinese characters, and definitions in French. Korean headwords are listed in the order of the Roman alphabet, not the Korean alphabet, which makes the dictionary hard to use for Korean readers. The dictionary marked a new phase in Korean lexicography and became a model for other bilingual dictionaries in the years to come.²⁰

Born in London but educated in the United States, Horace Grant Underwood arrived in Korea as a missionary in 1885. His interest in Bible translation led him to write the first dictionary of Korean and English: *A Concise Dictionary of the Korean Language in Two Parts: Korean-English & English-Korean*, published in Yokohama in 1890. The first part, 196 pages long, has 4,839 Korean headwords and the second part, 293 pages long, has 6,712 English headwords. In the preface, Underwood writes that one of the great difficulties he faced in compiling the Korean-English part was the 'chaotic condition of Korean orthography', which forced him to resort to *Jeon-un-gpyeon* (see Chapter 10) as the basis of Korean spelling, comparing its spellings with those of the *Dictionnaire coréen-français*.²¹

The next decade was a productive period in the history of Korean bilingual lexicography. An *English-Corean Dictionary*, compiled by the British vice-consul James Scott, was published in 1891 by the Church of England Mission Press. The dictionary, 345 pages long, contains about 10,600 headwords with Korean equivalents, but no other information. *Parvum vocabularium Latino-Coreanum ad usum studiosae juventutis Coreanae* ('Little Latin-Korean dictionary for the use of the young students of Korea') containing about 9,000 Latin headwords, published in 1891 in Hong Kong, has been attributed to two different compilers, both French priests: either Marie-Nicolas-Antoine Daveluy (who taught Latin to Korean theology students) or Michel Alexandre Petitnicolas (who compiled a Latin-Korean dictionary). The original compilation has been dated as early as 1851, but a manuscript was burnt during the persecution of Christians in 1866, and the text that was eventually printed may have been reconstructed from materials that survived the persecution.²² The Canadian missionary James Gale, who had assisted

²⁰ Ha, 'Pyojejo daeyeog hanja-eoui tansaeng-gwa hanbuljaeon-ui gachi', 107.

²¹ Underwood, *Concise Dictionary of the Korean Language*, ii–iii.

²² For the attribution to Daveluy and the date of 1851, see Cho, 'Seoyang-in pyeonchan-ui gaehwagi daeyeog sajeon-e daehan jonghabjeog geomto', 61; for the attribution to Petitnicolas, see Ahn and Kim, 'Parvum vocabularium Latino-Coreanum', 14; for the burning of the manuscript, see Dallet, *Hangug cheonjugyohoesa*, 137.

Underwood in the Korean–English part of his work, published his own *Korean–English Dictionary*, also in Yokohama, in 1897. The first edition consists of two parts: ‘Korean–English Dictionary’, presenting the Korean headword in Hangeul characters followed by a note on vowel length, transcription in Hanja, and English definition, and ‘Chinese–English Dictionary’, presenting the Hanja characters first.²³ Gale acknowledged the *Dictionnaire coréen–français* and Underwood’s and Scott’s dictionaries as sources of his first part, and Giles’ Chinese dictionary (see Chapter 15) as a source of his second part.²⁴ According to the preface to the third edition, the first edition contains about 35,000 headwords, the second about 50,000, and the third about 75,000.²⁵ *Corean Words and Phrases: Being a Handbook and Pocket Dictionary for Visitors to Corea and New Arrivals in the Country* was published in 1897 by John Hodge. The second edition (1902) contains 2,647 English headwords. Besides providing Korean definitions, entries include English phrases and sentences with Korean pronunciation in Roman transliteration and Korean equivalents, for the use of visitors ‘confronted by the many perplexities of a stranger in a strange land’.²⁶ Finally, French–Korean lexicography continued in the form of the *Petit dictionnaire français–coréen* (1901), compiled by Charles Alévêque, who was a teacher of French. It contains about 7,000 French headwords, with Korean equivalents and information about their pronunciation.

These bilingual dictionaries were published at a time of rapid change. The Korean language lacked prescriptive rules of grammar, usage, and orthography, and the writing system used in official documents was changing from Hanja to Hangeul. In the absence of any language standardization, dictionary compilers sometimes provided different forms of Korean equivalents for a foreign headword, which in turn helps identify some linguistic features of the language at that time.

Monolingual Dictionaries and Linguistic Independence

The rise of cultural nationalism and heightened awareness of the national language led to attempts to codify and standardize Korean.²⁷ At the forefront

²³ Gale, *Korean–English Dictionary*, 1–836 (‘Korean–English’), 837–1096 (‘Chinese–English’).

²⁴ Gale, *Korean–English Dictionary*, ii; in hindsight, Gale saw the *Dictionnaire coréen–français* as the most important source of the first edition (Gale, *Unabridged Korean–English Dictionary*, [v]).

²⁵ Gale, *Unabridged Korean–English Dictionary*, [v].

²⁶ Hodge, *Corean Words and Phrases* (1902), unpaginated preface.

²⁷ Duval, ‘Korean lexicography’, 904–5, gives an overview.

of this move was the Korean grammarian Ju Si-kyeong, who had long been planning the compilation of a Korean monolingual dictionary to be called *Malmoi* ('Collection of words') and who emphasized the critical need for a Korean dictionary in his articles of 1897 and 1906 in *Tongnip Sinmun*, the first Korean and English newspaper. He believed that 'unified language and grammar would lead to the unity and solidarity of the Korean people'.²⁸

After the Japanese occupation of Korea in 1910, however, a new dictionary project which aimed to survey Korean customs and institutions was started in 1911 by the Japanese Government-General of Korea and completed in 1920. The compilation of *Joseon-eo sajeon* ('Joseon language dictionary') was intended as a means to effectively rule the country. This bilingual dictionary contained a total of 58,639 Korean headwords with Japanese definitions.²⁹

A monolingual Korean dictionary project based on Ju Si-kyeong's own orthography and grammar was started in 1911 by a group of scholars but was interrupted by Ju's death in 1914. The manuscript of the first book is extant. Interest in Korean lexicography was rekindled in 1929 when a committee for compiling a Korean dictionary was formed within the Society for Joseon Language Research. The society was reorganized in 1931 into Joseon-eo Hakhoe (Joseon Language Society), renamed in 1949 the Hangeul Hakhoe (Korean Language Society), to further the research and standardization of the language. The new society published 'Standardization of Korean Orthography' in 1933 and 'Corrected List of Standard Joseon Words' in 1936.

The dictionary project was also transferred to the new society, but the compilation of the dictionary came to a halt when the key members were thrown into prison in 1942 and all the manuscripts were confiscated by the Japanese police. In 1938, four years before this incident, the Japanese government had enacted a nationwide ban on the use and education of the Korean language, as a way of obliterating Korean culture and identity. The manuscripts, consisting of about 16,500 pages in 17 volumes, were discovered in a warehouse of Seoul Station in September 1945, immediately after Korea's independence from Japan.

The dictionary project was resumed promptly, and the first three volumes of *Joseonmal Keunsajeon* ('Big Korean dictionary') were published in 1947, 1949, and 1950. The compilation continued after the Korean War, and the fourth, fifth, and sixth volumes were published in 1957 under the title *Keunsajeon* ('Big dictionary'). The dictionary, 3,558 pages long, contains 164,125 headwords.

²⁸ Lee, 'Brief history', 136; see also Choi, 'Wongolo nam-eun choechoi ulimal sajeon, *Malmoi*'.

²⁹ Heo, 'Gug-eosajeon pyeonchan jeongchaeg mich geu yeogsa', 68.

This dictionary was a great cultural achievement of the new Korea that took almost thirty years of concerted efforts, ushering in a rebirth of the national treasure, Hangeul.

Meanwhile, during the period of Japanese rule, Sim Uilin created the first monolingual Korean dictionary, *Botonghaggyo Joseon-eo Sajeon* ('Elementary school Korean dictionary'), in 1925.³⁰ This learners' dictionary, containing 6,106 headwords mostly taken from elementary school Korean textbooks, was especially designed for elementary school students, teachers, and those learning Korean through self-study. Thirteen years later, Moon Seyoung published *Joseon-eo Sajeon* ('Korean language dictionary'). In the preface to the dictionary he wrote, 'A country with a history of 5,000 years, we are disgraced by the absence of a dictionary.' He saw dictionary compilation as a practical action of intellectuals showing the excellence of Korean and the pride of the nation. He is said to have collected words from 1917 to 1928 and written definitions from 1929 to 1936, before editing and proofreading the dictionary with the help of Lee Yoonjae. *Joseon-eo Sajeon* is the first Korean dictionary that used the Joseon Language Society's 'Standardization of Korean Orthography', contributing to the spread of the standard language. The dictionary, 1,634 pages long, contains a vocabulary of about 100,000 words.

After the Korean War

Many large Korean dictionaries came to be published by commercial publishing houses. *Gug-eo Saesajeon* ('New Korean dictionary'), edited by the Society of Korean Language and Literature, was published in 1958. The dictionary, 1,002 pages long, contains about 150,000 words. In the same year *Pyojun Gug-eosajeon* ('Standard Korean dictionary'), edited by Shin Gicheol and Shin Yongcheol, was published. In 1975, these two brothers published *Sae Urimal Keunsajeon* ('New big Korean dictionary'), 3,856 pages long, containing about 310,000 words. The most successful dictionary was *Gug-eo Daesajeon* ('Grand Korean dictionary'), which was edited by Yi Heeseung and published in 1961. The one-volume dictionary, which was reprinted more than thirty-two times until 1982, contains 257,854 headwords. Its revised second edition of 1982, 4,504 pages long, was expanded to 420,000 words. These and other dictionaries produced by commercial publishers are said to have been based on *Keunsajeon*.³¹

³⁰ Pak, 'Sim uilin-ui *Botonghaggyo Joseon-eo Sajeon* (1925) ui bunseog', 112.

³¹ Yoo, 'Hangug-eosajeon pyeonchan-ui hyeonhwang-gwa ilonjeog jeongae', 190.

In 1967, ten years after the completion of *Keunsajeon*, the Hangeul Hakhoe planned a new project to supplement it, for instance by including more scientific and technical terms. Although the compilation process was interrupted several times, *Urimal Keunsajeon* ('Big Korean dictionary'), boasting more than 450,000 headwords, came to be published in four volumes over 1991 and 1992. In 1996 a CD-ROM edition was released under the title *Hangeul Urimal Keunsajeon*.

Corpus-Based Korean Dictionaries

The year 1986 heralded a new era of Korean lexicography. In that year the Korean Lexicographic Society, reorganized in 2002 as the Korean Association for Lexicography (KOREALEX), was founded at Yonsei University and initiated the idea of compiling a Korean dictionary on the basis of corpora. The society also began publishing its journal *Sajeon peonchanhak yeonku* ('Studies in lexicography', retitled the *Journal of Korealex* in 2003). The 1990s were a turning point in dictionary compilation and publication. Since then much research has been carried out on modern Korean lexicography.³²

Yonsei Hangug-eosajeon ('Yonsei Korean dictionary'), published in 1998, was a pioneering project, because at that time everything from corpus building to corpus analysis tools for data extraction had to be made from scratch in-house. Another dictionary project, begun in 1992 at Korea University, led to the publication of three volumes of *Goryeodae Hangukeodaesajeon* ('Korea University grand dictionary of the Korean language') in 2009. The dictionary is based on a corpus of 100 million words and contains 386,889 headwords in 7,535 pages. The move from commercial to academic institutional publication contributed to the development and quality of dictionary compilation.

The National Institute of Korean Language started a corpus-based Korean dictionary project in 1992, a year after its founding. After eight years of intensive work, the most authoritative Korean dictionary, *Pyojun Gug-eo Daesajeon* ('Grand dictionary of Standard Korean'), was published in 1999. It contains about 480,000 words and purports to be the norm for orthography, grammar, usage and foreign language notation, but it also incorporates a wide variety of words such as jargon, old words, dialectal words, and North Korean words. A revised edition containing about 510,000 words was released in 2008, but in an electronic version only. Since then, the dictionary

³² See Heo, 'Gug-eosajeon pyeonchan jeongchaeg mich geu yeogsa', and Heo, 'Gug-eosajeon pyeonchan yeongusa'.

has become available on the web, updated every three months, and recently through mobile apps. Ironically, the publication of this dictionary by a national institute adversely affected dictionary publishing. It became the acknowledged authority on standard Korean usage, sweeping the dictionary market.

The Korean War and the division of Korea accelerated the linguistic discrepancies between South and North Korea. Scholars on both sides saw the need to launch a project as a way of narrowing down language differences between the two Koreas.³³ The project for compiling a unified Korean dictionary, titled *Gyeoremal Keunsajeon*, was first proposed in 1989. Both parties signed a letter of intent about the project in 2004, and compilation started in 2005, with publication then expected in 2019. However, the compilation progress was brought to a halt in 2010 for political reasons.

The advent of the Internet and rapid advance of new technologies have changed the lexicography of Korean just as they have changed that of other languages. Since 1999, monolingual and multilingual Korean dictionaries have become available online.³⁴ At first these dictionaries were electronic versions of paper dictionaries, but later various search functions and multimedia elements were added to online dictionaries. Korean dictionaries have moved rapidly into the digital medium and are now evolving into more advanced versions.

Vietnamese

JOHN D. PHAN

In Vietnam, as in the rest of East Asia, the language of administration, religion, intellectual discourse, and the arts was Literary Chinese. This means that the crafting of dictionaries in Vietnam developed largely in tandem with an effort to comprehend and compose a language distinct from the common speech (and one shared with contemporary Korea, China, and Japan). While monolingual lexicographic materials in Literary Chinese have a long history in Vietnam, bilingual (Literary Chinese to vernacular Vietnamese) dictionaries became possible only with the development of the *Chữ Nôm* 字喃 vernacular script (hereafter, *Nôm*), which adapted Chinese characters to represent the Vietnamese language.³⁵ With

³³ Han, 'Features and procedures', 6; see also Lee, 'North Korean lexicology and lexicography'.

³⁴ See Kim and Han, 'Development of Korean online dictionaries'.

³⁵ Phan, 'Rebooting the vernacular'.

the development of *Nôm*, a variety of lexicographic texts began to emerge primarily to support the study of Literary Chinese.

Lexicography Before the Nineteenth Century

The earliest Vietnamese texts with dictionary-like functions were glossaries and interlinear translations of Literary Chinese texts. The earliest known separate wordlist containing Vietnamese is a Chinese text called the *Annan yiyu* 安南譯語 ('Annamese translations'; Vietnamese *An nam dịch ngữ*), which was probably produced around the beginning of the seventeenth century by the Ming Dynasty Interpreters' Institute (for the activities of which, see also Chapter 10) and was transmitted as one in a set of thirteen glossaries called the *Huayi yiyu* 華夷譯語 ('Sino-Barbarian translations'; Vietnamese *Hoa di dịch ngữ*).³⁶ The text is composed of 716 basic Chinese words transliterated into Vietnamese via an approximately homophonous Chinese character (not to be confused with the native *Nôm* script) and may have been used as a kind of crib sheet for diplomats. It uses a conventional organizational structure based on thematic categories: 'Heavenly patterns' 天文 ('thiên văn'), 'Earthly patterns' 地理 ('địa lý'), and so on.

This convention was also adopted by Vietnamese lexicographers, notably in the most significant of their works before the nineteenth century, the metrical dictionary *Chỉ nam ngọc âm giải nghĩa* 指南玉音解義 ('Explication of the guide to jewelled sounds'). Probably compiled by the seventeenth-century Queen Trịnh Thị Ngọc Trúc 鄭氏玉竹, and presented as an expanded explication of an older text supposedly dating to the Hàn dynasty, the *Chỉ nam* glosses Literary Chinese entries into vernacular Vietnamese *Nôm*, with occasional pronunciation hints given in a style similar to that of the *Annan yiyu*. Unlike the *Annan yiyu*, the purpose of this dictionary was almost certainly to help in reading and writing Literary Chinese; however, a significant side effect of its compilation was the broadening and enrichment of a vernacular literary mode.³⁷ The dictionary was an unprecedented tool for bridging Literary Chinese and vernacular Vietnamese, and the conventions of multi-topic thematic categories and a poetic metre would remain hallmarks of Vietnamese lexicography until the advent of a Romanized alphabet.

³⁶ Maspero, 'Études sur la phonétique historique de la langue Annamite', 8.

³⁷ See Phan, 'Chữ Nôm and the taming of the South', and Phan, 'Rebooting the vernacular'.

Historical Context for Nineteenth–Twentieth-Century Lexicography

The *Chi nam ngọc âm giải nghĩa* proved greatly popular, but it seems exceptional for its era. During most of the seventeenth century, classical literacy waned as Vietnam (then called Đại Việt) was split between the ‘old kingdom’ in the north, and a new southern realm.³⁸ By the turn of the nineteenth century, however, Đại Việt was reunified under Nguyễn Phúc Ánh 阮福暎, who became the first Nguyễn emperor.³⁹ The unification of a new dynasty meant a new centralized focus on the standardization of various literary traditions, including lexicographic texts. Education in Literary Chinese and the civil service exams (a system of standardized tests for political appointment) were, furthermore, reinvigorated by Nguyễn Phúc Ánh’s immediate successor, providing renewed demand for dictionaries to help in the study of the Confucian classics. Finally, the increasing presence of Westerners, culminating in French colonization in the 1860s, produced its own series of European-style dictionaries, using a Romanized script called Quốc Ngữ. French colonization ended with the conclusion of the First Indochinese War in 1954, after which Vietnam was partitioned into northern and southern halves. With the end of the Second Indochinese War and the victory of Communist forces in 1975, the country was reunified. A large number of South Vietnamese resettled, primarily in France and the United States, continuing an independent lineage of lexical practices outside Vietnam.

Traditional Lexicography in the Nineteenth Century

The major impetus for lexicographic production in the nineteenth century was to master Literary Chinese in preparation for the civil service examinations or to read Buddhist scripture. Traditional elementary education in early modern Vietnam was conducted at the village level.⁴⁰ While students at the secondary level concentrated on reading full texts such as the classics, philosophical treatises, documents on law and administration, and religious literature, students at the elementary level used a range of books which presented basic Literary Chinese knowledge. Often, these textbooks came to

³⁸ See Taylor, ‘Literacy in early seventeenth-century northern Vietnam’.

³⁹ Baldanza, *Ming China and Vietnam*, 1–6.

⁴⁰ Nguyễn, ‘The *Sanzijing* (三字經) and primary education’, 34.

be glossed with vernacular Vietnamese translations using *Nôm*, and thus functioned essentially as dictionaries of Literary Chinese.

The *Sơ học vấn tân* 初學文津 ('Literary ford for early education'), published in 1874, is an example of textbook–dictionary hybridization. In content, it is a textbook covering Chinese and Vietnamese history. In form, however, the text is arranged into four-character Literary Chinese phrases, each translated into vernacular Vietnamese *Nôm*. This essentially makes each four-character Literary Chinese line a dictionary entry. Other texts imported from China also functioned in this way, due to glosses added during the nineteenth century. The best-known of these was the *Sanzijing* 三字經 ('Three character classic'; Vietnamese *Tam tự kinh*), attributed to the thirteenth-century statesman and polymath Wang Yinglin 王應麟, which was widely used at the elementary level in nineteenth-century Vietnam. A number of Vietnamese versions of the *Sanzijing* bear vernacular glosses, as well as full translations.

Beyond these hybridized lexicographic textbooks were true glossaries and primers produced locally, such as the *Tam thiên tự* 三千字 ('Three thousand characters'), attributed to the eighteenth-century literatus Ngô Thì Nhậm 吳時壬, which is essentially a Sino–Vietnamese lexicon formed of strings of bilingual four-character phrases, pairing basic Literary Chinese terms with their equivalent in vernacular Vietnamese (*Nôm*). Literary Chinese entries and their vernacular equivalents alternate, and a linked structure of inter-linear rhymes facilitated rote memorization of the text. Take, for example, the opening, provided here in modern Vietnamese transcription (Chinese words in uppercase, Vietnamese equivalents in lowercase), with a translation:

THIÊN trời ĐỊA đất	HEAVEN heaven EARTH earth
CỬ cất TỒN còn	LIFT lift PERSIST persist
LỤC sáu TAM ba	SIX six THREE three
GIA nhà QUỐC nước	HOUSE house STATE state

A wide range of versions circulated during the nineteenth century.⁴¹ Other glossaries using a similar model also emerged, including the *Nhất thiên tự* 一千字 ('One thousand characters') and later, the *Ngũ thiên tự* 五千字 ('Five thousand characters'). Note that similarly titled glossaries, but with different lexical inventories, also exist in the Chinese tradition.

The format of a rhyming glossary was popular enough for the fourth Nguyễn emperor, Tự Đức 嗣德, to produce a similar text (published in 1883) titled *Tự Đức thánh chế tự học giải nghĩa ca* 嗣德聖制解義歌 ('Tự Đức's

⁴¹ Nguyễn, 'Vietnamese phonology and graphemic borrowings', 164.

sagely regulated explicatory verses’) or simply the *Tự học giải nghĩa ca* 自學解義歌 (‘Explicatory verses for self-study’), divided into seven thematic sections: nature, society, government/education, tools, grasses and trees, birds and beasts, and vermin. Thus, like the *Chi nam ngọc Âm giải nghĩa* before it, the *Tự Học Giải Nghĩa Ca* belongs to the most popular category of lexicographic materials in traditional Vietnam, which Trần Trọng Dương calls ‘multi-topic categorical dictionaries’.⁴² These were not designed to be used as look-up tools, since characters were not classified by radical or any other indexical system besides the broad thematic categorization (despite the fact that Chinese precedents for such classification were well known). Instead, the thematic organization and metrical form suggest that students learned their contents by heart, thus studying vocabulary and content across a variety of fields in one stroke.

Perhaps the best-known multi-topic categorical dictionary of the nineteenth century is Phạm Đình Hổ’s 範廷琥 *Nhật dụng thường đàm* 日用常談 (‘Common speech for daily use’), compiled between 1821 and 1827, and comprising 32 thematic sections, across which are distributed some 2,480 individual entries. Within each category, entries are headed by a Literary Chinese term in large font, followed by a smaller-print prose definition in vernacular Vietnamese Nôm. Despite its name, the dictionary was not meant for daily speech, but contains many abstruse literary terms. It was probably one of the most popular glossaries for Confucian education during the nineteenth century and was reprinted at least nine times.⁴³

The popularity of multi-topic categorical dictionaries continued even into the late colonial period but, perhaps unsurprisingly, died off once the civil service examinations were terminated in 1919 – and with them, an entire educational system based on Literary Chinese.

Colonization, European Dictionaries, and European Influence on Vietnamese Lexicography

European missionaries had been developing dictionary materials to help facilitate their own religious and political objectives in Đại Việt as early as the seventeenth century. Missionary work naturally required language proficiency, and the linguistic analyses conducted by early missionaries laid much of the groundwork for modern Vietnamese linguistics. The earliest

⁴² Trần, *Khảo cứu từ điển song ngữ Hán Việt Nhật dụng thường đàm*, 117; Trần, 中世紀越南漢字詞典的類型與特點 (*Zhong shiji Yuenan hanzi cidian de leixing yu tedian*).

⁴³ Trần, *Khảo Cứu Từ Điển Song Ngữ Hán Việt Nhật Dụng Thường Đàm*.

important work centred on developing a Romanized script for Vietnamese and was conducted by Portuguese missionaries.⁴⁴ Two Portuguese missionary priests, Gaspar do Amaral and Antonio Barbosa, created a bidirectional Portuguese and Vietnamese glossary, now lost, in the early seventeenth century. Drawing heavily on their work, the French missionary Alexandre de Rhodes systematized various Romanized transcriptions of Vietnamese to produce the earliest printed dictionary using the Vietnamese alphabet now called Quốc Ngữ: the trilingual *Dictionarium Annamiticum Lusitanum et Latinum* ('Annamese–Portuguese–Latin dictionary'; for which see also Chapter 29), published in 1651. The work of these missionaries introduced a new kind of dictionary to Vietnam, focusing on spoken language and, crucially, organized alphabetically.

By the nineteenth century, French and Vietnamese Catholics produced a number of dictionaries of Vietnamese and European languages. Notably, the missionary Jean-Louis Taberd produced a Vietnamese–Latin dictionary in 1838, *Dictionarium Anamitico–Latinum* ('Annamese–Latin dictionary'), and a companion, *Dictionarium Latino–Anamiticum* ('Latin–Annamese dictionary'), which had a thematically ordered English–French–Latin–Vietnamese wordlist as an appendix. The *Dictionarium Anamitico–Latinum* was actually Taberd's edit of a dictionary compiled, by 1773, by his confrere Pierre Pigneau de Behaine, best known for his long friendship with Nguyễn Phúc Ánh and his efforts to help establish the Nguyễn dynasty.⁴⁵ The dictionary is arranged alphabetically using Quốc Ngữ, but (unlike de Rhodes' dictionary) each entry is headed by a printed Nôm character, followed by a Quốc Ngữ transcription and a Latin definition. Taberd's work was in turn expanded upon by J. S. Theurel (1877), and Theurel's recension was in its turn used as the base for a Vietnamese–French dictionary of 1904, P. G. Vallot's *Petit dictionnaire annamite–français*.⁴⁶ The *Dictionarium Anamitico–Latinum* is introduced by a long discussion of the Vietnamese language, including an analysis of its tonal system.⁴⁷ Most dictionaries that followed would also include a structural analysis of the language, especially as it related to choices in alphabetical orthography, and these analyses were deeply influential in the development both of modern Vietnamese orthography and of modern Vietnamese linguistic analysis.

⁴⁴ Jacques, *Portuguese Pioneers of Vietnamese Linguistics*.

⁴⁵ Mai Quốc Liên, 'Lời nói đầu', in Taberd, *Dictionarium Anamitico–Latinum* (2004).

⁴⁶ For the Theurel edition, see Phan, *Mission and Catechesis*.

⁴⁷ Taberd, *Dictionarium Anamitico–Latinum* (1838), vi–vii.

The Behaine family of dictionaries was not the only effort at this time. The first Vietnamese–French dictionary appears to have been produced by a missionary named Théophile Legrand de la Liraye in 1868.⁴⁸ Legrand de la Liraye also produced a lexicon of Chinese characters glossed with their Sino-Vietnamese pronunciations, *Pronunciation figurée des caractères chinois en mandarin annamite* (1875). The preface of the dictionary remarks that ‘European dictionaries of Chinese only give the phonetics of the hieroglyphs [Chinese characters] in Mandarin Chinese’, and that the lack of a proper guide to Vietnamese pronunciations of Chinese characters leads to great embarrassment.⁴⁹ The lexicon itself is noteworthy for its collection of formal Sino-Vietnamese pronunciations, but also for its inventory of Chinese characters, many of which appear to be variant forms specific to Vietnamese usage.

In 1877, the Mission de Saigon published a small dictionary titled *Dictionnaire annamite–français* by Marie-Antoine-Louis Caspar, which apparently did not use Nôm characters but relied only on Quốc Ngữ. In 1893, Jean-François-Marie Génibrel produced a *Vocabulaire annamite–français*, based on the framework of Caspar’s dictionary, but adding Nôm characters; in 1898 he published his own *Dictionnaire annamite–français*, calling it a second edition on the grounds that it was fundamentally a revision of Caspar’s. Like Taberd’s, Génibrel’s dictionary was arranged alphabetically, but each entry was also headed by the Nôm character for the word, followed by its Quốc Ngữ transcription and a French definition. The title page describes the Vietnamese language as ‘la Langue Annamite vulgaire’ – the Annamese vernacular – and promises a second volume which would provide the Chinese characters ‘necessary to study the Tứ Thư, or Four Books of classical Chinese’.⁵⁰ It is clear that Génibrel conceived of the Vietnamese language as consisting of a vernacular as well as a classical variety – in the form of Literary Chinese. Likewise, Jean Bonet’s *Dictionnaire annamite–français (langue officielle et langue vulgaire)* (1899–1901) distinguishes purely native Vietnamese vocabulary; the numerous Chinese loanwords that were now to be considered integral to the Vietnamese lexicon; and Literary Chinese words, pronounced

⁴⁸ Nguyễn, ‘Vietnamese lexicography’, 1–2.

⁴⁹ Legrand de la Liraye, *Pronunciation figurée*, 3, ‘Les dictionnaires européens chinois ne donnent la phonétique des hiéroglyphes qu’en mandarin chinois. On se trouve donc en Cochinchine très embarrassé.’

⁵⁰ Génibrel, *Dictionnaire annamite–français*, title page, ‘comprenant 1° tout les caractères de la Langue annamite vulgaire . . . 2° les caractères chinois nécessaires à l’étude des Tứ Thư ou Quatre livres classiques chinois’.

according to Vietnamese phonology, and ‘used since time immemorial for the composition of official documents’.⁵¹

Of particular significance is the work of Vietnamese lexicographers such as Pétrus Trương Vĩnh Ký 張永記 and Paulus Của. As evident by their names, both were Catholic converts with connections to the French, and both produced dictionaries seeking to elevate the status of the Quốc Ngữ alphabet. Pétrus Ký published a wide array of texts for learners of both French and Vietnamese, including a well-known *Dictionnaire français–annamite* of 1885, with a supplement, *Vocabulaire annamite–français*, in 1887. This text, which targeted French and other European students of Vietnamese, makes use of a pared-down version of the traditional thematic arrangement, employing three broad categories, ‘Heaven’, ‘Earth’, and ‘Man’, followed by a short list of religious terms. Within each broad category are a number of thematic subcategories, and within these each entry is ordered alphabetically. The *Vocabulaire annamite–français* thus expresses an intellectual loyalty to traditional forms of Vietnamese lexicography, and a belief in their practical usefulness – while at the same time embracing the alphabetical principle.

Paulus Của’s two-volume *Dictionnaire annamite: Đại Nam quốc âm tự vị* (‘Annamese dictionary: character lexicon of national pronunciations of Đại Nam’, 1895–6), includes Nôm characters as well as Quốc Ngữ transcriptions. Although Của originally intended to compose a Vietnamese–French dictionary, on the advice of the director of the School of Interpreters in Saigon, he decided to focus on systematizing the Quốc Ngữ alphabet, producing an unprecedented monolingual Vietnamese dictionary instead.⁵² The dictionary is organized alphabetically, but each entry is headed by a Nôm character, which is then defined in vernacular Vietnamese written in Quốc Ngữ. Của’s work self-consciously attempts to standardize Quốc Ngữ, and in many instances common multiple spellings are identified in the text as non-standard variants. For example, under an entry for a word spelt *bực*, Của writes ‘see *bặc*’, thus establishing many of the orthographic conventions which survive to the present day.

The works of Petrus Ký and Paulus Của were influential, and a number of early twentieth-century dictionaries built on them, including the work of the

⁵¹ Bonet, *Dictionnaire annamite–français*, ii, ‘Les vocables d’origine purement annamites . . . Les mots de provenance chinoise dont les peuples d’Annam ont enrichi leur idiome par des emprunts successifs . . . Les termes de la langue écrite chinoise prononcés à la manière annamite et employés depuis un temps immémorial pour la rédaction des documents officiels.’

⁵² Bùi Đ. T., ‘Lời giới thiệu’ (‘Introduction’) in Của, *Đại Nam quốc âm tự vị* (1998).

French priest Gustave Hue, who lauded the richness and complexity of Vietnamese in his *Dictionnaire vietnamien–chinois–français* (1937) and criticized those who derided it as simple or impoverished when compared with French.⁵³ This invigorated interest in the vernacular language dovetailed with the meteoric rise in the popularity of Quốc Ngữ, which was accepted as the de facto script of modern Vietnamese by the 1930s.⁵⁴

The missionary lexicographical tradition thus met with a fertile climate of nationalist interest in the vernacular language. Many of the pioneers of this time were Vietnamese intellectuals who had inherited traditional forms of education and had also been exposed to contemporary French and other European modes of knowledge. Towards the end of the colonial period, a number of Vietnamese scholars produced erudite dictionaries combining these multiple epistemological and methodological traditions. Of these, one of the most important was Đào Duy Anh 陶維英. Although critical of traditional education and culture, he was also disturbed by what he perceived to be a rushed and unexamined abandonment of the past.⁵⁵ This educated ambivalence – itself emblematic of Vietnamese late colonial intellectuality – is neatly expressed by his two great lexicographic achievements: one Sino–Vietnamese dictionary, and one French–Vietnamese dictionary, based on the *Larousse du 20e siècle*. The former, *Hán Việt từ điển* ('A Sino–Vietnamese dictionary') is a dictionary of Literary Chinese (including modern neologisms); however, it is organized alphabetically. Notably, each character is transcribed into Quốc Ngữ using its orthodox Sino–Vietnamese pronunciation – another reminder that Literary Chinese was viewed as the Vietnamese classical language well into the twentieth century, as well as of the persistence of classical erudition into the twentieth century, even in the case of so great a champion of change as Đào Duy Anh. Definitions are in vernacular Vietnamese, written in the now-dominant Quốc Ngữ. Đào Duy Anh's French–Vietnamese dictionary (the *Pháp-Việt Từ Điển*) was published later (1936; ²1958), and was essentially a translation of the *Larousse de 20e siècle*. At the time of its release, it was perhaps the most exhaustive dictionary of Vietnamese that had ever been produced. Đào conceived of these two dictionaries as companions to one another, with Literary Chinese and French as intellectual languages that hovered to either side of an enriched and expansive vernacular Vietnamese language. In many ways, therefore, his

⁵³ Nguyễn, 'Vietnamese lexicography', 3. ⁵⁴ Phan, 'Secularization of the sinograph'.

⁵⁵ McHale, *Print and Power*, 84–5.

work unites the epistemological modes of traditional and European lexicography, which converged in colonial Vietnam.

These works set the stage for future refinement, expansion, and systematization of the Vietnamese language. This is especially true starting in the tumultuous years of the First Indochinese War, and may be seen in monolingual dictionaries or orthographic guides such as Nguyễn Văn Minh's 1941 *Văn-liệu Từ-điển* ('Dictionary of literary materials') or Lê Ngọc-Trụ's *Việt-ngữ chánh-tả tự-vị* ('Orthographic lexicon of Vietnamese').⁵⁶ The trend to expand and systematize the Vietnamese language grew during the partition of the Second Indochinese War, alongside a revived cultural interest in traditional literature and language that led to new scholarship and new lexicographic materials of Sino-Vietnamese and especially Chữ Nôm.

Today, the dictionary industry is thriving, and dictionaries are produced in large numbers in multiple languages, by national research institutes and also numerous private and semi-private publishers. Post-war diaspora intellectuals have continued to maintain a distinct lexicographic lineage, producing dictionaries and lexica of both modern Vietnamese and Chữ Nôm in the United States, France, and elsewhere, in print and online. In Vietnam, national organizations dedicated to the Vietnamese language, linguistics, and lexicography such as the Institute of Linguistics (*Viện ngôn ngữ học*; established 1968) and the Institute of Lexicography and Encyclopedias (*Viện từ điển học và bách khoa thư*; established 2008) both conduct academic research into fields relevant to lexicography, and also produce large numbers of dictionaries for the regular market. Finally, the Institute for Sino-Nôm Research (*Viện nghiên cứu Hán-Nôm*; established 1979) reprints pre-modern lexicographic materials as well as producing new dictionaries to help in the study of Literary Chinese and vernacular Vietnamese Chữ Nôm.

⁵⁶ Nguyễn, 'Vietnamese lexicography', 3–4.

Turkish and Persian from c. 1700

LUCIANO ROCCHI AND ARTHUR DUDNEY

This chapter picks up the stories which were told in Chapter 11, but with two differences of approach. Whereas the great importance of Maḥmūd al-Kāšgarī's *Dīwān Luġāt at-Turk* requires attention to the early lexicography of the Turkic languages as a group, modern Turkish can be treated in isolation from its cousins, as it is here. And whereas the period covered in Chapter 11 ended at a point when Persian shared a writing system with Turkish, and when Persian was a language of high prestige in the Ottoman empire, the two languages have grown apart culturally, so that the present chapter is divided into two sections, one on Turkish and one on Persian.

Turkish

LUCIANO ROCCHI

(TRANSLATED BY MASSIMO VERDICCHIO)

The Eighteenth Century

After the great period of the seventeenth century, which culminated with the publication of the monumental *Thesaurus* of F. Meninski (see Chapter 11), with the next century we enter, as it were, a slow phase, with an Ottoman Turkish lexicographical output of no great significance – at least not until we reach the last decade of the century. In fact, looking at the first half of the century, only two works are worth remarking; curiously, they both came out in the same year, 1730.

The first, called *Grammaire turque*, was printed by İbrāhīm Müteferriḳa, a diplomat of Hungarian origin who had converted to Islam and who founded the first printing shop in Istanbul in 1729; he published seventeen works in the span of fourteen years.¹ The work appeared anonymously but

¹ Watson, 'İbrāhīm Müteferriḳa', 436.

we can attribute it with certainty to Jean-Baptiste Holdermann, a French Jesuit who declared explicitly that he was the author in a letter addressed to his confreres.² Holdermann's work contains a section on grammar; a Turkish–French glossary ('Recueil des noms & des verbes'); a phraseological section; and, finally, about twenty dialogues in both languages. The glossary takes up more than a quarter of the entire volume (59 pages out of 200) and comprises a little fewer than 1,300 nouns, subdivided into many lists according to specific semantic fields, with the addition of about a hundred verbs. The Turkish lemmata are printed both in the Arabic–Ottoman alphabet and in the Latin alphabet, following orthographic rules that recall those of French. The transcription in Latin characters makes it possible, as usually happens in *Transkriptionstexte*, for us to account for certain phonetic phenomena that the traditional Arab–Persian alphabet keeps hidden. For instance, the word for 'dragon' is rendered in Ottoman letters with *ajder* (faithfully following the Persian source *aždar*), while in Latin letters it is rendered with *acder* (<adgedér>).³ This way we have solid proof that the sound [ʒ], foreign to the Turkish phonetic system, shifted to [dʒ] in the common pronunciation. Holdermann's lexical repertory is also important from the historical point of view, giving us the first known evidence of certain words (see below).

The second work that came out in 1730 is the Latin–Turkish–German *Compendiosum lexicon* of Johann Christian Clodius, a work closely connected to Meninski's, as the author clearly states in the preface, even if it is substantially less voluminous than the Polish lexicographer's. The entries are arranged in the alphabetical order of the Latin lemmata, while the Turkish equivalents are registered both in the Arabic–Ottoman alphabet and in the Latin alphabet. Consulting the work and searching for words in it are fortunately made much easier thanks to a *triplex index* added at the end of the volume, which consists of a dual list of Turkish words present in the dictionary (first in the Ottoman alphabet and then in the Latin) and of a third list of the German equivalents, with page references. On the basis of these indices the Turkish entry count amounts to more than 15,000 individual entries. Clodius' work covers more or less the same material as Meninski's (however, mostly with a different transcription), but we also find new entries. For instance, to confine ourselves to those of Italian origin, Clodius registers *ambra* 'amber' (apparently a *hapax legomenon*; < Italian *ambra*); *kolaz(i)yon* 'afternoon snack' (< Venetian *colaziòn*, standard Italian *colazione*); *kumkuma*

² Menz, 'Idioms and dialogues', 147. ³ Holdermann, *Grammaire*, 82.

'kind of container for liquids' (< Italian *cuccuma*); and *çurma* 'galley slaves' (not in Meninski; < Venetian *zurma*, standard Italian *ciurma*).⁴

In the last decades of the eighteenth century we find two important grammatical-lexicographical works on the Turkish language, both penned by European clergymen who conducted their missionary activity in the Ottoman empire, the Italian Bernardino Pianzola and the Frenchman Pierre François Viguier. Pianzola's work came out in three consecutive editions (*Dizionario grammatiche* 1781, *Dizionario grammatiche* 1789, *Grammatica dizionarij* 1801), personally edited by the author, who made many changes. The first edition is divided into two volumes. One contains a Greek-Italian dictionary; grammatical notes concerning Italian, vulgar Greek, and Turkish; and a series of dialogues in these languages. The other contains a quadrilingual dictionary (Italian, Latin, Greek, and Turkish) in the alphabetical order of the Italian lemmata. In the next edition, in three volumes, Pianzola takes out the Latin (both in the grammatical exposition and in the dictionary) and makes ample changes in the dialogue section. Finally, in the last edition, in four volumes, the most important new development is the addition of a Turkish-Italian section. The whole Turkish lexicon present in Pianzola's work (taking into account all three editions) comprises about 6,200 lemmata.⁵ In short, we could say that the Italian missionary is an excellent witness to the Ottoman language spoken at the time, since he gives many forms typical of the oral register, characterized by phenomena such as syncope, anaptyxis, metathesis, epenthesis, the dropping of sounds, and so on. Furthermore, his lexicon is full of loanwords from European languages, many of which do not seem to be attested in other sources.⁶ Finally, I would like to point to an interesting lexeme which is also not documented elsewhere: *asilmamlık* 'independence'. This seems to be a magnificent example of a word whose coinage follows the dictates of twentieth-century linguistic reform *avant la lettre*, to express the concept which in Ottoman is usually rendered with the Arabism *istiklal*.

Pierre François Viguier, the Prefect Apostolic in Istanbul during the last quarter of the eighteenth century, was the author of a work that also contained a grammatical section, a phraseological-dialogical one, and a wordlist, namely the *Éléments de la langue turque*, but, unlike Pianzola's, it dealt only with Turkish and not with other languages. The lexicographical

⁴ Rocchi, 'Italianismi', 891, 898, 923.

⁵ Rocchi, *Il lessico turco nell'opera di Bernardino Pianzola*, is a critical edition with commentary.

⁶ See the selection in Rocchi, 'Importance', 97–101.

section, which occupies 102 double-column pages, is titled 'Essai de vocabulaire François–Turc', and consists of more than 8,000 lemmata, according to Viguier himself.⁷ The Turkish words are transcribed in Latin characters on the basis of the rules of French orthography. Overall, Viguier's lexicon reflects the literary language, but there are also, here and there, examples of the spoken language. As for loans from European languages, their number is not as large as that of those registered by Pianzola, but some are very important from the cultural point of view, for instance *kazeta* 'gazette' (= modern Turkish *gazete*), a first Turkish lexicographical example of this important Italianism.⁸

As for lexicographic works compiled by Turks, these are essentially repositories of Persian or Arab vocabulary translated and illustrated with Turkish equivalents. The single exception is due to Mehmet Es'ad Efendi, an Ottoman religious (*Şeyhülislâm*) who also studied poetry and music. Between 1725 and 1732, he compiled a lexicon in which, for the first time in Ottoman history, Turkish words constitute the headwords and the equivalents are Arabic and Persian. The work, *Lehceü'l-luğât*, was printed only in 1795 (and again in 1801), decades after the death of the author. The lemmata discussed by Es'ad Efendi amount to fewer than 3,700 and they belong in large part to the original Turkish lexical stock. There are, however, also loans from Western languages, for example, *kalafatlamak* 'to caulk'; *marangoz* (which Es'ad Efendi indicates as *lafz-ı Rûmî* 'Greek word') 'ship's carpenter'; and *urba* (< Italian *roba*) 'dress'.⁹

Nineteenth-Century Monolingual Lexica

At the beginning of the nineteenth century, the first so-called *ğalaât* glossaries begin to appear, that is, lexical lists in which the 'erroneous' form (*ğalat*) – that is, the one actually spoken – of an Arabic or Persian word is compared to its correct literary form.¹⁰ It goes without saying that these 'glossaries of errors' are of great value for historical research into the integration of the Arab-Persian element to popular Turkish pronunciation; they constitute, *mutatis mutandis*, a kind of Turkish counterpart to the *Appendix Probi* in the Romance field.

⁷ Viguier, *Éléments*, 457, 'Le Vocabulaire, ainsi réduit à une espèce de Dictionnaire usuel, est l'assemblage de plus de 8000 mots effectifs.'

⁸ Stachowski, *Lexique turc dans le 'Vocabulaire' de P. F. Viguier*, 166.

⁹ Es'ad Efendi, *Lehce* (ed. Kirkkılıç), 373, 477, 677. ¹⁰ Tietze, 'Fremden Elemente', 105–6.

But the major turning point in the development of a Turkish lexicography, in the modern sense of the word, was constituted by the *Tanzimat* ('Reforms'), the name with which one refers to that great process of political, economic, and social modernization of the Ottoman empire that was accomplished on the Western model by the sultans Abdülmecid I (1839–61) and Abdülaziz (1861–76). This process of reform also, naturally, involved the educational system, where ample space was given to the teaching of the Turkish language. Article 18 of the constitution promulgated in 1876 prescribed explicitly that the officers of the state ought to know the Turkish language: 'A prerequisite for Ottoman subjects employed in the service of the State is that they should know Turkish, which is the official language of the State.'¹¹ One of the most important representatives of the political and intellectual world of the time, Ahmet Vefik Paşa, who was minister of education, grand vizier for two brief periods, and president of the first Ottoman parliament in 1877, and also founded the first Turkish theatre, was also the author of the first monolingual Turkish dictionary, *Lehce-i Osmânî*, published in 1876. It is divided into two sections: the first (*cüz-i evvel*) contains words of non-Arabic or Persian origin, and the second (*cüz-i sâni*) contains Arabisms and Persianisms. The importance of this work is not only historical but also more properly lexicographical because, along with colloquialisms and provincialisms, it also includes a great number of foreign terms of Western origin, more than 600 in all.¹² Several of them are of great documentary value: for example, among Anglicisms we find *gonbot* 'gunboat' and *ıskaru* 'screw (of a propeller)', and among Italianisms *bilyeto* 'ticket' (< Italian *biglietto*) and *salsa* 'sauce' (< Italian *salsa*).¹³

However, Sami Frashëri, known in Turkey as Şemseddin Sami, is the scholar to whom we owe the most important work in this field. An Albanian intellectual and philosopher, he was one of the leading lights of the Albanian National Awakening. He moved to Istanbul in 1872, worked in the government press office, and came into contact with the Ottoman circles whose members were fighting to reform institutions at the local and national level.

¹¹ 'Tebâ'a-i Osmâniyenin hidemât-ı devlette istihdam olunmak için devletin lisân-ı resmîsi olan Türkçeyi bilmeleri şarttır' (Lewis, *Language Reform*, 16).

¹² Yıldız, 'Vefik Lehce'.

¹³ Vefik, *Lehce* (ed. Toparlı), 161, 185, 55, 330. Of these words, *gonbot* represents a first short-lived adaptation from English, and the form which became usual in standard Turkish is *gambot*; *ıskaru* is an isolated form (Redhouse, *Lexicon*, III registers the word as *ıskuru*) and is not normally used; *bilyeto* and *salsa* are still close to Italian *biglietto* and *salsa*, but the obligatory forms in everyday Turkish are *bilet* and *salça* (the latter probably via Greek *saltsa*).

Although he died at the young age of fifty-four, he left more than fifty works: literary (novels and plays), political, scientific, and linguistic. Among these, his lexicographic works are most important: a French–Turkish dictionary and a Turkish–French one; a big encyclopedia in six volumes (the first of its kind printed in Turkey); and, above all, the one that has rightly become his most famous work, the Turkish *Kâmûs* – the Arabic word *qāmûs* meaning ‘ocean’ (from which the Turkish word is derived) has also come to mean ‘dictionary, lexicon’ following the great Arabic dictionary of al-Firūzābādī, *Qāmûs al-Muḥīṭ* (see Chapter 8). Sami’s work represents the highest point of Ottoman Turkish lexicography after the *Tanzimat* and immediately before the end of the empire. It is a noble attempt to give the Turkish people a ‘national’ dictionary that, as Sami explains in the Introduction, gathers, first of all, the ‘genuinely Turkish’ words, obtained not only by consulting written sources but also by travelling throughout the various regions of the empire and talking to people of all social classes. To these are added Arabic and Persian words only to the extent that they were part of the spoken language, excluding those rare and obscure words typical of ‘Ottoman’ dictionaries. It is important to note that in the title Sami employs the adjective *türki* and not *osmanlı*. ‘Sami was indisputably the most competent Ottoman scholar using the Turkish language. His “Turkish lexicon” . . . , still a standard reference for studies involving Ottoman Turkish, was a monumental achievement for its time with its technical rigor and its careful deployment of modern linguistic protocols.’¹⁴

Nineteenth-Century Bilingual Lexica

The most important bilingual lexica are still those by European scholars, of which three are, for their size, particularly distinguished. The first is the Turkish–French *Dictionnaire turc–français* by J. D. Kieffer and T. X. Bianchi. The original manuscript of the work was by Jean Daniel Kieffer, who was for many years the secretary-interpreter of the French royal court for Oriental languages and professor at the Collège de France. After his death in 1833 it was taken up and reworked by another French scholar, Thomas Xavier Bianchi, who had already published French–Turkish dictionaries. As the subtitle indicates, the work is destined for ‘consuls and diplomats, business people, seafarers, and other travellers to the Levant’.¹⁵ However, the extent of the

¹⁴ Ersoy, ‘Şemseddin Sami’, 33–4.

¹⁵ Kieffer and Bianchi, *Dictionnaire turc–français*, title page, ‘à l’usage des agents diplomatiques et consulaires, des commerçants, des navigateurs et autres voyageurs dans le Levant’.

vocabulary, with the inclusion of a massive adstratal body of Arabisms and Persianisms, and of numerous terms of European origin, makes this dictionary a useful instrument of work even for academics.

The second of these dictionaries is the *Dictionnaire turc-arabe-persan/Türkisch-arabisch-persisches Handwörterbuch* of the German Orientalist Julius Theodor Zenker. The title 'concise dictionary' (*Handwörterbuch*) given to it by the author must not deceive us, nor must the relatively limited number of pages (984). Every page has the text printed on three columns and the type is very small (which makes reading very difficult); the total word count is about 45,000 entries. Zenker's work not only contains Turkish elements together with Arabic and Persian elements, as is clear from the title, but is also distinguished from other compilations of its kind by the large number of Chagatay terms (classified by the author as *turc-orientale*, 'East Turkic'), so that it is of interest not only to Ottomanists but also to Turcologists in the widest sense of the term.

The dictionary that surpasses all the others by far for its wealth of vocabulary is the famous Turkish-English *Lexicon* of Sir James William Redhouse, which came out in 1890. The author had a troubled childhood and adolescence. As an orphan at a very young age, he was expelled from school, and took a job as a deckhand on a merchant ship that landed at Istanbul in 1826.¹⁶ There he found work in the Naval School, where he learned Turkish and other Oriental languages. Later he entered the service of the Ottoman government; he acted as interpreter and translator for both the Sublime Porte and the English Foreign Office, and had many diplomatic assignments. As for his activity as a scholar, after he had published some grammars and lexica of minor importance, the American Board of Commissioners for Foreign Missions commissioned what would become his *magnum opus*. 'It was computed that the work would contain about 100,000 Turkish words of Turkish, Arabic, Persian and European origin'¹⁷ – these words from the preface to his *Lexicon* suffice to illustrate the veritable lexical mine contained in this work. It can be said without fear of contradiction that Meninski's *Thesaurus* and Redhouse's *Lexicon* are the two monuments with which anyone who undertakes the study of Ottoman-period lexicography must reckon.

I cannot conclude this section without alluding to a work that, even though less voluminous than the others, still has some importance, namely, the dictionary by Charles A. C. Barbier de Meynard, which came out in 1881

¹⁶ Findley, 'Redhouse', 574–5. ¹⁷ Redhouse, *Lexicon*, xi.

and in 1886. The French Orientalist identified it in the subtitle, with too much modesty, as a ‘supplement to the dictionaries published to this day’.¹⁸ In fact it is an excellent account of the lexicon spoken in Turkey towards the end of the nineteenth century, and its entries are phraseologically rich and full of information not to be found in other authors.

It should be added that in all the above-mentioned bilingual lexica, the Turkish terms are given both in Arabic-Ottoman and in Latin orthography (the latter according to orthographic norms specific to each author). It goes without saying that this situation allows us to draw precious indications as to the actual pronunciation of the words in question.

Twentieth- and Twenty-First-Century Monolingual Lexica

With the fall of the Ottoman empire at the end of the First World War, and the birth of the Republic of Turkey, we witness a profound transformation of Turkish society in all areas, thanks to a series of radical reforms promoted by the founder of the new state, Mustafa Kemal Atatürk. The most important of these reforms was to the language: the Turks called it *Dil Devrimi*, the ‘Linguistic Revolution’. At first, the alphabet was changed. The traditional Arabic-Ottoman one was abolished and, starting from 1928, one based on Latin characters was introduced. In the following years, with the foundation of *Türk Dil Kurumu* (the ‘Turkish Language Association’), an enormous work of ‘purification’ of the language was undertaken, according to the dictates of Atatürk himself. The vast number of lexemes of Arabic and Persian origin that had become part of the Turkish vocabulary over the centuries were replaced by ‘purely Turkish’ (*öztürkçe*) words, both by drawing from spoken Anatolian and by the creation of neologisms, many of which proved to be ephemeral. ‘The often bizarre, sometimes tragicomic, but never dull story of the Turkish language reform’ has been brilliantly described by Geoffrey Lewis.¹⁹ However, this reform had indubitable merits, such as encouraging the Turks to take an interest in their language and to promote lexicographical activity.

Starting from the 1940s, the *Türk Dil Kurumu* set itself the goal of providing the Turkish nation with a normative dictionary that could equal in scientific rigour those of the more evolved Western societies. The project

¹⁸ Barbier de Meynard, *Dictionnaire*, title page, ‘Supplément aux dictionnaires publiés jusqu’à ce jour’.

¹⁹ Lewis, *Language Reform*; the phrase cited is from p. 1.

was completed in 1945 with the publication of the first edition of *Türkçe Sözlük* ('Turkish dictionary'). Other editions followed, at intervals from three to ten years. At the time of writing (2017) the most recent one to be printed is the eleventh (2011). Every edition was revised, amplified, and updated. The word count of the first edition ran to a total of 25,574 entries (32,104 words), and with the most recent edition we have arrived at 77,005 entries (92,292 words). To be sure, the work of revision has not only concerned the addition of new words: the structure of entries has also undergone important changes. For instance, in the seventh edition (1983), most of the entries were enriched by a phraseological part and by literary quotations by the most important Turkish writers. Again, in the first seven editions, only the language of origin of loanwords was given, while in the eighth edition the etymology is quoted in its complete form.

Among the most important editorial initiatives of the Türk Dil Kurumu has been the publication of two large-scale lexical collections. The first of these is a Turkish historical dictionary that came out at first in four volumes between 1943 and 1957 as *XIII Asırdan Tarama Sözlüğü*, then, in a completely revised edition, in eight volumes between 1963 and 1977 as *XIII. Yüzyıldan Tarama Sözlüğü*. The purpose of the work was to gather words no longer in usage in the spoken language and belonging to the lexicon of 'pure Turkish' – thus excluding words of foreign origin such as Arabisms and Persianisms – on the basis of a corpus of Ottoman literary texts that covers a period from the thirteenth to the nineteenth centuries. The number of these texts, which was 160 in the first edition, was brought to 227 in the second. The lemmata are given in transcription according to the modern Turkish alphabet, but in the literary quotations that are attached to each entry the headword in question is also given in Arab-Ottoman script. The second is a dialect dictionary of Anatolian Turkish, *Türkiyede halk ağzından söz derleme dergisi*, which first came out from 1939 to 1957 in six volumes. The first three contained the main lexical corpus, the fourth a supplement, the fifth a reverse index, and the sixth 'folkloristic words'. This work also came out in a greatly enlarged second edition in twelve volumes, *Türkiye'de halk ağzından söz derleme sözlüğü*, published between 1963 and 1982.

With the advent of the computer age, the Türk Dil Kurumu has established its own site on which digitized texts of several of its dictionaries are available, both those I have mentioned and others of a more specialized nature. The wordlists of all these dictionaries have been collected in a single enormous online lexicon, *Büyük Türkçe Sözlük*, that can be freely consulted and that, at the present time (2017), runs to more than 600,000 words. As far as

the substandard lexicon is concerned, we should mention *Argo* by Hülki Aktunç, who is not a professional linguist but a poet, a writer, and an essayist, a passionate scholar of his own language. It handles a repertory of 5,000 lemmata belonging to the argotic register, with etymological information and, in many entries, illustrative quotations from popular literature.

Twentieth- and Twenty-First-Century Bilingual Lexica

Of course this is not the place to make a list, even a summary one, of the numerous bilingual Turkish dictionaries that came out after the reform of the language. So, I will only mention some of those works that seem particularly relevant. In first place on chronological grounds, I have to mention the Turkish–Italian lexicon of the Orientalist Luigi Bonelli that came out in 1939. As far as we know, it is the first work published in Europe which took the lexicon of the ‘new Turkish’ (*Yeni Türkçe*) that began to appear in the 1930s into critical consideration. Bonelli was rightly prudent in welcoming these neologisms, including them in the wordlist, but marking them with an asterisk and referring them to the Ottoman Turkish entry that they should have replaced (for instance, *alan* = *saha* ‘field, area’, *anlam* = *mana* ‘meaning’, and *araç* = *vasıta* ‘means’ have been established in usage, but others have turned out to be ephemeral, such as *alpay* = *fatih* ‘conqueror’, *anlatık* = *fıkra* ‘anecdote’, and *aran* = *itidal* ‘moderation’).

Towards the middle of the twentieth century, a group of scholars, among whom the name of Andreas Tietze stands out, conceived the ambitious project of preparing a new edition, revised and updated, of the classic Turkish–English dictionary by Sir James Redhouse. The *New Redhouse Turkish–English Dictionary* came out in 1968, and constitutes the first serious attempt to present both the Ottoman lexicon and the modern Turkish one in a single fully comprehensive volume. As the preface states, ‘the intention has been to include every word, and as nearly as possible every set phrase or locution, that has been used in standard Turkish . . . in the last two hundred years’ (viii). On the basis of this enormous work, the authors themselves prepared another of more modest proportions aimed at illustrating the lexicon of present-day Turkish, the *Redhouse Contemporary Turkish–English Dictionary*, which was subsequently re-elaborated and substantially enriched in a new edition edited by Serap Bezmez and C. H. Brown. This latter version of

Redhouse, the *Redhouse Turkish–English Dictionary* of 1999, represents, in my view, one of the best reference books available to translators and to anyone else who is interested in the Turkish language.

Finally, I must mention the work of Karl Steuerwald, a German linguist who lived for a long time in Turkey, where he held many academic positions. His outstanding *Türkisch–Deutsches Wörterbuch*, which came out in 1972 and in a revised and extended edition in 1988, offers extensive coverage of the obsolete Ottoman lexicon (classified as ‘veraltet’ or ‘historisch’), and it is a model of scientific lexicography, with entries accurately structured in their various meanings and enriched by excellent phraseological material.

Twentieth- and Twenty-First-Century Historical-Etymological Dictionaries

On the situation of the etymological lexicography of Turkish, which was not very brilliant until the end of the twentieth century, please allow me to refer to one of my previous articles.²⁰ I will confine myself here to the two major works published in the last few years. The first is the monumental etymological lexicon of Andreas Tietze, *Tarihi ve Etimolojik Türkiye Türkçesi Lugatı*, which has had a somewhat turbulent editorial history. The first volume of the work, comprising the entries from A to E, came out in 2002 (during the lifetime of the author, who died in 2003) as a joint publication of the Istanbul publishing house Simurg and the Austrian Academy of Sciences in Vienna; the second (F–J) followed seven years later, in 2009, as a sole publication of the Viennese academy. Then, for reasons of which I am not aware, the latter chose not to proceed with the remaining volumes, and the project passed into the hands of the Turkish Academy of Sciences, which decided to republish the entire work from the beginning. The first four new volumes (A–L; the entries from A to J are reprinted without revision from the previous edition) appeared in June 2017 (although the date of the imprint is 2016), and the publication of the other five, to make up a total of nine, is due to take place in 2018 and 2019. Tietze’s work is a worthy climax to the learned activity of the great Austrian scholar, which will, I believe, for a long time remain a required reference book in the field of Turcology. The second is the work of Sevan Nişanyan, a scholar and writer of Armenian origin, who edited a historic-etymological dictionary of contemporary Turkish, published in a paper edition as *Sözlerin Soyağacı* in 2002 and in an online edition from 2007. The

²⁰ Rocchi, ‘Turkish historical lexicography’, 203–6.

importance of this work consists above all in the fact that, for the first time in Turkish lexicography, the words are given a historical depth, since it provides the date of their first occurrence. From this point of view, the consultation of the online version is indispensable; it is regularly updated, so that many entries appear redated on the basis of new research conducted by Nişanyan.²¹

Persian

ARTHUR DUDNEY

The extensive corpus of Persian-language dictionaries compiled in the eighteenth and nineteenth centuries was largely produced in South Asia, and was with few exceptions concerned with poetic language. This second section of Chapter 17 traces the development from these literary lexica to the first comprehensive dictionary compiled in Iran through modern lexicographical methodology, namely the *Lughatnāmah* project of ‘Alī Akbar Dihkhudā. In the eighteenth century, dictionaries engaged with the key question in contemporary literary aesthetics, namely how to integrate recent innovation into a literary tradition whose authority was derived from a corpus of classical poets. The tradition worked through *asnād* (the singular is *sanad*), namely citations demonstrating that authoritative poets had used a word or expression in a particular meaning. Most but not all dictionaries – *Burhān-i Qāṭi* ‘(Decisive proof, 1651: see Chapter 11) being a conspicuous exception – included such quotations with every entry. Lexicographers were concerned both with correcting the work of earlier scholars as well as with integrating new words and expressions into their dictionaries. European efforts to produce an accurate Persian–English dictionary, again largely concentrating on British India rather than Iran, bore fruit in the late nineteenth century. This chapter is not concerned with dictionaries relating to particular works (for example, those defining obscure words in early classics such as Sa‘dī’s *Gulistān* or Firdawsī’s *Shāhnāmah*), but such texts were common, and the connection between lexicography and literary commentary that they imply is notable.²²

²¹ However, a careful analysis of early documents will make further important antedatings possible: for instance, the Italianism *dama* ‘draughts, checkers’, the Hispanism *kanarya* ‘canary’, and the Persianism *maslahatgūzar* ‘chargé d’affaires’ are dated by Nişanyan to 1876 (from Vefik, *Lehce*), but the first is attested in Bernardo da Parigi (1665) (Rocchi, ‘Bernardo da Parigi’, 265), the second in Holdermann, *Grammaire*, 79, and the third in a transcribed text of 1567 (Bombaci, ‘Lettera turca’, 139).

²² These are discussed in detail (in Persian) in Naqavī, *Farhangnīvisī*, 181ff.

The Eighteenth Century

Viewed from Iran, the eighteenth century was a time of stagnation in Persian-language lexicography, but on the Indian subcontinent this was arguably the golden period for critical reflection on the meaning of words. In India, lively debates about authority in the Persian literary tradition were being adjudicated in part through lexicography. Lexicographers worked to correct the mistakes of their predecessors, particularly those of the seventeenth century, but also to include the work of recent poets (*al-mut'akhhirūn*) as authorities.²³ The difference in the leading contemporary Persian literary style, known as *tāzah-go'ī* ('fresh-speaking') or *ṭarz-i tāzah* ('the fresh style'), from classical Persian poetry was notable. This divergence presented literary scholars with a problem in applying standards to tell good poetry from bad. Dictionaries functioned as repositories of correct usage and loomed large in literary debates.²⁴ In many cases, the goal on the part of those in favour of *tāzah-go'ī* was to knit the recent tradition to its classical past by arguing that what seemed to be a recent innovation in usage had actually been used by a classical poet.

The most important figure in the debates over aesthetic innovation was arguably the philologist Sirāj al-Dīn 'Alī Khān Ārzū, who spent most of his working life in Delhi. As a lexicographer, he was engaged both with correction and expansion of the lexicon. Ārzū's *Sirāj al-Lughat* ('Lamp of language', 1734) is significant because it revises *Burhān-i Qāṭi* and another important seventeenth-century dictionary, *Farhang-i Rashīdī* (for which see Chapter 11), as well as drawing on many others. Ārzū reckons *Farhang-i Rashīdī* to be the most carefully compiled dictionary available and *Burhān-i Qāṭi* the most thorough.²⁵ Manuscripts of *Sirāj al-Lughat* are relatively rare and it has never been printed, though it is of obvious scholarly value because of its corrections to some of the key works in the earlier lexicographical tradition.²⁶ The work's second volume, *Chirāgh-i Hidāyat* ('The lamp of guidance'), is about one-fifth the length and is usually treated as a stand-alone work. By contrast to the first volume of *Sirāj al-Lughat*, it is very common in manuscript. In the nineteenth century, it became even more widely known when it was lithographed in the margins of *Ghiyās al-Lughāt* (see below). *Chirāgh-i Hidāyat* accounts for words that do not appear in earlier lexica and whose

²³ Tauer, 'Learned literature', 431. ²⁴ Kinra, 'Cultures of comparative philology', 248.

²⁵ Quoted in Naqavī, *Farhang-nivīsi*, 111.

²⁶ Blochmann, 'Contributions to Persian lexicography', 27.

meanings are either unfamiliar or debated. It grapples with particularly Indian usages and adjudicates whether these should be permissible or not.

Ānand Rām Mukhliṣ and Ṭek Chand Bahār, two Hindu bureaucrats who were friends of Khān-i Ārzū, also compiled important dictionaries in Delhi around this time. Mukhliṣ wrote *Mir'āt al-Iṣṭilāḥ* ('Mirror of expressions') in 1745.²⁷ It differs from other pre-modern Persian dictionaries in that it contains a large proportion of what we could call 'proto-anthropological' observations that are anchored in experience rather than in poetic language. (Nonetheless many of these words are still illustrated with poetic quotations as in other dictionaries.) Mukhliṣ had a particular interest in administrative terminology as well as in words and expressions having to do with painting, clothing, handicrafts, animals, flowers, hot beverages (particularly coffee), games, and so on. The work also contains adages and long digressions describing, for example, particular people that Mukhliṣ knew such as Ārzū (briefly defining the term *ārzū* as 'hope and desire' in as many words serves as an excuse to launch into several hundred words of praise for his friend and teacher) or objects such as the Peacock Throne.

Ṭek Chand Bahār was a friend of both Ārzū and Mukhliṣ. In compiling *Bahār-i 'Ajam* ('The spring of Persian'), he drew extensively on Ārzū's work. He revised *Bahār-i 'Ajam* seven times between 1751 and 1781, and greatly increased its scope into one of the longest dictionaries then available.²⁸ A crucial source that came to Bahār's notice between revisions was *Muṣṭalahāt al-Shu'arā* ('Expressions of the poets') by Siyāl Kotī Mal Vārastah, who was also a Hindu. *Muṣṭalahāt al-Shu'arā* was based on original research rather than being a revision of previous dictionaries. Many scholars have incorrectly inferred that Vārastah's reference to fifteen years' research with Iranians demonstrates that he lived in Iran for that period.²⁹ A better explanation is that he had access to Iranian emigres and manuscripts in Lahore (present-day Pakistan) where he lived. *Bahār-i 'Ajam* made such liberal use of Vārastah's dictionary that it effectively absorbed it.³⁰

²⁷ A modern edition was published in 2013. ²⁸ Naqavī, *Farhangnivīsī*, 110.

²⁹ Rieu, *Catalogue*, II.503; Naqavī, *Farhangnivīsī*, 158. Vārastah states that the title of the work is a chronogram for the year he started compiling the work (1180 = AD 1766) but does not give the year he finished it. According to one source, he died the same year. There is no reason to assume, *pace* Blochmann, 'Contributions to Persian lexicography', 30, that compilation of the work took fifteen years or that those years were spent in Iran.

³⁰ Blochmann, 'Contributions to Persian lexicography', 29.

Multilingual Dictionaries

Because New Persian (namely the language that has been in use for the past millennium) was formed in part by an influx of Arabic words into Middle Persian, the Persian lexicographical tradition has always in some sense been multilingual. Here I discuss dictionaries that define Arabic words in Persian as well as developments in presenting Indic words as part of Persianate discourse. European dictionaries of Persian are discussed below in their own section, but the scope of this chapter necessitates leaving out other language pairings such as Persian–Turkish; nor can it include specialized lexica such as *materia medica* or taxonomies of metaphorical language. However, it is worth noting in passing that the rhymed wordlist or *niṣāb* was a form that had currency in Persian from the thirteenth century into the twentieth. A *niṣāb* was a list of words presented in poetic form to allow for easy memorization by students. *Niṣābs* were often bilingual in Persian and Arabic, but in other parts of the Persian-using world they incorporated Indic, Turkish, and European words. In nineteenth-century India, there were Persian-language *niṣābs* of English, such as the *Niṣāb-i Inglīṣī* (1853).³¹

Persian–Arabic dictionaries, or, rather, dictionaries that define Arabic words in Persian, were commonly compiled in this period, not only since formal Persian prose contains a significant proportion of Arabic-derived vocabulary but also because Arabic remained the language of certain kinds of technical writing, such as philosophy, wherever Persian was used. Indeed, among Persian users in the eighteenth century in South Asia the key axis of the process that might be called ‘vernacularization’ was not translating Persian texts into Indic languages, as the modern perspective might expect, but rather making Arabic texts available in Persian translation for wider accessibility. Arabo-Persian lexicography largely took the form of commentaries on and translations of the fourteenth-century Arabic dictionary *Qāmūs al-Muḥīṭ* by Majd al-Dīn Muḥammad ibn Ya‘qūb al-Fīrūzābādī (see Chapter 8), who like many Arabic-language lexicographers and philologists was himself ethnically Persian.³² The *Tarjumān al-Lughat* (1705) by Muḥammad Yaḥyā bin Muḥammad Shafīq Qazvīnī, compiled in Iran, and the *Qābūs* (1737) by Muḥammad Ḥabībullāh Isfahānī Qannawjī Dihlavī, compiled in India, are two important examples of commentarial translations of the *Qāmūs*.³³ Like their source, they are organized in rhyme order rather than alphabetically. In the nineteenth century, the *Muntahā al-Arab fī Lughāt*

³¹ Rieu, *Supplement*, 121. ³² Haywood, *Arabic Lexicography*, 83–90.

³³ Munzavī, *Farhangnāmah-hā*, 238–40.

al-‘Arab (‘The final need for Arabic words’, 1825), compiled by ‘Abd al-Raḥīm Ṣafīpūrī under the auspices of the East India Company’s Fort William College in Calcutta, despite its errors, became the most important resource for subsequent dictionaries. It is in alphabetical order, as had become common in monolingual Persian dictionaries by this point.³⁴

In the eighteenth century, wordlists in Persian defining Indic terms, such as Mīrzā Khān’s *Tuhfat al-Hind* (‘India’s gift’, c. 1700), became common because Persian was the link language shared by Indian scholars (see Chapter 18). Ārzū’s *Navādir al-Ālfāz* (‘Wonders among words’, 1743) is usually taken to be the first critical dictionary of Urdu but it is, according to its author, a lexicon of ‘Indic words that people of the provinces say rather than the Persian, Arabic or Turkish [synonyms]’. It is therefore a specialized lexicon of Persian that consists of vernacular words that have entered the Persianate cultural sphere. *Navādir al-Ālfāz* connects with Ārzū’s larger philological programme. For example, the entry for *ast* ‘is’ clearly is intended to elucidate the fundamental nature of language rather than simply define the word.³⁵ Various wordlists compiled for the East India Company that used Persian as a link between a given Indian language and English demonstrate the continued importance of Persian as a glossing language in the colonial period. For example, *Muṣṭalahāt-i Thagān* (‘Expressions of the Thugs’, c. 1834) formed the basis for Colonel William Sleeman’s *Ramaseeana, or, A Vocabulary of the Peculiar Language Used by the Thugs* (1836).³⁶

Nineteenth-Century Compendia

As in the eighteenth century, lexicographers in the nineteenth century continued to engage with the classical tradition by revising earlier dictionaries and illustrating meanings with poetic quotations. Lexicography remained prestigious in South Asia, as is suggested by the fact that the *navāb* (ruler) of Awadh, Ghāzī al-Dīn Ḥaydar Shāh, wrote – or more likely put his name on – a dictionary called *Haft Qulzum* (‘Seven seas’), which was largely derived from *Burhān-i Qāṭī*.³⁷ Indeed, many dictionaries appeared with *Burhān-i Qāṭī* as their main source, which is unsurprising given how many times it was printed in India.

³⁴ Perry, ‘Dictionaries: ii. Arabic/Persian’. ³⁵ Ārzū, *Navādir al-Ālfāz*, 23.

³⁶ Sachau, Ethé, and Beeston, *Catalogue*, I, cols. 1023–4; other examples in Rieu, *Catalogue*, II.517, and Ethé, *Catalogue*, I, col. 1359.

³⁷ Naqavī, *Farhangnīvīsī*, 217.

The most important dictionary of the early nineteenth century was *Ghiyās al-Lughāt* by Ghiyās al-Dīn Muḥammad Jalāl al-Dīn bin Sharīf al-Dīn Rāmpūrī. It was compiled in 1826 in Rampur (modern-day Uttar Pradesh, India). This dictionary became ubiquitous because it was useful for students in its presentation and balanced classical and recent usage.³⁸

The strangest engagement with *Burhān-i Qāṭi* in this period was Mīrẓā Asadullāh Khān Ghālib's *Qāṭi-i Burhān* ('Chopping of the proof', 1861), which takes some 400 words and idioms from the seventeenth-century dictionary and criticizes them on aesthetic grounds, particularly as Indianisms unknown to the 'true Persian' of Iran. Ghālib was the most accomplished Urdu and Indian Persian poet of his time, and he claimed to have a knowledge of contemporary Iranian idioms on the basis of having been taught by 'Abd al-Ṣamad, a historically untraceable, and likely imaginary, Iranian tutor.³⁹ Ghālib's foray into scholarship was ill advised. Since Ghālib offered force of personality in place of evidence in support of his idiosyncratic judgements, *Qāṭi-i Burhān* was roundly criticized by his contemporaries.⁴⁰ The work is significant in demonstrating the continuing importance of the earlier lexicographical tradition, but many of Ghālib's interpretations were smacked down in the acrimonious pamphlet war that followed.⁴¹

With approximately 50,000 entries, the *Farhang-i Ānandrāj* (1889) is the largest dictionary compiled in the Persianate world before the twentieth century.⁴² Besides its length, the work is remarkable because of where it was compiled, namely at the edge of the Persian-using world, in what is today the south-eastern Indian state of Andhra Pradesh. The dictionary is dedicated to the maharajah of Vizianagaram, Ānand Gajapatī Rāj. The collation of every available Persian dictionary, including some European sources, was undertaken by the maharajah's chief secretary (*mīr munshī*) Muḥammad Pādshāh b. Ghulām Muḥī al-Dīn, known by his penname Shād. Because of its completeness, it was important as a source for Dihkhudā (see below).

Remarkably, one cannot point to a single important comprehensive dictionary compiled on Iranian soil in the two centuries after Muḥammad Qāsim Surūrī's *Majma' al-Furs* (Isfahan, 1599). However, the Iranian

³⁸ Blochmann, 'Contributions to Persian lexicography', 31; Dabīr Siyāqī, *Farhang-hā-yi Fārsī*, 176–8; there is an edition of 1958 by Muḥammad Dabīr Siyāqī (about whom see below).

³⁹ Ḥālī, *Yādgar-i Ghālib*, 13–14; Pritchett, *Desertful of Roses*, commentary on Ghazal 139, verse 1.

⁴⁰ Blochmann, 'Contributions to Persian lexicography', 19–20.

⁴¹ Dabīrsīyāqī, '*Borhān-e qāṭe*'.

⁴² Baevskii, '*Farhang-e Ānandrāj*'; Shād, *Farhang-i Ānandrāj* (1956–8), is an edition.

lexicographical tradition that was dormant in the seventeenth and eighteenth centuries re-emerged somewhat in the nineteenth century. Muḥammad Karīm Tabrīzī's *Burhān-i Jāmi* ('Comprehensive proof', 1833) was a reworking of *Burhān-i Qāṭi* and *Farhang-i Jahāngīrī*.⁴³ The most celebrated dictionary of this period is Rīzā Qulī Khān Hidāyat's *Farhang-i Anjumanārā-yi Nāṣirī* ('The gathering-illuminating dictionary of [the reigning Qajar Shah] Nāṣir', 1871), published in Tehran. Although Hidāyat was tutor to the Qajar princes, and the leading literary scholar of his time, the dictionary was mostly a rehashing (albeit with significant corrections) of *Farhang-i Rashīdī*, *Burhān-i Qāṭi*, and *Farhang-i Jahāngīrī*. Hidāyat's main contribution was adding new verse citations and less usefully words drawn from the *Dasātīr*, a sixteenth-century document purporting to be the writings of ancient prophets but now universally acknowledged to be a forgery.⁴⁴

European Scholarship

The German-born, Calcutta-based Orientalist Henry Blochmann observed in 1868 that 'There exists no reliable Persian dictionary . . . This is a matter of surprise, as there are most excellent sources from which a good Persian dictionary could be compiled.'⁴⁵ In his time, European lexicography of Persian was still mostly derived from seventeenth-century precursors such as Meninski's Persian and Turkish dictionary and Edmund Castell's *Lexicon heptaglotton* (see Chapter 11).⁴⁶ Blochmann's article 'Contributions to Persian lexicography' was a deeply researched attempt to identify all of the sources that would be useful for a new Persian–English dictionary based on a proper critical engagement with the indigenous lexicographical tradition rather than repeating previous European lexicographers' mistakes. He proposes ten key works as a basis for the dictionary (the three written after 1700 are *Sirāj al-Lughat*, *Bahār-i 'Ajam*, and *Ghiyās al-Lughāt*).⁴⁷ Although Blochmann compiled no dictionary himself – the article was the result of six years of study and ends with a plea to fund such a compilation – his research was the first proper

⁴³ Dabīrsīāqī, '*Borhān-e jāmi*'. ⁴⁴ 'Affī, '*Anjomanārā, Farhang-e*'.

⁴⁵ Blochmann, 'Contributions to Persian lexicography', 1.

⁴⁶ A four-volume revision of Meninski's work with an English translation, by the Orientalist and pioneer of Indo-European comparative philology William Jones, with contributions by the Oxford Orientalist Joannes Uri, was advertised in 1771 in Jones' *Grammar* (sigs. Qq2r–v), but was abandoned thereafter.

⁴⁷ Blochmann, 'Contributions to Persian lexicography', 3.

analysis in English of the intertextuality of the Persian lexicographical tradition.⁴⁸ The article was also significant for reflecting on regionalism in Persian usage.

The culmination of Persian–Latin lexicography was the *Lexicon Persico–Latinum etymologicum* by Johann August Vullers. The work was published in Bonn between 1855 and 1864, and made use of sources from the indigenous tradition, namely *Burhān-i Qāṭi*, *Haft Qulzum*, *Bahār-i ‘Ajam*, the seventeenth-century Persian–Turkish *Farhang-i Shu‘ūrī* (printed in Istanbul in 1742), and from European scholarship, namely the lexicon of Richardson (described below) and those of Meninski and Castell.⁴⁹ Despite this range of sources, the dictionary was the result of problematic scholarship. Blochmann harshly criticises Vullers for reproducing errors in his source material and introducing many others because of, for example, not knowing Hindi/Urdu and therefore having misread Indic words in the texts. In particular, Blochmann writes, the manuscript of *Farhang-i Shu‘ūrī* that Vullers used ‘must have been the worst possible in existence’.⁵⁰

Persian dictionaries in most modern European languages were relatively slow to arrive compared to those in English. French and German lexica, for example, began appearing only in the late nineteenth and early twentieth centuries, with continental scholars up to that point continuing to use Latin as the target language. The first comprehensive bilingual Persian–French dictionary appears to be the 1864 *Iṣṭilāḥāt-i Farānsah–Fārsī* (‘French–Persian expressions’) by the Qajar courtier Muḥammad Ḥasan Ṣanī‘ al-Dawlah, and a number of others appeared in Europe subsequently.⁵¹ It appears that no significant Persian–German dictionary was published before the twentieth century. Here we are primarily interested in Persian to English lexicography because the practical needs of colonialism led it to be more complete than in other European languages. Persian lexicographical scholarship in English in the nineteenth century was basically a product of the British empire in India, whose administration required significant engagement with Persian, even at this late moment, when the Mughal emperors had become figureheads. The East India Company officially replaced Persian with English and vernacular languages in 1837, but this apparent watershed did not obviate the need for Persian education (and Persian reference books) into the twentieth century. The company was also involved in printing Persian scholarly books, for example the edition of *Burhān-i Qāṭi* published by Thomas Roebuck in

⁴⁸ Blochmann, ‘Contributions to Persian lexicography’, 7, 72. ⁴⁹ Paul, ‘Vullers’.

⁵⁰ Blochmann, ‘Contributions to Persian lexicography’, 19, 42ff.

⁵¹ Beni, ‘Historique des dictionnaires bilingues’.

1818, which became widely available in Europe despite being full of errors.⁵² Important Persian dictionaries of Arabic, such as the *Ṣurāḥ*, were also printed in Calcutta around the same time.⁵³

The first comprehensive Persian–English dictionary was published by John Richardson in 1777–80, and remained the basis for Persian lexicography in English for more than a century. The first volume was a Persian–English dictionary and the second was English–Persian. It was revised twice, first by Charles Wilkins in 1806–10, and later by Francis Johnson in 1852, who abandoned the English–Persian component and published the dictionary under his own name. Johnson observes in the preface to his own edition that Richardson’s dictionary was cumbersome to use because it was folio-sized and besides was ‘little else than an abridgement of the Oriental Thesaurus of Meninski’ with some material from Castell and Golius that omitted the Turkish and combined entries for homonyms.⁵⁴ He condemns Richardson’s English–Persian dictionary as ‘unidiomatic’ and adds that it was therefore rarely used. Johnson claims to have added 30,000 words to Richardson’s stock, primarily by using Arabic dictionaries, but Blochmann criticizes his work, although not as fiercely as he did that of Vullers, for the repetition of mistakes.⁵⁵

Blochmann’s proposal for an entirely new Persian–English dictionary never bore fruit, but at the end of the nineteenth century lexicographers had come close to producing definitive dictionaries. The Calcutta-born scholar John Thompson Platts was the first holder of a post dedicated to teaching Persian in a British university (at Oxford, from 1880). He is the author of the still-standard *Dictionary of Urdu, Classical Hindi, and English* (1884). Although he did not prepare a Persian dictionary, his Urdu–English dictionary is so concerned with word origins that it in effect functions as a well-researched Persian dictionary. In 1892, Francis Joseph Steingass revised Johnson ‘by reducing the Arabic element and increasing the Persian’ while drawing heavily on Vullers and adding in material from *Bahār-i ‘Ājam* that Vullers had neglected as well as incorporating more recent works.⁵⁶ Like Johnson’s dictionary, his work is both a Persian and an Arabic dictionary in so far as it contains the Arabic vocabulary encountered in Persian texts.

⁵² Blochmann, ‘Contributions to Persian lexicography’, 18.

⁵³ F. Johnson, *Dictionary, Persian, Arabic, and English*, ii.

⁵⁴ F. Johnson, *Dictionary, Persian, Arabic, and English*, i.

⁵⁵ F. Johnson, *Dictionary, Persian, Arabic, and English*, ii; Blochmann, ‘Contributions to Persian lexicography’, 41–2.

⁵⁶ Steingass, *Comprehensive Persian–English Dictionary*, v.

Twentieth-Century Developments

The Persian lexicographical tradition culminates in the *Lughatnāmah* begun by ‘Alī Akbar Dihkhudā. Dihkhudā was the Tehran-born son of a landowner from Qazwin and became a bureaucrat, editor, critic, and politician. He had the idea for an authoritative Persian lexicon in 1914 during a brief period of exile among the Bakhtiari tribes in west-central Iran.⁵⁷ He published the first volume in 1934 and continued working on the dictionary for the rest of his life. With the support of the National Assembly, he created a foundation in 1945 to manage the project. Dihkhudā eventually returned to politics, supported the reformist prime minister Mohammed Mossadegh, and was shaken by the 1953 *coup d’état* that overthrew Mossadegh. He devoted the remaining three years of his life to the dictionary, which was far from finished at his death. Muḥammad Mu‘īn was named director of the foundation in accordance with Dihkhudā’s wishes, and in 1957 Parliament transferred the foundation to the Faculty of Literature, University of Tehran. Mu‘īn’s colleagues, who included the lexicographers Muḥammad Dabīr Siyāqī and Muḥammad Ja‘far Shahīdī, eventually completed the work in 1980.⁵⁸ The foundation continues to update the dictionary, which was revised and reissued in 1994 and 1998, when it was also made available on CD-ROM and subsequently online.

Dihkhudā’s circle spearheaded lexicographical efforts in twentieth-century Iran. Mu‘īn published his own encyclopedic dictionary in six volumes between 1963 and 1973 (of which two volumes were edited by Shahīdī). It is the first Persian dictionary which fully embraces the best practices of modern European lexicography: Mu‘īn familiarized himself with the work of dictionary publishers such as Brockhaus in Germany and Larousse in France, a preferred model being the *Petit Larousse illustré* of 1961.⁵⁹ These lexicographers also edited pre-modern Persian dictionaries for publication in Tehran, such as Mu‘īn’s edition of *Burhān-i Qāṭi‘* and Dabīr Siyāqī’s of the *Ghiyās al-Lughāt* of Ghiyās al-Dīn and the *Farhang-i Ānandrāj* of Shād. Bilingual dictionaries in this period expanded from the diplomatic French–Persian dictionaries of the nineteenth century. For instance, Sulayman Hayyim’s Persian–English dictionary (Tehran, 1934–6) has some 50,000 entries and has been reprinted many times. Hayyim was a Jewish translator and playwright who also wrote French and Hebrew dictionaries of Persian (the latter is unpublished).⁶⁰

⁵⁷ Sīrjānī, ‘Dehkodā’. ⁵⁸ Sīrjānī, ‘Dehkodā’; *Dehkhoda Lexicon Institute*.

⁵⁹ Moussākhānī, ‘Dictionnaire encyclopédique persan de Mohammad Moīn’.

⁶⁰ Netzer, ‘Haim’.

The last important Persian dictionary to be compiled on the Indian sub-continent was *Aṣaf al-Lughāt*, written in Hyderabad between 1906 and 1921 by Aḥmad al-‘Azīz al-Nā’iṭī Walā (d. 1924). The work was still incomplete at the author’s death and has never been finished. Its enormosity is clear from that fact that it broke off at *jīm* (less than a quarter of the way through the alphabet) but was 10,200 pages long. Despite being incomplete, it was used by Dihkhudā because it drew upon many other dictionaries intelligently and comprehensively.⁶¹

The lexicography of Persian in Afghanistan and Tajikistan, the two present-day countries besides Iran that have Persian as an official language, developed only in the twentieth century. In Afghanistan, the colloquial language diverges from standard Iranian Persian and contends with the uneasy linguistic duopoly of Persian (known officially as ‘*dari*’ since 1958) and Pashto in public life. Iranian dictionaries are typically used, and the *Lughāt-i ‘Āmiyānah-yi Fārsī-i Afghānistān* (‘General dictionary of the Persian of Afghanistan’, Kabul, 1961) by ‘Abdullāh Afghānīnavīs appears to be the only general-purpose Dari dictionary available. Tajik lexicography was necessitated by the change to Roman and then Cyrillic script, in 1928 and 1940 respectively, as well as the ‘vernacularization’ of written Tajik compared to the transnational classical standard. Šadr al-Dīn ‘Aynī, regarded as the father of modern literary Tajik, began compiling the *Lughat-i Nīm-Taḥṣīlī barā-yi Zabān-i Tājīkī* (‘Semi-explicitary dictionary for the Tajik language’) in 1938, but it was not published during his lifetime and did not appear in print until 1976.⁶² In the meantime, the Tajik Academy of Sciences had published *Farhang-i Zabān-i Tājīkī* (‘Dictionary of the Tajik language’) in 1969. Numerous classical Persian dictionaries were published in Cyrillic editions in the twentieth century and there was a strong Tajik–Russian lexicographical tradition.⁶³

⁶¹ Mahmoodi-Bakhtiari, ‘*Aṣaf al-lughāt*’; Dabīr Siyāqī, *Farhang-hā-yi Fārsī*, 194–5.

⁶² Perry, ‘Modern Iranian lexicography’.

⁶³ Borjān, ‘Tajikistan v.: Dictionaries’; Burjiyān, ‘*Farhangnīvisī*’.

South Asia from c. 1750

WALTER HAKALA AND LISA MITCHELL

South Asia, a region that has been defined variously to include all or parts of present-day Afghanistan, Pakistan, India, Nepal, Bhutan, Bangladesh, Sri Lanka, the Maldives, and occasionally Myanmar and the Tibet Autonomous Region, is home to about a quarter of the world's population and is the scene of much of its linguistic diversity. Each of these nations has adopted a different set of official languages. Some have enshrined multilingualism at the national or provincial levels, while others have chosen a single language to serve as an official national language. Sri Lanka, for example, amended its constitution in 1987 to name both Sinhalese and Tamil as official and national languages, with English retained as a 'link language', and the 8th Schedule to the Indian Constitution currently designates twenty-two official languages, with an official 'three-language formula' implemented in 1968 by India's Ministry of Education.¹ Bhutan, at the other extreme, has fixed Dzongkha as the sole national and official language, with Chökê serving as a liturgical language, English as a 'necessary foreign language', and other languages as local mediums.² This linguistic pluralism is typical across the region, with the expression *har das kos bolī badaltī hai* (every twenty miles the dialect/speech changes) or some variant common throughout northern India.³ Rather than attempt a comprehensive survey of this vast region, this chapter will examine a representative sample of works in order to illustrate how vocabularies, wordlists, and dictionaries are both reflections of and sites for the articulation of broader material, political, and ideological changes in South Asian society.

The lexicography of modern South Asia derives from three major philosophical traditions: the Sanskritic, Persianate, and European. Our use of these

¹ DeVotta, 'Ethnolinguistic nationalism', 133–4; Karunaratne, *Teaching of English*, 10; Schiffman, *Linguistic Culture and Language Policy*, 167–8.

² Wangdi, 'Language policy and planning in Bhutan'.

³ Washbrook, 'To each a language of his own', 181. The phrase in Gujarati is *bar gaue boli badlay*; Dwyer, *Gujarati*, 5.

three terms as a convenient shorthand does not, however, correspond with any assertion of atavistic isolation, uniformity, or historical coherence. South Asia has been the site of deep debates on the nature of the impacts of Sanskritic, Persianate, and European models on knowledge production and everyday practices within the subcontinent.⁴ Tempting though it may be to view the history of South Asian lexicography as developing along a single track – be it from manuscript to print, poetry to prose, idiosyncrasy to standardization – our examples illustrate the significant variations that characterize the ways in which South Asian lexicographers have contributed to the construction of the ‘modern’ field of lexicography in South Asia.⁵ This allows us to avoid viewing the lexicographic experiments that took place during this period as a series of apparent failures, false starts, and lagging indicators, diverging in some way from a universalized European model. Moreover, the number and variety of languages represented by South Asian lexicography are daunting: readers seeking more comprehensive surveys of South Asian lexicography are encouraged to turn to the excellent chapters on the subject in the third volume of Hausmann et al.’s landmark *Wörterbücher: Ein Internationales Handbuch zur Lexikographie*.⁶ Our focus will instead be on illustrating the range of ways that South Asian lexicographers have shaped and responded to new forms of identity, political aspiration, and social organization, articulating these changes through their practices of lexicography.

The very notion of treating the languages of South Asia as discrete and potentially equivalent subjects of study is the result of a series of political and ideological transformations that saw language increasingly become a marker of distinct social identities.⁷ At the beginning of the period covered in this chapter, lexicographers operated within a multilingual system that assumed

⁴ On the process of Sanskritization, see Srinivas, *Social Change in Modern India*, 1–88; for a critique of Sanskritization, see Singh, *Modernization of Indian Tradition*. On the distinction between Islamic and Islamicate influences on the subcontinent, see Wagoner, ‘Sultan among Hindu kings’ and, by analogy, for that between the terms *Persian* and *Persianate*, see Hodgson, *Venture of Islam*, II.293. For contrasting summaries of the debate over the British colonial role in the construction of knowledge in (and about) South Asia, see Dirks, *Castes of Mind*, 303–15, and Pinch, ‘Same difference in India and Europe’.

⁵ On the tendency of area studies scholarship to promote a ‘denial of coevalness’, see Fabian, ‘The other revisited’, 143.

⁶ Balasubramanian, ‘Bilingual lexicography of the Indian subcontinent’; Haywood, ‘Arabic lexicography’; Haywood, ‘Bilingual lexicography with Arabic’; Katre, ‘Lexicography of Old Indo-Aryan’; Krishnamurti, ‘Dravidian lexicography’; Perry, ‘Modern Iranian lexicography’; Singh, ‘Lexicography of New Indo-Aryan’; Strandberg, ‘Lexicography of Middle Indo-Aryan’; Zide, ‘Lexicography of other languages of the Indian subcontinent’.

⁷ See especially chs. 1 and 2 of Mitchell, *Language, Emotion, and Politics*.

linguistic forms would serve limited and complementary roles. By the turn of the twentieth century, however, dictionaries had become vital symbols of the political aspirations of competing communities, operating both within and across national boundaries. An ideology of linguistic nationalism has at its basis the twin notions that each nation must possess its own distinctive national language and that ready equivalents in that language must be identified for any term existing in another. Similarly, the contents of these works expanded from the predominantly literary concerns of earlier lexicographers (whose works were devoted to cataloguing lexical options for use in a wide range of poetic metrical contexts and were themselves often written in verse form) towards a view of language that reflected the claims of elites to represent diverse subject populations and establish parity with the languages of other ethno-national populations. Aural and oral mnemonic technologies that were embedded in often idiosyncratic and experimental generic forms gave way as lexicographers took advantage of printing, cheap paper, and collective financing, culminating in a new reliance on visual modes of reference. The production of the modern monolingual comprehensive dictionary became the *sine qua non* of political assertions of cultural differentiation and group cohesion while their bilingual counterparts announced a group's claims to equal standing within a community of peers.

Sanskrit and Persianate Lexicographical Influences

South Asian traditions of vocabularies in verse may be traced back at least two millennia through early Sanskrit examples such as the *Nighaṇṭu* and Amarasiṃha's *Amarakośa* (see Chapters 4 and 7).⁸ The primary concern of pre-colonial lexica was the cataloguing of terms employed in the production and appreciation of *kāvya* (literary works in verse form), with particular attention to documenting two major modes of equivalence: homonymy (*anekārtha*, *nānārtha*) and synonymy (*ekārtha*, *saṁānārtha*, *nāmamālā*). Such lexica were of great assistance to the efforts of authors who sought to demonstrate their learning through elaborate forms of wordplay such as bi-textual *śleṣa* oronymy (double meaning verse, also *dwisandhana kāvya* in Sanskrit, or *dwyarthi kāvya* in Telugu), *śatartha kāvya* (the composition of a single verse with a hundred meanings), *citra kāvya* ('picture' verses, in which lines of intersecting letters or syllables arranged in an often complex pattern

⁸ Amarasiṃha, *Cōśha*, is a translation of the *Amarakośa*. For general introductions to Indic lexicography, see Vogel, *Indian Lexicography*; Patkar, *Sanskrit Lexicography*.

can be read in multiple directions), compositions in which each verse reads the same both forwards and backwards, or compositions in which an entire class of letters is proscribed (e.g., *niroshthya kāvya* in which all labial vowels and consonants are proscribed, or *niranunāsika*, in which no nasal sound appears).⁹

Typically written in verse and arranged by theme, *nighaṇṭu* and *kośa* lexical texts employ a variety of mnemonic devices that enable oral transmission and memorization.¹⁰ Students continued to commit works such as the *Amarakośa* and later bilingual adaptations, such as the Telugu *Śiva Āndhra*, to memory well into the nineteenth century.¹¹ What is striking about the various regional 'bilingual' adaptations of Sanskrit lexica from today's perspective is the fact that words from Sanskrit and from regional languages were grouped together, offering poets sets of acceptable words that could be used in compositions rather than providing translations from one language to another. Following the pattern set by the *Amarakośa*'s three chapters (*kāṇḍa*) and twenty-five sections (*varga*), these works often organized their contents in hierarchical topical divisions, ranging from terms related to the divine and deities, celestial bodies, terrestrial phenomena, occupations, and animals, to collections of polysemous and other miscellaneous terms.¹²

Persianate texts began to influence lexicographical practices in South Asia from at least the thirteenth century. The *Niṣāb al-ṣibyān* ('Wealth of children'), a 200-verse Arabic–Persian vocabulary in verse, was composed in 1220 in what is now south-western Afghanistan. This work inaugurated a popular genre of versified multilingual vocabularies for children, frequently employing one or more Indic languages to gloss or provide synonyms for Persian, Arabic, or, beginning in the nineteenth century, English terms.¹³ Two sixteenth-century works, the *Qaṣīdah dar luḡhāt-i hindī* ('Ode on Hindi terms'), a medical vocabulary composed by the first Mughal emperor's court physician Yūsuf bin Muḥammad Yūsuf Kḥurāsānī, and *Ajay cand nāmah* (Book of Ajay Cand), a vocabulary by a Hindu *kāyasth* (scribe) that organizes its contents thematically in sections named for the rooms of a palace, modify the format of the *Niṣāb al-ṣibyān* by providing Hindawī (northern Indian)

⁹ Bronner, *Extreme Poetry*, 128–32; Warder, *Indian Kāvya Literature*, III.223–8; Wagoner, 'Telugu literature of the Qutb Shahi period', 100.

¹⁰ See, e.g., Sankara Reddy's comments about Telugu lexica in *Reference Sources in Telugu*, 23.

¹¹ Brown, *Essay on the Language and Literature of the Telugus*, 5.

¹² Birwé, 'Amarakośa'; Nair and Kulkarni, 'Knowledge structure'. Compare, e.g., Patkar, *Sanskrit Lexicography*, 132, 135, 148; Sankara Reddy, *Reference Sources in Telugu*, 24–5.

¹³ Farāhī, *Niṣāb al-ṣibyān* (1923), is an edition; see also Sherānī, 'Dībācah-i avval', 5–6.

equivalents for Arabic and Persian terms. These authors may have been responding to the growth of a bureaucracy across much of the Indian subcontinent that employed scribes literate in Persian.¹⁴ The expansion of a Persianate high culture accompanied the rise of Muslim political hegemony across much of the Indian subcontinent. The first Persian–Sanskrit lexica appeared as early as the fourteenth century, and a number of such works were produced in the sixteenth and seventeenth centuries under direct and indirect Mughal–Timurid patronage.¹⁵ Some of these texts were analogous to the Sanskrit *nāmamālā* genre, which had been adapted as early as the 1365 *Śabdavilāsa* (‘Sport of words’; also called *Pārasīnāmamālā*, ‘Garland of Persian words’) by Salakṣa to gloss Persian (and in later vocabularies Turkish) terms in Sanskrit.¹⁶ Through Persian, South Asian authors, beginning around the turn of the fifteenth century with Badr-al-Dīn Ibrāhīm’s *Farhang-i zafāngūyā wa jahānpūyā* (‘Eloquent and world-seeking dictionary’, written before 1433), adapted the Persianate *farhang* and Perso-Arabic *luḡhat/luḡhāt* genres of defining dictionaries for use in South Asia, deploying various forms of alphabetical ordering and, occasionally, Indic equivalents as glosses for terms drawn from Persian, Arabic, Turkic, and other languages.¹⁷

Multilingual to Monolingual

Like most other lexicographic traditions, South Asian lexicography from 1750 has been dominated by bilingual lexicography. The ideological and practical aspects of this multilingualism, however, have shifted over time. With the notable exceptions of Sanskrit, Persian (for which see Chapters 11 and 17), and Arabic (for which see Chapters 8 and 19), the late appearance of monolingual dictionaries in the vernacular languages of South Asia corresponds with patterns typical across most of Europe.¹⁸ But even mono- and bilingual works of the nineteenth century give emphasis to the tendency of South Asian spoken languages to draw their vocabulary from a variety of sources.

C. P. Brown’s 1854 *Dictionary of the Mixed Dialects and Foreign Words Used in Telugu* demonstrates the attitudes of those European lexicographers who

¹⁴ Alam, ‘Pursuit of Persian’; Hakala, ‘Ajay čand nāma’; Hakala, ‘Qaṣīda dar luḡāt-i hindī’.

¹⁵ See Patkar, *Sanskrit Lexicography*, 133–4, 145, 148–50, 160–2.

¹⁶ Truschke, ‘Defining the other’, 638, 641.

¹⁷ Baevskii, *Early Persian Lexicography*, 94–104; see also Chapter 11 in this volume.

¹⁸ For Arabic, see also Versteegh, *Arabic Linguistic Tradition*, 23–34. Cf. Malkiel, ‘Typological classification of dictionaries’, 11; McArthur, *Worlds of Reference*, 24, 76, 82, 89. See also Zgusta’s helpful discussion of what he calls the ‘quasi-bilingual’ dictionary in *Lexicography Then and Now*, 293.

sought to describe the range of languages one encounters on a daily basis in South Asia in support of the project of administration. Brown begins his preface to the dictionary by characterizing government business in southern India as taking place in ‘Tamil, Telugu, Kannadi [Kannada], Malayalam, or Marata [Marathi]’ among Hindus while Muslims ‘still speak Hindustani [Hindi-Urdu] as well as the local native tongue’. The Telugu he describes as containing ‘a variety of Persian and Arabic expressions’ while ‘in later years some English words have crept into use’.¹⁹ This practical approach of Brown to ‘this inelegant but useful dialect’ may be differentiated from the classicizing efforts of other European scholars and, subsequently, nationalists who sought to isolate ‘pure’ and ‘chaste’ forms of languages and purge them of their ‘foreign’ accretions.²⁰ But even Brown, who was fascinated by the amalgam of words that comprise the vocabularies of South Asia languages, would encourage his readers to make every effort to study the ‘pure Telugu words’ – those ‘not derived from Sanscrit’ – included as synonyms of the ‘foreign’ terms glossed in his bilingual dictionary: ‘This will aid those who wish to get rid of these foreign words. Perhaps one word in twenty may merit preservation. But in talking and writing Telugu we ought to avoid this mixed dialect.’²¹

Print and Paper Technologies

Though the first appearance of print in South Asia may be traced to texts of the mid sixteenth century, their circulation was almost exclusively confined within European communities in coastal areas of the subcontinent. Large-scale networks of manuscript production and distribution dominated book production in South Asian languages well into the nineteenth century.²² Despite a few scattered efforts to develop typefaces for the Arabic, Eastern Nāgarī (Bengal-Assamese), and Tamil scripts, and the official promulgation of texts produced using these fonts by institutions such as the Fort William College in the opening years of the nineteenth century, their costs were prohibitive to all but a very small elite. One such early printed South Asian dictionary was the *Shams al-lughāt* (‘Sun of words’) – or, as the English-language title-page called it, *Shums-ool-Loghat, or, a dictionary of the Persian*

¹⁹ Brown, *Dictionary of the Mixed Dialects*, iii.

²⁰ Brown, *Dictionary of the Mixed Dialects*, iii.

²¹ Brown, *Dictionary of the Mixed Dialects*, xxi, iv. See also Trautmann, ‘Dr Johnson and the pandits’.

²² Shaw, ‘South Asia’.

and Arabic Languages, the Interpretation being in Persian – compiled under the direction of Joseph Baretto and printed in Calcutta in 1806. While sumptuous in its production, using fine paper imported from Britain, it was deemed ‘too full of typographic errors, to render [its] use desirable’.²³

Beyond the high cost of production and a propensity for errors, the slow spread of movable-type printing may also be explained aesthetically in so far as it provided a poor approximation of the appearance of handwritten manuscripts. Graham Shaw has argued that the introduction to South Asia in the 1820s of lithography, with its visual resemblance to manuscripts and flexibility in accommodating a variety of scripts (especially those that incorporate cursive elements such as Persianate *nasta‘līq* and Indic *kaithī*) represents ‘the true dawn of printing in the subcontinent’.²⁴ Others have argued, adapting social theorists such as Durkheim and Habermas to the South Asian context, that the adoption of print was emblematic of a broader ideological shift during the nineteenth century away from modes of knowledge transmission that privileged oral (or anthropocentric) interactions towards a more print-oriented ‘public sphere’.²⁵

With the spread of print technologies, many features of pre-colonial South Asia lexicographic genres that facilitated the memorization of texts came to be derided as redundant. The frequent and masterful application of word play in many South Asian literary traditions was enabled in part by the proliferation of thematically arranged *kośa* vocabularies in verse, with their strings of synonyms on important topics and capacity for the simultaneous communication of multiple meanings. Print technology transformed the virtues of this synonymy and homonymy into faults. For G. V. Rāmamūrti, the author of the 1913 *Memorandum on Modern Telugu*, who sought to bring literary Telugu closer to the spoken language,

The Pandit swallows all this stuff and prescribes it to his pupil for his literary salvation. Variety is a merit in his own language; it is a demerit in the modern dialect. The variants in a living language will develop into synonyms expressing delicate differences in meaning, whereas the Pandit’s variants remain a useless load on the memory.²⁶

Even pre-modern texts could be transformed through print technologies. H. T. Colebrooke’s 1808 edition of the famous *Amarakośa* Sanskrit lexicon in

²³ Blochmann, ‘Contributions to Persian lexicography’, 2.

²⁴ Shaw, ‘South Asia and the history of the book’, 18–19.

²⁵ See, e.g., Robinson, ‘Technology and religious change’; Green, ‘Uses of books in a late Mughal *takiyya*’.

²⁶ Rāmamūrti, *Memorandum*, 40.

verse devotes 393 pages to the original text in *devanāgarī* script with interlinear English translations and an extensive commentary. It concludes, however, with a 219-page alphabetical index – comprising more than one-third of the total length of the entire codex – in an attempt to make this ancient text useful for visual rather than aural consultation.

The long South Asian tradition of composing and memorizing multilingual vocabularies would likewise be adapted to fit new material and ideological conditions. Works that would once have been prepared in verse form now began to be organized as parallel wordlists. While lithography reduced the cost of preparing texts in multiple scripts and encouraged new kinds of equivalences, the growing popularity of multilingual wordlists encouraged lexicographers to treat South Asian languages as distinct entities, physically separated on the page by white space and functionally equivalent. Some of these may have been intended for students seeking to learn English, though many wordlists, such as Ī. V. Sitārāmasvāmī's trilingual Telugu, Hindi, and Persian *Trībhāṣa-maṇjari* (1890), did not include English at all. It was not uncommon for such vocabularies to include three, six, or even more language varieties, each lined up in parallel, and each presented as a separate language.²⁷ Through this process of equating terms – that is, by listing terms drawn from increasingly differentiated 'languages' in separate, but functionally equivalent columns – lexicographers would transform the nature of multilingualism in South Asia. Rather than serving as complementary registers whose appropriateness in a particular context was determined by their value in completing a specific communicative task, languages were beginning to be claimed as 'mother tongues' – markers of collective, political identities.

A limited supply of paper throughout the eighteenth century further hampered early colonial experiments in printing. Much of the stock was of necessity imported (at great cost) from Europe. In eastern India, two paper mills in Calcutta and Patna were evidently in operation by the 1790s. When the Baptist Press was established in 1800 on the Danish island colony of Serampore, it was at first compelled to employ the low-quality paper produced at Patna. After experiments with handmade paper production on a small scale and suffering a disastrous fire in which much of its imported paper stock was lost, the press in 1820 installed a steam engine imported from England in what would be the first mechanical paper mill in India.

²⁷ An early example was the eleven-language list published in [Schultze and Fritz], *Orientalisch- und Occidentalisches Sprachmeister*, between 212 and 213, and then as Abel, *Symphona symphona*, for which see Chapter 30 in this volume; for additional examples, see Mitchell, 'Parallel languages', 458.

The second half of the nineteenth century witnessed a proliferation of steam paper mills in Bombay, Lucknow, Ahmedabad, Surat, Gwalior, and the colonial capital Calcutta.²⁸ By the first decade of the twentieth century, the traditional production of paper by hand had become ‘a dying industry’ in Bengal despite having been in a ‘flourishing condition’ a mere generation earlier.²⁹ Paper famines during the two world wars may also explain, for instance, why Sayyid Aḥmad Dihlavī waited until 1918 to bring out two Urdu dictionaries (the second edition of his four-volume *Farhang-i āṣafīyah* and *Lughāt al-nisā*’, a ‘Women’s vocabulary’).³⁰ Improvements in the use of grass and bamboo as materials for paper beginning in 1920 encouraged the production of larger and less expensive works such as two massive monolingual Urdu dictionaries, N. Ḥasan’s *Nūr al-lughāt* (1922–31) and ‘Abdul Majid’s *Jāmi al-lughāt* (1935).³¹

South Asian Lexicography and the Birth of Historical Linguistic Methods

The European encounter with South Asian lexicographical practices was central to the development of comparative and historical linguistics and the birth of the concept of the Indo-European language family. Early Jesuit missionaries had begun producing bilingual Tamil–Portuguese vocabularies and glossaries as early as the mid sixteenth century (see Chapter 30). Gaston-Laurent Coeurdoux, a French Jesuit missionary who spent the bulk of his life in southern India, responded in 1767 to queries sent by the Abbé Barthélemy of the Académie des Inscriptions et Belles-Lettres asking for information about the Sanskrit language, with references to his consultation of a Sanskrit dictionary written in the Telugu script, ‘*Amara sim houam*’ (in other words Amarasiṃha’s *Amarakośa*), as well as his production of a draft of ‘un petit dictionnaire *Têlongou–François* et *Samscroutam*’.³² He then offered a detailed description of the similarities between Sanskrit, Latin, and Greek, illustrating his discovery with parallel wordlists and outlining the grammatical and phonological relationships between the three languages, and argued that the obvious similarities must be due to their ‘common origin’, rather than to commerce, intercourse, proximity, religion, or invasion.³³

²⁸ Ramaseshan, ‘History of paper’, 116–17. ²⁹ Mookerji, *Monograph on Paper*, 3.

³⁰ Ramaseshan, ‘History of paper’, 118. ³¹ For both, see Bailey, *Jami ‘ul-lugāt*’.

³² Coeurdoux, ‘Réponse’, 648.

³³ Coeurdoux, ‘Réponse’, 651–67, esp. 660, ‘une commune origine’. See also Godfrey, ‘Sir William Jones and Père Coeurdoux’, 58, and cf. Arlotto, ‘Jones and Coeurdoux’.

Although Coeurdoux's observations pre-date by almost two decades the similar (and much more well-known) observations of the British philologist Sir William Jones, it was not until 1808 that the French Jesuit's observations were finally published. Jones' theory of the common origins and kinship of Latin, Greek, and Sanskrit was presented in his 1786 third anniversary address to the Asiatick Society of Bengal, which he had founded in 1784. Like Coeurdoux, Jones was influenced by his study of Persian and Sanskrit lexicography, particularly his engagements with the Sanskrit *vyākaraṇa* (language analysis) tradition's efforts to account for the entire lexicon through the generation of a comprehensive list of verbal roots (*dhatu pātha*) and a set of concise rules (*sūtras*) for transforming these roots into every other linguistic element present within the language (see Chapter 30 for Jones and the missionary tradition of Sanskrit studies).

European Influences

If Jones found lexicographical authority in the classicism of Sanskrit, which he found 'more perfect than the *Greek*, more copious than the *Latin*, and more exquisitely refined than either', John Gilchrist represents a competing approach that drew from the authority of everyday speech.³⁴ Gilchrist complained about the challenges in procuring vernacular lexicographic works as he set about in 1785 to compile what would become his *Dictionary, English and Hindoostanee* (published in parts between 1786 and 1798). In an effort 'to fathom, as far as possible, the real depth of that tongue', he decided to leave the army detachment in which he served as an assistant surgeon to reside in the then-independent Indian state of Awadh. 'At so considerable a distance from all my own countrymen', he hoped to 'faithfully dedicate without the possibility of interruption, every moment I could safely snatch from the devouring jaws of Indian slumbers, to my projected work'.³⁵ In his dictionary's preface, he writes that he asked native speakers for dictionaries and grammars: 'Those from whom I required them stared with astonishment, and answered interrogatively, if it was ever yet known in any country, that men had to consult vocabularies, and rudiments for their own vernacular speech.' Exasperated, he insists to his interlocutors that 'in the extended succession of five or six centuries, some thing or other must have been produced on the modern Hindoostanee philology, by the Moosulmans

³⁴ Steadman-Jones, *Colonialism and Grammatical Representation*, 49–60, 77–8.

³⁵ Gilchrist, *Dictionary*, vii.

who introduced it'.³⁶ At last, he depicts them as first producing for him the *Khāliq bārī*, a well-known trilingual Persian–Arabic–Hindi vocabulary in verse, which Gilchrist dismisses as 'a *Tom Thumb* performance' and 'the shrivelled mummy of an old meagre school glossary', and, secondly, the *Tuhfat al-hind* ('Gift from India'), a brief Braj–Persian glossary appended to a late seventeenth-century treatise.³⁷

Disappointed, he describes how he was compelled 'to extract *viva voce*' from the 'learned *Hindoostanees*' he employed (at, he insists, 'considerable expence [*sic*]') 'every known word in their voluminous tongue'. With pathos, he complains,

I a solitary individual, whom no one had preceded, must now complete such a system in a foreign tongue, without the smallest help from those even who had used it so long . . . To select words from the *Hindoostanee* writers, as they occurred in their compositions, seemed so endless a task, that I at once took the alphabet regularly so. My auditory were severally instructed to furnish me with every signification they could possibly attach to such sounds or words as *a*, *ab*, *abab*, *ababa*, *abach*, *abad*, *abada*, *abaf* &c. &c. so on, ringing the changes in this manner progressively, with every letter. One or two syllables commonly led the way to a numerous tribe of words, till I at last in this manner compiled, in the space of a few months by incessant application, the whole of the *Hindoostanee* Dictionary noticed hereafter, with more ease, and precision, and probably much sooner than, I could have accomplished it by any other mode, in a living speech of which neither a vocabulary nor grammar existed.³⁸

Notwithstanding claims to complete originality, Gilchrist is compelled to acknowledge a few pages later that his English wordlist was derived from a very celebrated source: 'from Johnson every word was explained in succession, to a number of learned *Hindoostanees*, who furnished the synonymous vocables in their own speech'.³⁹

The adaptation of European works for bilingual South Asian lexicography is a pattern that has been repeated in subsequent projects. Bilingual editions of British pedagogical texts were prepared for South Asian users. *Mavor's Spelling-Book, with its Goojrathee Translation* (1837) was one such work, providing interlinear transliterations and translations for Indian children seeking to learn English through their mother tongue. Karsandās Mūlji's *A Pocket*

³⁶ Gilchrist, *Dictionary*, vii.

³⁷ Gilchrist, *Dictionary*, vii–viii. On these works, see Hakala, 'Authorial problem'; Keshavmurthy, 'Mīrẓā Hān ibn Fahr al-Dīn, *Tuhfat al-hind*'.

³⁸ Gilchrist, *Dictionary*, vii–viii. ³⁹ Gilchrist, *Dictionary*, xiv.

Dictionary, Gujarati and English (1862) draws upon Webster, Ogilvie's *Imperial Dictionary*, and Roget's *Thesaurus*.⁴⁰ In addition to a primary corpus of terms drawn from contemporary newspapers, Munshī Ziyā' al-Dīn Aḥmad Kḥān, in his slim 1915 *Akḥbārī lughāt, ma'rūf bēh Kalīd-i akḥbār binī* ('Newspaper dictionary, known as "The key to newspaper viewing"'), lists several English-language works from which he took assistance for his Urdu dictionary. These include the *Encyclopædia Britannica*, *Whitaker's Almanack*, Walter Bagehot's *English Constitution*, *Pears Cyclopedia*, and *The Essays of Elia*.⁴¹ The *Pañjāb Yūnīwarasiṭī Aṅgrezī–Pañjābī kosha/English–Punjabi Dictionary* first published by Punjab University in 1968 adopts the *Concise Oxford Dictionary* as its model.⁴² P. G. Deshpande and B. Deshpande's *Universal English–Gujarati Dictionary* (1988), consisting of 34,000 English terms and phrases with Gujarati equivalents, is 'based on original Oxford source material'.⁴³

Orientalists, Anglicists, Missionaries, and Nationalists

By the beginning of the twentieth century, elites had begun to employ lexicography to demarcate distinct linguistic communities and align these communities with broader political projects. Dictionaries compiled in accordance with 'modern' Western models were perceived as an especially effective means for corporate groups to gain patronage from, and recognition by, the state, as were claims to antiquity as etymological research began to be incorporated into these political projects. Beginning in the nineteenth century, Indian authors began preparing vocabularies in verse (B. V. Mishra's *English–Hindi Vocabulary in Verse* of 1902 is an example) that would cater to the growing demand for educational materials in English, which in the 1830s replaced Persian as the official language of higher administration throughout the regions of India directly ruled by the British East India Company.⁴⁴ European scholars continued to produce monumental dictionaries of South Asian languages – most notably, Otto von Böhtlingk and Rudolf von Roth, whose *Sanskrit-Wörterbuch* (1855–75) comprises 7 volumes and 5,000 large pages.⁴⁵ However, the so-called Anglicists within the colonial state famously argued for the utility of educating 'a class who may be interpreters between us and the millions whom we govern; a class of persons Indian in blood and

⁴⁰ Múlji, *Pocket Dictionary*, vi. ⁴¹ Kḥān, *Akḥbārī lughāt*, 1.

⁴² Singh, 'Lexicography of New Indo-Aryan', 2515. ⁴³ Bender, 'Review'.

⁴⁴ King, *One Language, Two Scripts*, 53.

⁴⁵ For it, see Zgusta, 'Copying in lexicography', 148–9.

colour, but English in tastes, in opinions, in morals, and in intellect'.⁴⁶ The philological approach of classical Orientalists, especially those working in continental Europe, rested on the Romantic belief 'that the primary goal of research should be the discovery of the earliest testimonies of the human mind'. For Indologists of this persuasion, 'Modern texts were seen as derivatives of lesser interest except as they started new or independent traditions'.⁴⁷ Missionaries, by contrast, had goals both more immediate and eschatological: their focus was on communication with the communities in which they worked, and their vocabularies, glossaries, and phrasebooks tend to give emphasis to spoken language over classical literature (see Chapter 30).

Professionalization and the Rise of Corporate Lexicography

This is not to suggest that the flow of knowledge was always from Europe to Asia. Joseph Taylor and William Hunter adapted Mirzā Jān Ṭāpish Dihlavī's Urdu dictionary of idioms with Persian glosses, *Shams al-bayān fī muṣṭilāḥāt al-hindūstān* ('The sun of speech, on the idioms of Hindustan', completed c. 1792) into their *Dictionary, Hindoostanee and English* (1808).⁴⁸ Unlike Gilchrist and the other authors mentioned above, however, their wholesale incorporation of the earlier work was unacknowledged: Ṭāpish had been convicted in 1800 of conspiring against the British and was imprisoned in Calcutta until 1806 or 1807. The title page credits Captain Joseph Taylor with preparing the work 'for his own private use'; it was subsequently 'Revised and prepared for the press, with the assistance of learned natives in the College of Fort William, by W. Hunter'.⁴⁹ The list of 'about 8000 Marat,ha [*sic*] words' of the bidirectional Marathi–English *Dictionary of the Marat,ha Language: In Two Parts*, published in Bombay in 1824, was partly generated by enlisting 'an intelligent Brahman' to prepare Marathi equivalents of the Sanskrit words in the *Amarakośa*, after which the compiler, Vans Kennedy, translated these Marathi terms into English.⁵⁰ C. P. Brown engaged pandits in an attempt to collate and rearrange the terms that appear in thematically organized pre-colonial Telugu verse lexica such as

⁴⁶ Macaulay, *Prose and Poetry*, 729. On the debates between the so-called Anglicists and Orientalists regarding the state's patronage of English-language education in South Asia, see also Zastoupil and Moir, *Great Indian Education Debate*; Hall, *Macaulay and Son*, 225–34.

⁴⁷ Gaeffke, 'Rock in the tides of time', 70.

⁴⁸ Ṭāpish Dihlavī, *Shams al-bayān* (1979), is an edition of the source text.

⁴⁹ Hakala, *Negotiating Languages*, 76–114.

⁵⁰ Kennedy, *Dictionary of the Marat,ha Language*, iv (word count), ii (*Amarakośa*).

the *Āndhra nāma sangrahamu*, *Āndhra nāma śēṣamu*, *Vēṅkaṭṣāndhram*, and *Sāmba nighaṇṭuvu* in alphabetical order. The fruits of their labour, however, were never published, in part because the primary functions of the pre-colonial *nighaṇṭu* in comprehending, composing, and recalling *kāvya* literature were largely incompatible with the primary interests of colonial lexicographers in producing dictionaries to assist in language acquisition and translation.⁵¹ The few case studies described below, however, will demonstrate the increasing role played by the British in organizing, financing, and lending credibility to new dictionary projects of the nineteenth and twentieth centuries.

New Investments in Antiquity: The Case of the Dravidian Language Family

The ‘discovery’ of the Indo-European language family and the idea that languages bore traces of kinship relationships to one another led to new interest in the histories and antiquity of languages as evidence of the histories of distinct ‘nations’. Methodologically, these discoveries relied upon the new comparative philology made possible, in part, by the use of wordlists. Francis Whyte Ellis published the first proof of the Dravidian language family in his ‘Note to the Introduction’ to A. D. Campbell’s 1816 *Grammar of the Teloogoo Language*, refuting the notion that Telugu, Tamil, and the other languages of southern India were descended from Sanskrit.⁵² Like Coeurdoux and Jones, Ellis relied on insights drawn from *vyākaraṇa* analytic categories that divided lexical items into distinct categories that distinguished Sanskrit (*tatsama*) and Sanskrit-derived (*tadbhava*) vocabulary from lexical items not derived from Sanskrit (*deśya*). Ellis showed that the list of roots for Sanskrit words was distinct from the list of roots for Tamil, Telugu, and Kannada words, and that the roots for the latter group were cognate to one another.⁵³ Although Telugu was central to this discovery, the implications for the reorganization of linguistic and ethnological knowledge were profound, placing Tamil, as the Dravidian language with the longest written record, in a new position of pre-eminence and inspiring new interests in the purification of languages.

The expansion of intellectual horizons in South Asia during nineteenth-century encounters with European legal discourse, administrative practices, educational institutions, and literary genres motivated desires for new lexical

⁵¹ Mitchell, *Language, Emotion, and Politics*, 169–77.

⁵² Trautmann, *Languages and Nations*, 151–85.

⁵³ Trautmann, ‘Hullabaloo about Telugu’, 60.

items in South Asian languages. As education and literary production had historically been dominated by Brahmins, lexical innovation within South Indian languages had drawn from Sanskrit since well before the twelfth century. But, by the nineteenth century, there was growing resentment among many non-Brahmin South Indians, prompting efforts such as the ‘pure Tamil’ movement of the early twentieth century.⁵⁴ In 1937, T. Nilambikai Ammaiyar, the first female lexicographer of Tamil, published a Sanskrit–Tamil dictionary intended to assist Tamilians in avoiding Sanskrit terms. Such efforts were also addressed towards English lexical items. The University of Madras’ *English–Tamil Dictionary*, published in three volumes from 1963 to 1965 and modelled on the 1958 printing of *The Concise Oxford Dictionary*, had a similar purpose in offering pure Tamil equivalents, some newly coined for the purpose.⁵⁵ In the 1950s and 1960s the Committee for Implementing Tamil as an Official Language sought to compile a ‘Tamil official language dictionary’.⁵⁶ Although a comprehensive dictionary was never completed, numerous official glossaries of administrative words and phrases were compiled, such as the Government of Tamil Nadu’s *Tamīlnāṭṭu āṭṭic corṭaḷ* (‘Tamil Nadu administrative vocabulary English–Tamil’, 1968).⁵⁷ In many cases, early inscriptions were culled for lexical items that could be revived or repurposed for contemporary administrative use.⁵⁸ Once glossaries were created, it was yet another step to overcome what Sumathi Ramaswamy has described as ‘years of bureaucratic dependence on English’ before ‘instilling in its place the new habit of using Tamil’, into which laws and statutes might then be translated.⁵⁹

Competing Priorities: The Example of Assamese

The competing priorities and shifting alignment of groups such as the Orientalists, Anglicists, missionaries, and nationalists may be illustrated in the history of Assamese lexicography. *A Spelling Book and Vocabulary, in English, A’sa’mese, Singpho, and Na’ga*, compiled and published in 1839 by the American Baptist missionary Miles Bronson, appears to be the earliest-known lexicographical work describing Assamese. In 1867, Bronson’s American Baptist Mission published *A Dictionary in Assamese and English*, compiled with the assistance of Jaduram Deba Barua. Bronson argued in

⁵⁴ Ramaswamy, *Passions of the Tongue*, 145–9. ⁵⁵ James, *Colporul*, 234.

⁵⁶ James, *Colporul*, 236. ⁵⁷ For additional examples, see James, *Colporul*, 237.

⁵⁸ Sankaranarayanan, ‘Preparation of administrative terminology’.

⁵⁹ Ramaswamy, *Passions of the Tongue*, 163–4.

the introduction to his dictionary that Assamese was grammatically distinct from the Bengali language that had been adopted by the British as a language of administration. He claimed to have recorded the 14,000 terms defined in his dictionary 'as they dropped from the lips of the people', preferring those 'in daily use by the people' over those familiar to Bengali scholars. While including the 'more common Sanscrit words that are used in the *Puthis* [a popular narrative genre] and therefore known to the people', he warns that 'they are often used in Assamese, with a modified meaning, and a different pronunciation'.⁶⁰ It was largely through the efforts of Christian missionaries such as Bronson to establish Assamese as a distinct language that the General Department of Education of Bengal would resolve in 1873 to replace Bengali with Assamese as the official medium of instruction in Assam.⁶¹

Hemchandra Baruwā, the author of *Hema Kosha; or, An Etymological Dictionary of the Assamese Language* (published posthumously in 1900), used Webster's *American Dictionary of the English Language* and H. H. Wilson's *Dictionary, Sanscrit and English* as models for his dictionary, which contained 22,386 Assamese words explained with Assamese and English equivalents.⁶² He would break, however, with Bronson and the broader missionary approach by adhering to the instructions of the colonial government in 1873 and 1885 that Assamese should follow the Sanskrit orthography rather than spoken pronunciation.⁶³ In doing so, Hemchandra's lexicographic project was similar to the Sanskrit relexification efforts pursued by nationalist lexicographers working with other Indian languages in the late nineteenth and early twentieth centuries.⁶⁴ One colonial official, critical of Hemchandra's penchant for employing Sanskritic vocabulary in his children's primer of 1873, described a passage in that work in which more than forty of sixty-seven words were pure Bengali.⁶⁵

The *Chandrakānta abhidhān*, published in 1933 under the supervision of Debesvar Chaliha through the Assam Sahitya Sabha (Assam Literature Society), follows in the tradition of Bronson and Hemchandra by providing Assamese and English equivalents for its 36,819 entries, and adds citations

⁶⁰ Bronson, *Dictionary in Assamese and English*, iv; see also Sengupta, *Becoming Assamese*, 98.

⁶¹ Sengupta, *Becoming Assamese*, 87.

⁶² For the dictionary, published as Barua, *Hema Kosha*, see the biographical introduction by P. R. T. Gurdon (i–iii); word count from Gurdon, 'Review', 236; see also Singh, 'Lexicography of New Indo-Aryan', 2509.

⁶³ Sengupta, *Becoming Assamese*, 88, 157.

⁶⁴ Sengupta, *Becoming Assamese*, 156–9; Bhattacharya, 'Experts question roots of "Modern" Assamese'. Cf. Orsini, 'What did they mean by "public"?', 413–15.

⁶⁵ Sengupta, *Becoming Assamese*, 89.

from literary sources. Unlike its predecessors, the *Chandrakānta abhidhān* was compiled and published without missionary or government funding, representing for one reviewer ‘the cultural advance of the Assamese people’.⁶⁶ Revised by B. K. Kakati in 1962, it served as the standard work of Assamese lexicography for most of the twentieth century.⁶⁷

Assamese was being employed as early as 1795 as a defining language for the *Bar Amra*, a lexicon of Ahom preserved in multiple manuscript copies.⁶⁸ By 1920, it would share this role with English for Rai Sahib Golap Chandra Barua’s *Ahom–Assamese–English Dictionary*. Works produced after independence, such as the *Prasāsaniya paribhāṣā*, or *A Glossary of Administrative and Other Terms* (1966) and the *Hindī–Asamīyā Kośa* (‘Hindi–Assamese dictionary’, 1966), reflect the changing priorities of the Assamese state government. The *Asamīyā Baṛo Śabdakosha* (‘Assamese–Bodo dictionary’, 1987), H. W. Marak’s *Ku-bidik: A Garo–English–Assamese Dictionary* (2000), B. S. Kro’s *Akemi Karbi lamthe amarjong: A Karbi to Karbi–English–Assamese Dictionary* (2002), V. L. Kholar and D. U. Maslai’s *Tiwā mātpāḍiḥ: A Tri-Lingual Dictionary, Tiwā–Assamese–English* (2004), and Saikia et al.’s *Deurī–Asamīyā citra śabdāwalī* (‘Deuri–Assamese pictorial glossary’, 2014) demonstrate the growing importance of Assamese as a link language across the north-eastern region of India. Monolingual dictionaries of Assamese, however, are a recent phenomenon: the *Ādhunika Asamīyā śabdakosha* (‘Modern Assamese dictionary’) and D. Katakī’s *Natuna Asamīyā abhidhāna* (‘New Assamese dictionary’, 1996) are both approximately 700 pages in length and relatively comprehensive.⁶⁹

Moves Towards Standardization: The Case of Gujarati

Gujarati–English lexicography of the nineteenth century reflects the transition from early idiosyncratic wordlists prepared by individuals to the dominance of the modern standard dictionary form enabled by the growth of formal scholarly institutions on European models. The glossary of 463 terms appended to Robert Drummond’s 1808 *Illustrations of the Grammatical Parts of the Guzerattee, Mahratta & English Languages* has been called the first bilingual Gujarati dictionary.⁷⁰ It is more encyclopedic than properly lexicographic

⁶⁶ Gurdon, ‘Review’, 236 (word count), 237 (quotation).

⁶⁷ Singh, ‘Lexicography of New Indo-Aryan’, 2509. ⁶⁸ Morey, ‘Ahom and Tangsa’, 51–5.

⁶⁹ For the former, see Singh, ‘Lexicography of New Indo-Aryan’, 2509.

⁷⁰ The glossary is on unnumbered pages after Drummond, *Illustrations*, 36; for it, see Nagar, ‘Brief survey’, 22; Shastree, ‘Monolingual and bilingual dictionaries’, 30; Singh, ‘Lexicography of New Indo-Aryan’, 2511.

and, like other works of this period, acknowledges the variety of languages necessary for conducting business in western India. This shorter entry suggests the style of the work overall:

Bayillo or *Ba,elo*. (Guz.) Effeminate, weak; under the undue influence of woman; according to Indian illustration 'a husband in stays while his wife wears a turband'.

Drummond's entry *parsee*, by contrast, is seven pages long. Though little is known of the author (he was appointed as a surgeon in the Bombay Medical Establishment in 1798 and 'struck off the rolls of the Bombay Army on the 14th March 1809 having been lost at sea on his passage home'), we learn a great deal from the glossary about his personal opinions.⁷¹

Contrasted with this is Mirza Mahomed Cauzim's *Dictionary, Goojratee and English* of 15,000 entries (1846), which benefited from the assistance of the Jamsetjee Jeejeebhoy Translation Fund, a Parsi charitable organization established 'for the purpose of promoting the advancement of science and literature among their countrymen'.⁷² The fund purchased 'a manuscript dictionary compiled by the late Dulputram Bhugobhaee of Surat', which was of 'considerable assistance' to Cauzim's work. The work also benefited from the cross-pollination made possible from the publication of wide-ranging dictionaries in other South Asian languages. Cauzim, for example, acknowledged his debts to James Molesworth's Marathi–English dictionary (for which see below) and John Shakespear's *Dictionary, Hindustani and English* in the introduction to his dictionary, calling them 'two of the most comprehensive and elaborate lexica that have yet appeared in any of the Oriental languages'.⁷³ By the twentieth century, however, Gujarati lexicographers would not need to draw as extensively from dictionaries of other languages. The two-volume *Modern Gujarati–English Dictionary*, compiled by B. N. Mehta and B. B. Mehta (1925), contains 51,595 words and was an expansion of predecessors such as Narmada Śankara Dave's *Narma Kośa* (1861–73; 25,268 words), M. B. Belsare's *Pronouncing and Etymological Gujarati–English Dictionary* (1895; 35,138 words), and the Gujarat Vernacular Society's *Rūdhī prayoga Kośa* (1898; 35,678 words).⁷⁴

⁷¹ Mandlik, 'Editor's preface', xii.

⁷² Cauzim, *Dictionary, Goojratee and English*, xii; Shastree, 'Monolingual and bilingual dictionaries', 30–1.

⁷³ Cauzim, *Dictionary, Goojratee and English*, xii.

⁷⁴ Shastree, 'Monolingual and bilingual dictionaries', 32, 34; Singh, 'Lexicography of New Indo-Aryan', 2511–12.

Constructing a Regional Identity: The Case of Marathi

A Dictionary of the Mahratta Language (1810), compiled by William Carey in collaboration with the Maratha pandit Vaijnāth Śarmā, was among the many dictionaries produced for British students at the College of Fort William in Calcutta.⁷⁵ Carey's work was preceded by the anonymous *Gramatica marastta: a mais vulgar que se practica nos reinos do Nizamaxà e Idalxà* ('Grammar of Marathi, most commonly employed in the Nizām [Shahī dynasty of Hyderabad] and 'Ādil [Shahī of Bijapur] kingdoms'), published in Rome in 1778. The final section of the *Gramatica marastta* is a seven-page Marathi–Latin vocabulary organized thematically in fifteen sections and set in Roman type.⁷⁶ Carey's work, by contrast, was organized alphabetically, containing approximately 11,000 entries and employing the types in the Modi script cast for the missionary press at Serampore near Calcutta.⁷⁷ The cursive Modi script, despite then being the most popular means of writing Marathi, would eventually be phased out by the colonial government in favour of the *devanāgarī* script employed for Hindi and other languages across much of northern India.⁷⁸

Three works prepared under the supervision of James Molesworth – Jaganathshastri Kramvant's *Dictionary of the Marāṭṭa Language* (1829, with a supplement in 1831), Molesworth and Thomas and George Candy's *Dictionary, Murat, hee & English* (1831, revised edition 1857, 40,000 and 60,000 entries respectively), and Molesworth and Thomas Candy's *Dictionary, English & Marāṭhī* (1847) – have served as a foundation not just for subsequent Marathi dictionaries but also for the lexicography of other western Indian languages.⁷⁹ These works use the *devanāgarī* script and are based primarily on the Pune dialect of Marathi. Thomas Candy describes the process by which Molesworth commenced his work:

He took Todd's edition of Johnson, and transferred its columns to the pages of several blank folios, omitting only words altogether obsolete or unsuitable. With the English Dictionary thus arranged before him, he took up the Marāṭhī and English Dictionary, and went over its pages, transferring the words to the blank spaces opposite their corresponding terms in English.⁸⁰

⁷⁵ Naregal, *Language Politics*, 162–3.

⁷⁶ *Gramatica Marastta*, 39–45; see also Zwartjes, *Portuguese Missionary Grammars*, 67–75.

⁷⁷ Naregal, *Language Politics*, 162–3. ⁷⁸ Naregal, *Language Politics*, 164, 167.

⁷⁹ For Molesworth and his dictionaries, see N. G. Kalelkar, in *Molesworth's English and Marathi Dictionary* (1992), ten–eighteen [sic].

⁸⁰ T. Candy, in Molesworth, *Dictionary, English and Marāṭhī*, 1.

Molesworth's dictionaries would form the basis for most Marathi dictionaries to follow, including the seven-volume *Mahārāshṭra śabdakośa* ('Maharashtra dictionary', 1932–8), the most comprehensive and encyclopedic monolingual Marathi dictionary to date.⁸¹

From Poetry to Politics: Hindi, Urdu, and Hindustani

The colonial state saw Hindustani (originally a British term for Hindi-Urdu written in the Arabic script) as an expedient *lingua franca* widely understood across much of northern India. It was subsequently deployed as the official language of the lower levels of administration, with some colonial officials seeking to extend its use even to areas where it was not at the time commonly spoken. The middle decades of the nineteenth century correspond with the rapid expansion of vernacular printing and publication of monolingual dictionaries, particularly in Hindustani/Urdu. By the turn of the twentieth century, however, advocates for Hindi written in the left-to-right *devānāgarī* script were successful in displacing Urdu as the sole official vernacular, beginning with the Central Provinces (corresponding largely with present-day Madhya Pradesh) in the 1870s, Bihar in the 1880s, and the North-Western Provinces and Oudh (present-day Uttar Pradesh) in 1900.⁸²

In response to the changing needs of the colonial government, lexicographers began adapting existing genres to include terms drawn from English. In 1845, four years prior to the British annexation of the Punjab, Niyāz 'Alī Beg 'Nakhat' of Delhi published a monolingual Urdu dictionary, *Maḥẓan-i favā'id* ('Treasury of benefits'), described in its printer's colophon as 'including *iṣṭilāḥāt* [terminologies] and *muḥāvarāt* [idioms] and *miṣl* [proverbs] of the Urdu language with verses and poems of famous poets of India of a superior nature'.⁸³ Its recent editor has remarked that the *Maḥẓan-i favā'id* is similar in format to earlier texts that focused in part or whole upon Urdu idioms. These include the 290 entries of the *Shams al-bayān fī muṣṭalahāt al-hindūstān* of Ṭāpish, mentioned above; the fourth and fifth chapters of the *Daryā-yi laṭāfat* ('Sea of delicacy', 1808) by Inshā (on 'The idioms of the language of Delhi' and 'The language and idioms of Delhi's women'); a glossary of phrases used by women included in the 1828 *Dīvān-i rekḥṭī* ('Anthology

⁸¹ For it, see Singh, 'Lexicography of New Indo-Aryan', 2514.

⁸² King, *One Language, Two Scripts*, 69–79.

⁸³ See Husain, 'Pesh guftār', in Nakhat, *Maḥẓan-i favā'id* (1998), viii; the colophon is translated in Hakala, *Negotiating Languages*, 138.

of female-voiced poetry') by Sa'adat Yār K̲h̲ān 'Rangīn'; and the *Risālah-i qavā'id-i šarf o naḥv-i urdū* ('Treatise on the rules of Urdu grammar') by Imām Baḳḥsh 'Šahbā'ī, completed the same year as the *Maḳḥzan-i favā'id* and published in 1849.⁸⁴ However, the primary written sources of Nakhat's work – which, with 12,858 entries devoted to terminology, 116 to idioms, and 118 to proverbs, dwarfed these earlier specialized dictionaries of Urdu – were the *divāns* (collections of poetry) of Urdu poets. This preference for poetic citations over those drawn from prose conforms with established methods of Persianate lexicography.⁸⁵

Once published, however, the work seems to have fallen into obscurity, with no reprints until a new edition appeared in 1998. Despite (or perhaps because of) this obscurity, Sayyid Aḥmad Dihlavī borrowed heavily and without attribution from Nakhat's dictionary for his magisterial *Farhang-i āṣafiyah* ('Dictionary of the Aṣaf Jāhī dynasty', 1918). Sayyid Aḥmad, for example, uncritically reproduced errors in attribution for the citations appearing in Nakhat's earlier work.⁸⁶ Nevertheless, Sayyid Aḥmad complained in the 1918 introduction to the second edition of his *Farhang-i āṣafiyah* of himself being the victim of unprincipled plagiarists, accusing several individuals of having 'committed daylight robbery' by incorporating materials from his first edition into their own dictionaries.⁸⁷

Sayyid Aḥmad's career neatly encapsulates the synergies emerging between South Asian and European lexicographers during the nineteenth century. He served for seven years as assistant to S. W. Fallon, the primary author of the innovative *New Hindustani–English Dictionary, with Illustrations from Hindustani Literature and Folk-Lore* (1879–83). This assistantship provided professional accreditation for his future lexicographic endeavours; it also opened the door to opportunities within the colonial state. In addition to the variety of positions he maintained in the colonial education system, Sayyid Aḥmad was appointed superintendent of the Ethnographic Survey in Delhi.⁸⁸ He was given the title K̲h̲ān Šāḥib by the colonial government in 1914 and became a pensioner of the wealthy Nizām of Hyderabad.⁸⁹ However, since he was an Indian scholar who worked under British lexicographers, critics in the twentieth century derided Sayyid Aḥmad both for

⁸⁴ Inshā, *Daryā-yi laṭāfat*, 128–34; Rangīn and Inshā, *Divān-i Rangīn va Inshā*; Šahbā'ī, *Risālah-i qavā'id-i šarf o naḥv-i urdū*.

⁸⁵ Hakala, *Negotiating Languages*, 9–10. ⁸⁶ Hakala, *Negotiating Languages*, 134–47.

⁸⁷ Husain, 'Pesh guftār', in Nakhat, *Maḳḥzan-i favā'id* (1998), xv.

⁸⁸ Minault, 'Sayyid Ahmad Dehlavi', 182.

⁸⁹ Bukhārī, 'Maulvī Sayyid Aḥmad Dihlavī', I.17.

retaining too many elements of the pre-colonial Persianate tradition and also for violating the cultural mores of his readership.

Sayyid Ahmad could not resist inserting his own poetry among the citations included in his entries. While this was consistent with an earlier mode of patronage in which an author's status as a literary figure was the primary source of authority, it did not conform with the modes of accreditation valued in the emerging market for mass-produced pedagogical materials. By the early twentieth century, lists of diplomas and expert testimonials had become a standard component of dictionaries' front matter. As the scope and reach of the colonial education system grew, so too did the opportunities for patrons to recoup the substantial costs of producing the monumental dictionaries produced in the late nineteenth and early twentieth centuries. This encouraged the rise of a corporate mode of production, whether through the pooling of private capital investments in publishing houses or through the association of scholars with institutions dedicated to the compilation of dictionaries and ancillary activities like the editing of critical editions of texts.⁹⁰ The substantial team of scholars assembled by Fallon in the 1870s anticipated a further division and specialization of labour: some were tasked with collecting folksongs and proverbs from the countryside, while others were made responsible for researching etymologies or collating notecards. Dictionaries produced in this way were no longer the product of a solitary drudge, toiling alone, but rather could draw in their lexis upon a wider range of human experiences.

Maulvī 'Abdul Ḥaq, the venerable leader of the *Anjuman-i Taraqqī-yi Urdū* (Society for the Advancement of Urdu) in India and, following independence, Pakistan, attributed the inclusion of 'obscene' (*fōḥsh*) materials in the *Farhang-i āṣafiyah* to its author's years of 'companionship [*faiz-i ṣōḥbat*] with Dr. Fallon'.⁹¹ Sayyid Ahmad argued in the introductory essays of the *Farhang-i āṣafiyah* for the 'national' significance of his project as 'an extremely important national [*qaumī-mulkī*] work'.⁹² For both, the lexical domains staked out by lexicographers are as much an articulation of political intentions as the maps prepared by cartographers. Following Sayyid Ahmad's death in 1918, however, that map would undergo a dramatic series of changes: India was no longer a British colony, Pakistan had been established as an independent homeland for South Asian Muslims, Hindi in the *devanāgarī* script was declared the official language of India, and Urdu in the Arabic script the official language of Pakistan.

⁹⁰ Stark, *Empire of Books*, 64–83. ⁹¹ Hakala, *Negotiating Languages*, 153.

⁹² Dihlavi, *Farhang-i āṣafiyah*, 1.51.

The *Nāgarīpracārīnī sabhā*, a group founded in 1893 to promote the use of *devanāgarī* over the more popular Perso-Arabic and *kaithī* scripts used in Hindi-Urdu writing, sponsored and published Śyām Sundar Dās' *Hindī śabdasāgara* (4 vols.; 1916–28, ²1965–75). This monumental work, based on Dās' efforts over decades to collect early 'Hindi' manuscripts, was part of a greater nationalist effort to claim the different languages spoken across northern India (for instance Braj Bhāṣā, Avadhī, Marwaṭī, and Bhojpurī) as different varieties of a single, 'national' language – Hindi. The *Nāgarīpracārīnī sabhā*, despite its historical ecumenism, prescribed a very specific modern form of Hindi: its grammar, like that of the prestige dialects of Urdu, would be derived from the speech of the Delhi region. In a pattern that became familiar across the Indian subcontinent, its proponents would insist that terms of Arabic, Persian, and Turkish origin, many thousands of which were commonly understood across northern India, be replaced with Sanskrit-derived *shuddh* ('purified, unpolluted') Hindi equivalents.

South Asian Lexicography After Independence

Maulvī 'Abdul Ḥaq, whom we have just seen criticizing the *Farhang-i āṣafiyah*, was himself involved in compiling the *Lughat-i kabīr-i Urdū* ('Great dictionary of Urdu'), a project that may be traced back to 1917 when the Niẓām of Hyderabad endowed the Anjuman with a fund to develop a dictionary of scientific and technical terms.⁹³ Following Ḥaq's death in 1961, the project, a historical dictionary of the Urdu language modelled on the *Oxford English Dictionary*, was published between 1973 and 2010 in twenty-two large volumes. That Ḥaq could also laud as 'extremely praiseworthy' Sayyid Aḥmad's examination of 'certain expressions and Hindi terms which up to now no Urdu lexicographer has recorded' suggests that linguists of the post-independence period would share the ambivalence regarding prescriptivist and descriptivist ideals that has been amply documented in English.⁹⁴ In June 2017, a full century after it was first conceived, the online version of 'Abdul Ḥaq's dictionary, known now by the title *Urdū lughāt: tārikhī uṣūl par* ('Urdu dictionary: on historical principles'), was inaugurated by the government of Pakistan.⁹⁵

Beginning as early as the middle of the nineteenth century with organizations such as the Vernacular Translation Society, and reaching a peak in the

⁹³ King, *One Language, Two Scripts*, 163; Datla, *Language of Secular Islam*, 73–4.

⁹⁴ See, e.g., Skinner, *Story of Ain't*.

⁹⁵ Salman, 'Multivolume Urdu dictionary set to go online'.

middle decades of the twentieth century with a flourishing of government-sponsored bilingual technical dictionaries, evidence of the failures of relexification projects in present-day India and Pakistan may be seen in the paucity of such works in recent decades and the persistence of English as a de facto official language in both countries some seven decades after their independence.⁹⁶

Despite a stated desire to limit Persian-, Arabic-, Turkish-, and English-derived words in its official glossary of administrative terms, the Indian Standing Commission for Scientific and Technical Terminology was nevertheless compelled to include the occasional ‘non-Hindi’ word – *ghairhājirī* (from the Persian *ghair-hāzīrī*) for ‘absence’, *teñdar* for ‘tender’, and *hisāb* (from the Arabic *ḥisāb*) for ‘account’ – in its *Samekita Praśāsana śabdāvalī*, *Āngrezī–Hindī/Consolidated Glossary of Administrative Terms, English–Hindi*. The primary objectives of the project, ‘to maintain as far as practicable a uniform all India character’ and ‘to achieve the maximum possible identity of terms in Indian languages’, had to be balanced against ‘simplicity, precision and easy intelligibility besides the currency that some of these terms have already gained in the various regions’. The latter option would be pursued, however, only in the absence of Sanskrit-derived equivalents.⁹⁷ Advocates for Hindi and Urdu sought to populate a single language with two separate lexica, each capable of expressing any concept in any context or medium as well as any other ‘national language’. For Urdu (and, subsequently, Hindi), this entailed the production of technical dictionaries providing equivalents for European terminologies, for instance in medicine (where Breton’s *Vocabulary* of 1825 is an early example); law (where nineteenth-century dictionaries by, for instance, Fallon have successors such as M. M. T. Khan’s *Dictionary of Law* of 2004); and many other fields (Latif et al.’s *Dictionary of Technical Terms: Sociology, English–Urdu* of 1970 is one example).⁹⁸ Once situated on the world stage, Hindi and Urdu lexicographers could direct their gaze inwards: the linguistic conquest of a multilingual nation-state requires the translation of less prestigious regional languages and even criminal cant into the national language.⁹⁹

The Cold War and realignment of post-independence India and Afghanistan with the Soviet Union encouraged a new focus by lexicographers

⁹⁶ See, for instance, Ramaswamy, *Passions of the Tongue*, 228–33; Rahman, ‘Language policy and education’, 383.

⁹⁷ *Samekita Praśāsana/Consolidated Glossary*, vii–viii.

⁹⁸ See, for example, Datla, ‘Making a worldly vernacular’, 85–95.

⁹⁹ For a Telugu–Urdu dictionary, see Redḍi, *Shabdārtha-Chintāmaṇi*; for criminal cant, see Akbar, *Muṣṭalahāt-i ṭhagī*.

on the production of bilingual Russian dictionaries. The career of the Soviet scholar Vasilij Matveevič Beskrovnyj is worthy of special mention. In 1932, early in his career, he produced a specialized vocabulary of 'social, political and economic terminology' in Hindi.¹⁰⁰ Two decades later, he had completed a 19,000-entry *Urdu-russkij slovar'* ('Urdu–Russian dictionary', 1951) with the assistance of V. Krasnodembskij. Unlike the great British dictionary projects of the Victorian era, Beskrovnyj preferred to draw his corpus from more contemporary sources, including modern literature, journals, newspapers, and scientific texts. In their coverage of contemporary language, one authority considered the dictionaries of Beskrovnyj 'markedly superior' to any Urdu–English or Hindi–English dictionary prepared in the first half of the twentieth century.¹⁰¹ His *Khindi-russkij slovar'* ('Hindi–Russian dictionary', 1953; ²1959; 35,000 and 40,000 entries respectively), was the product of nearly two decades of work conducted under the aegis of the USSR's Oriental Institute of the Academy of Sciences.¹⁰² Beskrovnyj led a subsequent effort to produce a vast comprehensive Hindi–Russian dictionary of 75,000 words in 2 volumes that was published in 1972, edited by A. S. Barkhudarov. M. G. Aslanov's massive *Afgansko-russkij slovar' (puštu)* ('Afghan–Russian dictionary (Pashto)', 1966; 50,000 words) and its Russian–Afghan counterpart of 1973 by K. A. Lebedev et al. encouraged the development of concise bilingual dictionaries prepared for Afghan schoolchildren seeking to learn Russian.¹⁰³ Indian scholars also endeavoured to produce bilingual Russian dictionaries. Weer Rajendra Rishi prepared the *Rūsī–Hindī śabdakośa* ('Russian–Hindi dictionary', 1957), a work to which the first Indian prime minister, Jawaharlal Nehru, contributed a glowing preface.

In light of their bilateral strategic alliance, it is interesting that Chinese scholars do not appear to have taken the same interest in the languages of Pakistan, though two recent Urdu–Chinese dictionaries are the result of increased scholarly investments in South Asian linguistics.¹⁰⁴ The *Wuerdu yu Han yu ci dian* ('Urdu–Chinese dictionary', 2014) is more comprehensive, with 65,000 Urdu entries on almost 1,500 pages. Sponsored by the government of the People's Republic of China, it was compiled by Kong Julan at Peking University. The university established its programme in Urdu in 1954 and began laying the groundwork in the 1980s for an Urdu–Chinese

¹⁰⁰ Pořízka, 'Hindī–Russian and Urdū–Russian dictionaries', 185.

¹⁰¹ Pořízka, 'Hindī–Russian and Urdū–Russian dictionaries', 185. ¹⁰² Vulfson, 'Slavica', 162.

¹⁰³ See Bečka, 'Afghan dictionaries'.

¹⁰⁴ The authors wish to thank Jinhee Song and Yan Liu for their assistance in examining the Chinese dictionaries cited here.

dictionary. This project was revived in 2004, requiring another ten years to complete. The *Han yu Wuerdu yu fen lei ci dian* ('Chinese–Urdu classification dictionary', 2013) by Chen Xiang is smaller, with 12,000 entries, and is part of a series of dictionaries devoted to less commonly studied languages. Like other vocabularies in the series, it organizes its contents into five general categories and forty-four specific themes such as art, public safety, environmental protection, and marketing.

Conclusion

Contemporary South Asian lexicographic experiments have drawn from a wide range of languages and resources. Classical forms of lexical analysis drawn from the Sanskrit *vyākaraṇa* tradition facilitated distinctions between distinct classes of vocabulary items, enabling the identification of language families and their associations with ethnological populations of speakers. It was not uncommon, however, for Sanskrit and regional forms of lexicography that modelled themselves on Sanskrit to group words together based on their acceptability for use in poetic compositions, organizing them by topic rather than alphabetically. Persian and later English lexical influences reshaped the languages of literature, bureaucracy, and administration, and introduced forms of lexicography that served new pedagogical functions, offering glosses, equivalences, and definitions for unfamiliar terms as the uses of these languages expanded beyond the ruling classes. Most significant for lexicographical trends during this period, however, has been the imagination of languages as parallel rather than complementary to one another, and the emergence of languages as foundations for political distinctions and claims. Movements to purify languages and purge them of 'foreign' influences, the resurrection and repurposing of archaic vocabulary drawn from inscriptions or early literary works to substitute for loanwords, and efforts to associate lexical items more closely with the identities of their speakers are key transformations during this period.

Arabic from c. 1800

JAN HOOGLAND

This chapter will cover the lexicography of the Arabic language(s) after 1800. Arabic is spoken by more than 200 million people – and as the *(s)* after *language* in the previous sentence indicates, there is not just one Arabic language. Considering Arabic means considering the different variants of the language: the classical language of the Qurʾān, called Classical Arabic; the mostly written (but also spoken) language of the present-day Arab world, called Modern Standard Arabic; and the spoken colloquial languages or dialects of the twenty-two countries where Arabic is the official language, as well as of countries such as Turkey, Iran, Afghanistan, and some areas in sub-Saharan Africa, where we find minority groups who speak varieties of Arabic.

A situation in which a written and a spoken variant of a language coexist is called diglossia.¹ In the case of Arabic,

the ‘high’ variant, Modern Standard Arabic, [is] the accepted language for any written and official spoken discourse. Meanwhile, a multitude of ‘dialects’ or ‘colloquials’ are the languages people actually speak. These are officially known as dialects, but are often mutually unintelligible, and can be considered different languages on purely linguistic grounds. It is hard to say how many exist, as there is a dialect continuum, but in many countries the dialect of the capital has the major status.²

Classical Arabic can be considered a dead language that is still being used for religious practice, to enjoy the classical cultural heritage, and in some cases for literary production. Nowadays, most lexicographic projects with Arabic involve Modern Standard Arabic or one of the dialects, and this chapter will therefore focus on these varieties rather than the classical language.

I am grateful to Tressy Arts for reading this chapter and correcting my English.

¹ Defined in Ferguson, ‘Diglossia’.

² Arts, ‘Making of a large English–Arabic/Arabic–English dictionary’, III.

Early Arabic Lexicography

As described in Chapter 8, the Arabs have a long tradition in lexicography, and with them lexicographers of other linguistic backgrounds belonging to the Islamic world, for instance Persian and Turkish (see Chapters 11 and 17). However, this long tradition is not an unqualified asset: it has also been a cause for a hesitant and less open attitude to modern principles in lexicography:

there was always a strong conservative influence to preserve and perpetuate the language of the Qur'ān and the Ḥadīth [registration of deeds and statements of the prophet Muḥammad]. This was due partly to the fact that the Qur'ān and prayer were not, normally, translated. The lexicographers helped to keep the written language static, and to aid the understanding of it, as the spoken dialects diverged more and more from it. So strong were religious sanctions on this point, and so well did the lexicographers do their work, that these spoken dialects were not able to develop into independent languages, as Italian, Spanish, Portuguese, and French were able to develop out of Latin.³

This attitude can be partially ascribed to the fact that Arabic is seen as divine in origin, since in the Qur'ān it is said that Arabic is the language in which the Qur'ān was sent down to earth.⁴

This linguistic conservatism causes a rigidity in language-users' minds – and in the policies of authorities, be they linguistic or non-linguistic, a reluctance to acknowledge changes in the language or to accept words being added to the language by its users. The consequences of this attitude for the development of modern terminology will be discussed later.

It should be added that most of these observations pertain to both monolingual Arabic lexicography and bilingual lexicography involving Arabic and another language from the Islamic world. When contacts with Europe intensified, through colonization, realization grew among Arabs themselves that terminological development had been sadly neglected for centuries, and bilingual lexicography was spurred into activity.⁵ For bilingual lexicography involving European languages the situation was totally different, since not many Arabic–European language dictionaries were published before the nineteenth century, and those that did exist were compiled by Western scholars as decoding dictionaries of Arabic (they are mentioned briefly in Chapter 14).

³ Haywood, *Arabic Lexicography*, 116. ⁴ Seidensticker, 'Lexicography: Classical Arabic'.

⁵ Khoury, 'Dictionnaires arabes bilingues', 42–4.

The Arabic Language and Lexicography

Before we turn to the modern dictionaries themselves, it is necessary to mention a few more characteristics of Arabic, since they have consequences for dictionary compilation and presentation of entries.

The first of these is the Arabic script. Arabic is written from right to left, with a consonant-based script where diacritic signs representing the vowels can be added to the consonants. These vowels are omitted in most written or printed texts, but there are at least three categories of texts that should contain full vocalization: the Qur'ān, textbooks for learning the language, and dictionaries. However, different dictionary compilers have made different choices, especially in the past when typesetting of vowelised Arabic script was very difficult. Some dictionaries contain unvowelised Arabic script with a transliteration in Latin characters, including the vowels. The fact that Arabic is written from right to left can complicate the typesetting of a dictionary in which the other language is oriented left to right, since many thousands of lines will be filled with two types of text, written in opposite directions. Again, different dictionary compilers have made different choices.

The second is word formation, which is built on a root-and-pattern system, explained thus by Ernest McCarus in his excellent overview in the *Encyclopedia of Arabic Language and Linguistics*:

The 'root' is a series of typically three consonants, occurring in a fixed sequence, that has lexical identity. The consonants making up a root are called 'radicals' and may number as many as three to five. Thus the root *k-t-b* means 'to write; writing', *s-k-n* means 'to dwell, reside', and *ṭ-ʿ-m* means 'to eat, savour'. A 'pattern' is a fixed framework of consonants and vowels that likewise has lexical meaning, e.g. the pattern *maF^caL* means 'noun of place'; the root is variable but the *ma* and the vowel *a* before the last radical are obligatory. Substituting a root *x* in place of the model root *f-ʿ-l* produces a noun meaning 'place where *x* takes place', like *maktab* 'office; desk'; *maskan* 'dwelling place, residence'; and *maṭ^cam* 'restaurant'.⁶

The third lexicographically significant characteristic of Arabic follows from the second. Modern dictionary compilers can decide to order the headwords of their dictionary in two ways (others were used in the past: see Chapter 8): either alphabetical order or an ordering in which all words belonging to a given root are presented together in the dictionary. The first principle

⁶ McCarus, 'Modern Standard Arabic'.

would order the example words from the paragraph above in the sequence *sakana*, *fa^cala*, *ta^cama*, *kataba*, *maskan*, *ma^cam*, *maktab* (Arabic alphabetical ordering), whereas the second principle would result in the sequence *sakana* + *maskan*, *ta^cama* + *ma^cam*, *fa^cala*, *kataba* + *maktab*. Some examples of root-ordered bilingual dictionaries are Hans Wehr's *Arabisches Wörterbuch* (1952) and *Dictionary of Modern Written Arabic* (1961), Jan Hoogland's *Woordenboek Arabisch–Nederlands* (2003), and Tressy Arts' *Oxford Arabic Dictionary* (2014). Examples of alphabetically ordered bilingual dictionaries are Lorenz Kropfitch's *Langenscheidt Handwörterbuch Arabisch–Deutsch* (1996) and Ruhi Baalbaki's *Al-Mawrid qāmūs 'arabīy-inklīzī* (1987). A root-ordered monolingual dictionary is Buṭrus al-Bustānī's *Muḥīṭ al-Muḥīṭ* (1867–70); alphabetically ordered is Ibrāhīm Muṣṭafā's *Mu 'jam al-wasīṭ* (1960–1).

Finally, the availability of large corpora makes it possible to assess the size of the vocabulary of prose in Modern Standard Arabic. By my calculation, the 24,000 entries of Hoogland's *Woordenboek Arabisch–Nederlands* (2003) sufficed to cover 99.95 per cent of the vocabulary of modern texts. Mark Van Mol stated that the 17,000 entries in his *Leerwoordenboek Arabisch–Nederlands* covered 99 per cent. Petr Zemánek states in the introduction to his Arabic–Czech dictionary, which is entirely corpus-based, that 21,000 entries suffice to cover the vocabulary of Modern Standard Arabic, noting that 15,000 of the 36,000 headwords in the 1985 edition of Wehr's *Dictionary of Modern Written Arabic* did not occur in his corpus of modern texts. It seems safe to conclude that a number of entries of little more than 25,000 suffices to cover the general vocabulary of present-day Modern Standard Arabic, and this figure is worth bearing in mind when considering the entry counts of some dictionaries of the modern language.

Monolingual Arabic Lexicography

Since neither Classical Arabic nor its present-day variant Modern Standard Arabic is the mother tongue of any Arabs, there has always been a need for tools to help users of Arabic to improve and expand their knowledge of the written language, namely monolingual dictionaries. Just as Arabic monolingual lexicography in the pre-modern era served to maintain a conservative form of Arabic, Ahmad Taher Hassanein has argued that modern Arabic–Arabic dictionaries generally resemble medieval Arabic dictionaries and can even be seen as an extension of the medieval tradition.⁷

⁷ Hassanein, 'Lexicography: monolingual dictionaries', 39.

The modern period starts with the publication in 1867–70 of the *Muḥīṭ al-Muḥīṭ* by Buṭrus al-Bustānī. Al-Bustānī was not just a lexicographer; he also compiled an encyclopedia (*dā'irat al-ma'ārif*) and was a leader of the Arab Renaissance (*Nahḍa*) movement that strove to revive, and create a love for, the Arabic language, making Arabic into a tool for Arabs to express their thoughts and ideas in the modernizing world of the nineteenth century.

Another major monolingual dictionary is the *Mu'jam al-wasīṭ*, edited by Ibrāhīm Muṣṭafā and issued by the Arabic Language Academy in Cairo in 1960–1. It runs to around 30,000 entries. Its perspective is modern in so far as it excludes entries for archaic terms that are irrelevant to modern usage. However, it is in need of an update to include vocabulary such as the Arabic words for 'notification' (*'i'lām*), 'scientific' (*'ilmīy*), 'to secularize' (*'almana*), and many others.⁸ A major shortcoming, also present in other Arabic dictionaries, both monolingual and bilingual, is the ordering of different senses of words: the most frequent senses are not presented at the beginning of the entry, and clear sense indicators are lacking.

Bilingual Dictionaries: The Nineteenth Century

For Western scholars, Georg Wilhelm Freytag was the link between the historic era and the era covered in this chapter. Freytag published an Arabic–Latin dictionary based on the *Qāmūs* of al-Fīrūzābādī (see Chapter 8), rearranging the entries, though he claimed that he had also consulted the *Ṣiḥāḥ* of al-Ġawharī (see Chapter 8) and the *Lexicon* of Golius (see Chapter 14). Since the *Qāmūs* is a lexicon without illustrative examples and quotations, Freytag's dictionary does not contain any examples either, and the different meanings of Arabic words are explained without quotations from Arabic texts. According to John Haywood, 'It was a mere word-list with meanings; whereas what was needed was a European "Lisān" [the reference is to Ibn Manẓūr, *Lisān*, for which see Chapter 8] which would not only give the meanings of words explicitly, but would also illustrate them amply by examples drawn from literature.'⁹

Edward William Lane attempted to fill that gap. His work can be considered a transition from the age when Latin was used as the language of science, as it still was by Freytag, to the age of the vernaculars, and as a passage from the Arab tradition of lexicography to the Western tradition.

⁸ See Abou El Aazm, 'al-Luġa al-'arabiyya'. ⁹ Haywood, *Arabic Lexicography*, 124.

Lane was a remarkable Arabist who published three major works which made him famous in different disciplines.¹⁰ He spent twelve years of his life in Egypt, in three different periods. His first visit to Egypt took place because the climate was considered beneficial to his health, but when there he was struck by a growing interest in the language and culture of Egypt, and his years in that country resulted first in a famous ethnographic study, *An Account of the Manners and Customs of the Modern Egyptians* (1836), for which he was praised in his own time, and was later described by Edward Said as an outstanding example of an Orientalist from Europe.¹¹ His second publication constitutes a landmark in the reception of literature from the Arab world in the form of an English translation of the Thousand and One Nights story collection, in English better known as the *Arabian Nights*. But Lane's *magnum opus* is beyond all doubt his *Arabic-English Lexicon*, which he had not completed at the time of his death in 1876, when he had reached the letter *qāf*, leaving the last seven of the twenty-eight letters of the Arabic alphabet incompletely treated. His nephew Stanley Lane-Poole saw the rest of the work through the press, but this last part is not of the same level of refinement and completeness as the rest. Lane started working on the *Lexicon* during his third, last, and longest period in Egypt, between 1842 and 1849, but continued to work on it in an admirably zealous way after his return to England until his death.

As noted in Chapter 8, Lane based his *Lexicon* on the *Tāǧ al-ʿarūs* by al-Zabīdī. He translated all the definitions of that monolingual dictionary into English, adding to it from his own readings of Classical Arabic. Since the *Tāǧ* was available only as a manuscript, Lane had his own copy of it written out by his assistant. It consisted of 13,000 folios and kept the assistant working for more than ten years. After returning to England he would regularly receive an instalment of this copy and thus be able to proceed with his work. Lane could commit himself to this immense task only because he was financed during the preparation and publication of the *Lexicon* by Algernon Percy, duke of Northumberland. Lane's *Lexicon* contains around 34,000 entries and is ordered by root. Although it is an Arabic-English lexicon, the columns are oriented from left to right, since the majority of the text is English. The Arabic text is not fully vocalized but it contains many vowels, which means the setting and checking process before printing must have been very time-consuming in those days. Haywood praises Lane's *Lexicon* as follows: 'It is difficult to conceive

¹⁰ A biography is Thompson, *Edward William Lane*.

¹¹ Said, *Orientalism*, 158–64 and *passim*.

a better dictionary in the accuracy of its definitions, and the fulness of its examples. It is surely one of the finest dictionaries ever written in any language.¹²

In 1881, the Dutch Arabist Reinhart Pieter Anne Dozy published an Arabic–French supplement to Lane’s monumental work and other dictionaries, titled *Supplément aux dictionnaires arabes*. Dozy included many post-classical words, particularly those peculiar to Arab Spain, as he was a specialist on the history of the Arabs in Spain.¹³ Dozy’s *Supplément* was based on a corpus of a rich variety of texts, also from a later period, which makes it modern in its time in two aspects: the use of a corpus and the inclusion of texts from a modern period in that corpus.

So far, all the dictionaries which have been mentioned cover the Classical Arabic language. When we proceed to discuss the twentieth-century dictionaries, the focus will switch to Modern Standard Arabic and the Colloquial Arabic languages.

The Twentieth Century

In the early stages of bilingual lexicography, the Orientalists followed Classical Arabic models such as the *Šiḥāḥ* of Ġawharī and the *Qāmūs* of Fīrūzābādī very closely with respect to their contents, although they adopted the modern arrangement of ordering the roots according to the first radical. In the third decade of the twentieth century, the German Arabist August Fischer believed that this practice was no longer adequate; what was needed was an Arabic lexicon based on historical principles, which should illustrate every word and meaning from actual use in literature. His plans were so detailed and thorough that he was able to convince the Arabic Language Academy of Cairo to enable him to realize them. However, his thorough working method was so time-consuming that he was unable to finish this immense task before his death.¹⁴

Haywood divides the twentieth century in terms of influential Arabic–European language dictionaries: the ‘Elias era’ between 1922 and 1952 and, implicitly, the ‘Wehr era’ after it.¹⁵ The ‘Elias era’ dawned in 1913, when Elias A. Elias published the first edition of his English–Arabic dictionary, followed in 1922 by the first edition of his Arabic–English dictionary; both have appeared in many editions. The latter was aimed at speakers of both Arabic

¹² Haywood, *Arabic Lexicography*, 125. ¹³ Haywood, *Arabic Lexicography*, 126.

¹⁴ Haywood, *Arabic Lexicography*, 125–6.

¹⁵ Haywood, ‘Bilingual lexicography with Arabic’, 3090.

and English, according to introductions in both languages, although the English introduction is rather short. However, in a dictionary meant to serve as an encoding dictionary for English, one might expect some grammatical information or usage notes on the English, but these are totally absent, nor is there essential information on the Arabic, so neither decoding nor encoding users are properly served. The macrostructure is vast – the English–Arabic dictionary apparently grew from 32,000 headwords in 1913 to 69,000 by 1962, and the Arabic–English dictionary from 45,000 words in 1922 to 64,500 by 1962 – but it is doubtful that all of Elias’ Arabic headwords were in actual use in Egypt at the beginning of the twentieth century.¹⁶ Some of them may have been newly coined as equivalents for English words in the English–Arabic part before appearing, after reversal, in the list of headwords of the Arabic–English part.

The end of the ‘Elias era’ is marked by the work of the German Arabist Hans Wehr, who in 1952 published the first edition of his Arabic–German dictionary under the title *Arabisches Wörterbuch für die Schriftsprache der Gegenwart* (‘Arabic dictionary for the modern written language’). Wehr had started working on this dictionary during the Second World War. It was then intended to facilitate the translation of Hitler’s *Mein Kampf* into Arabic, and subsidies were received for the compilation process from the German Ministries of Foreign Affairs and Education.¹⁷ Since the Arabic–German dictionary was intended as a basis for a German–Arabic dictionary, as *Mein Kampf* had to be translated from German into Arabic, this would seem a very time-consuming process. To my knowledge the German–Arabic dictionary based on Wehr’s work has never seen the light of day; nor do I know of the completion of the official translation of *Mein Kampf*, though prints of its Arabic translation (*kifāhī*) are freely available in different Arab countries.

Since Hans Wehr was physically handicapped as a result of polio in his youth, he was not drafted into the German army, though some of his assistants were. The compilation process had been completed during the war, but due to different complications in post-war Germany, the first edition of the Arabic–German dictionary did not appear until 1952. In 1959 Wehr published a supplement to the first German edition. In 1985 the supplement was incorporated into the dictionary, and a new German edition was published. In 1961 the first Arabic–English edition of Wehr’s dictionary appeared as *A Dictionary of Modern Written Arabic*, after it had been translated and edited

¹⁶ Entry counts from Haywood, ‘Bilingual lexicography with Arabic’, 3090, and Elias, *Modern Dictionary Arabic–English*, 9th edn (1962), title-page verso.

¹⁷ Haakh, ‘Adolf Hitlers “Mein Kampf” und seine arabischen Übersetzungen’.

by the American Arabist J Milton Cowan. Wehr's 1959 supplement was included in this first English edition, so until 1985 the English version was, in Cowan's words, 'more accurate and much more comprehensive than the original version, which was produced under extremely unfavourable conditions in Germany during the late war years and the early post-war period'.¹⁸

Arts describes Wehr's dictionary as follows:

Though groundbreaking at the time, it is nowhere near comparable to modern bilingual dictionaries for most other languages. Obviously the word list is outdated, but also the presentation of the entries is not really what one would expect from a modern dictionary: long lists of possible English translations are given for most Arabic words, without any guidance for the user on which translation to choose in which context; no word senses are distinguished. No examples are given and only very few collocations, many of which don't actually exist in modern Arabic. Yet, this is still the dictionary that everyone translating from Arabic into English will use.¹⁹

The latest Arabic–English edition, of 1985, contains around 36,000 entries, 3,000 of them being newly added. Of these entries, a considerable number are, as we have seen, not part of the general lexicon of Modern Standard Arabic. All Wehr editions are ordered according to root, and all are oriented left to right. Rather than providing vowel diacritics, they provide a transliteration in Latin script to illustrate which vowels should be inserted. This indirect way of presenting the Arabic text was chosen for practical reasons, since setting of unvowelled text was less complicated and thus more cost-effective.

Complementary to the Arabic–English edition of Wehr was *The Oxford English–Arabic Dictionary of Current Usage* by Nakdimon Shabbethay Doniach. A remarkable feature of this dictionary, which came out in 1972, is that the Arabic is handwritten. This gives a rather old-fashioned impression. It does not supply grammatical information about the Arabic, which makes its use for encoding difficult.

When Haywood wrote of the 'Elias era' in 1991, the 'Wehr era' was still ongoing, but in 2012, Tim Buckwalter and Dilworth Parkinson suggested that the 'Wehr era' might have ended in the Netherlands in the early 2000s with the publication of a new generation of Arabic–Dutch dictionaries.²⁰ These were compiled after a Dutch–Flemish special committee received a budget

¹⁸ Wehr, *Dictionary*, v.

¹⁹ Arts, 'Making of a large English–Arabic/Arabic–English dictionary', 110.

²⁰ Buckwalter and Parkinson, 'Modern lexicography', 539.

from the governments of the Netherlands and Belgium, and commissioned the compilation of a number of bilingual dictionaries with Dutch, as well as lexicographic tools to create them.²¹ Among them was a set of abridged Dutch–Arabic and Arabic–Dutch learners’ dictionaries compiled by the Belgian Arabist Mark Van Mol (2001), and a set of more exhaustive translators’ dictionaries compiled by a team of Arabists at the Radboud University Nijmegen, co-ordinated by Jan Hoogland (2003).

Buckwalter and Parkinson remarked after the online publication of the Arabic–Dutch dictionary of 2003 that its ‘structural framework and Arabic content’ might be a basis for a future Arabic–English dictionary and others.²² By the time their words were published, Oxford University Press had indeed obtained the rights to use the data for a new Arabic–English dictionary, whose compilation started in 2010. And thus we see history repeat itself, when after the translation of Hans Wehr’s dictionary into English fifty years earlier, the Hoogland Arabic–Dutch dictionary was translated, adapted, and expanded, to appear in 2014 as the *Oxford Arabic Dictionary*, ending the ‘Wehr era’ for Anglophones. It is difficult to establish which of the two translations into English underwent more editing but, having contributed to the *Oxford Arabic Dictionary*, I can state that it has been substantially augmented: I estimate that 85 per cent of it is a translation of its Arabic–Dutch predecessor. The compilation of the Arabic–Dutch dictionary, which started in 1997, made use of a corpus that was not lemmatized and had only limited tools for corpus research, whereas the *Oxford Arabic Dictionary* could benefit from a fully lemmatized corpus and more advanced technology. But the increase in entry count from the 24,682 entries of the Dutch–Arabic dictionary to a new total of 26,316 entries in the *Oxford Arabic Dictionary* has been relatively modest, although it was the result of an enormous effort, illustrating the law of diminishing returns in lexicography.

The *Oxford Arabic Dictionary* is a single-volume dictionary of 1,960 pages, 920 Arabic–English and 1,040 English–Arabic. The two parts were separately compiled: the English–Arabic part was based on the English framework of the publisher, to which Arabic translations were added.²³ Both are, like their Dutch predecessors, intended for use as both an active and a passive dictionary, for encoding and decoding. The ordering of the Arabic–English part is root-based, the columns are oriented left to right, and the Arabic text is fully vowelised; the English–Arabic part is, of course, alphabetically ordered. The

²¹ Tiberius et al., ‘OMBI bilingual lexical resources’.

²² Buckwalter and Parkinson, ‘Modern lexicography’, 553.

²³ See Arts, ‘Making of a large English–Arabic/Arabic–English dictionary’, 113.

online availability of the dictionary (to subscribers) makes it possible to update it regularly; the most recent update at the time of writing, of August 2017, saw 435 new additions to the Arabic–English part and 137 to the English–Arabic, covering recent terms such as *climate refugee* and *Brexit*.

By contrast, the ‘Wehr era’ has not really ended for speakers of German, who continue using the Arabic–German dictionary of Wehr in its fifth edition (1985) and the German–Arabic dictionary compiled by Götz Schregle (1974). Schregle also compiled an Arabic–German dictionary, which appeared between 1981 and 1992; like his predecessor Lane, he did not get past the letter *qāf*. Another German Arabist, Günther Krahel, compiled a German–Arabic dictionary in 1964 and an Arabic–German dictionary, in co-operation with Gharieb Mohamed Gharieb, in 1984, and an Arabic–German dictionary compiled by Lorenz Kropfisch in 1996 and published by Langenscheidt has been followed by pocket-sized German–Arabic and Arabic–German dictionaries compiled or updated by Kropfisch. So although the German-speaking community has an ample choice of Arabic dictionaries, no new dictionary has been published since 1996. I have received information that German students of Arabic have been (and maybe still are) using the Hoogland Arabic–Dutch dictionary, since the German and Dutch languages are so closely related that Germans can use a Dutch dictionary without great difficulty.

Arabic–French and French–Arabic dictionaries have been compiled for passive use, for instance by Daniel Reig (1983), but an active French–Arabic dictionary for speakers of French and an active Arabic–French dictionary for Arabs do not seem to exist. Given the high number of speakers of Arabic in France and the intensive contacts between France and the Arab world, this is remarkable. In present-day Spain there are several dictionaries available in both directions, for instance that of Federico Corriente Córdoba (2005–10), and there is also a series of dictionaries and other materials available for Moroccan Arabic, due to the proximity of Morocco and the high number of Spanish citizens of Moroccan origin.

The dictionaries named above were all produced in Europe with non-Arabs as initiators, but bilingual lexicographic projects have been undertaken in the Arab world too. However, there seems to be a segregation between the lexicographers of Arabic within the Arab world and those outside it. Khoury is very explicit in her observation: ‘Many bilingual dictionaries compiled by Arabs exist, but few equal those of the [European] Orientalists in importance. In the case of bilingual Arabic dictionaries including English, for instance, it is

difficult to find a dictionary comparable to those of Lane and Wehr.²⁴ I have never read any motivation or explanation for this segregation. In some Arab countries the facilities and budgets for research are very limited, but in others there seem to be no material constraints. It is not unlikely that the conservative attitude towards the Arabic language mentioned above plays a role in this reluctance to apply modern techniques to Arabic lexicography.

A number of lexicographers from the Arab world have already been mentioned. Elias was an Egyptian; several French–Arabic dictionaries were compiled by Lebanese lexicographers such as Yussof Mohamed Reda, Jabbour Abdel-Nour, and Souheil Idriss. Also Lebanese was Mounir Baalbaki, who founded a publishing house (Dar al-‘ilm li-l-malayīn) and compiled a series of English–Arabic dictionaries titled *Al-Mawrid*. The *Al-Mawrid* dictionaries exist in abridged and exhaustive editions, in separate volumes or combined, and they are updated on a regular basis. Arabic–English versions have been edited by Mounir Baalbaki’s son Ruhi Baalbaki; the English–Arabic *Al-Mawrid al-Akbar*, left unfinished at Mounir Baalbaki’s death, was finished by his son Ramzi Baalbaki, who has also updated and expanded the original English–Arabic *Al-Mawrid*. The list of *Al-Mawrid* dictionaries is impressive, but there have been criticisms.²⁵ One is that Mounir Baalbaki in his dictionaries introduced new Arabic words coined by himself as translations for English words for which there was not yet any Arabic equivalent. Other lexicographers have done the same.²⁶ Ramzi Baalbaki defended this practice, arguing that ‘compilers of bilingual dictionaries are not only entitled to coin words which may or may not gain currency, but that coinage becomes an essential duty of theirs, especially with a language like Arabic, where a huge number of terms, particularly scientific ones, are lacking’.²⁷ Thus speakers of Arabic searching for an Arabic equivalent for a certain concept can be confronted with a term invented and proposed by a dictionary compiler. And even a difference in source language (for instance, English or French) may result in different Arabic equivalents coined by two different dictionary compilers. According to Radia Benzehra,

today, English–Arabic dictionary editors have to deal with a huge number of lexical gaps that have cumulated over time. The lexical gap-filling process is carried out in a very unsystematic way that is far from creating an atmosphere of cooperation that ultimately contributes to creating unified

²⁴ Khoury, ‘Dictionnaires arabes bilingues’, 48.

²⁵ Arts, ‘Making of a large English–Arabic/Arabic–English dictionary’, 110.

²⁶ Hoogland, ‘Lexical gaps’.

²⁷ Baalbaki, ‘Coinage in modern English–Arabic lexicography’, 68.

English–Arabic lexical databases for lexicographic purposes ... No one dictionary agrees with any other on the treatment of the conceptual mismatches between the two languages.²⁸

This confusion is clearly one of the reasons why speakers of Arabic take recourse to foreign languages when talking about technical or scientific topics, either by putting Arabic aside or by code-switching between the two languages.

Dictionaries of the Arabic Dialects

As mentioned above, dictionaries of the Arabic colloquial languages or dialects do exist and there is ample reason for their existence. The dialects are no longer exclusively used for oral communication, and have for many years attracted the attention of mainly Western linguists, both from a research perspective and for practical communicative needs. This has resulted in the compilation of many dictionaries, of various standards. Probably every Arabic dialect has been codified in a dictionary of some sort. Some of these may cover vast geographic regions like ‘Gulf Arabic’, as in the dictionary compiled by Hamdi Qafisheh. Others may treat the local dialect of a specific group or of inhabitants in a specific city, for instance Cairene Arabic or Damascene Arabic. Hinds and Badawi’s Egyptian Arabic–English dictionary (1986), with 25,000 entries, is a much-praised example of lexicography that is primarily of a dialect, but which covers many Modern Standard Arabic lexical items as well and labels them appropriately.²⁹

A very interesting monolingual dictionary of Moroccan Arabic was published in 2017 by the Zakoura Foundation, a Moroccan NGO striving for the improvement of the education system in the country. To create a vehicle for codification of the Moroccan colloquial, the foundation took the initiative to have this dictionary compiled by a number of academic linguists from Morocco, the principal editors being Khalil Mgharfaoui, Abdellah Chekayri, and Abdelouahed Mabrouir. It bears the title *Qāmūs al-dāriġa al-maġribīya* (‘Dictionary of Moroccan Darija’). The dictionary contains 8,100 entries, and all explanations and definitions are in Moroccan Arabic. It contains a modest number of multi-word expressions and proverbs in addition to a considerable number of loanwords from French, all written in Arabic script, with only very few vowels, which leaves the pronunciation of many words unclear to a user

²⁸ Benzehra, ‘Issues and challenges’, 83, 85.

²⁹ Hoogland, ‘Lexicography: bilingual dictionaries’, 27.

who is unfamiliar with the language. The orthography of Moroccan Arabic as applied in this dictionary is not always in accordance with what can be considered as consensus by practice, but it should be noted that orthographic rules for writing the colloquials do not exist, and so every author may devise their own rules or just improvise: theoretically speaking, there is no correct or incorrect method. The reactions to the publication of this dictionary in Morocco were mixed.

New directions seem possible in the lexicography of the Arabic dialects, given the present rapidly developing situation with regards to their written use, both in print and on the Internet and social media. A kind of mixed dictionary containing more than one variety of Arabic might become more widespread in the future.³⁰

Lexicography and Terminology

The promotion of the Arabic language after the colonial period took place in the shape of a policy of Arabization, replacing the former colonial languages in administration and education with Modern Standard Arabic as a step towards the re-establishment and regaining of an autonomous Arab identity. The process of Arabization is considered to have been most successful in Syria, where it has covered all levels of education, including university education and scientific and technical fields such as engineering and medicine.³¹

We notice in most Arab countries a lack of unified terminology to express modern notions such as the products of modern technology and political and economic phenomena. This lack of terminology has been seen as a question for the language academies which were established in Damascus, Baghdad, and Cairo to revitalize the Arabic language after the period of colonial rule. The major problem with their work in this particular area has been a lack of co-ordination, so that they rarely achieve consensus among themselves concerning the choice of a new word. For example, 'television' is *telfāz* in Lebanon, while it is *tilivizyōn* in Egypt; the official term for 'radio' is *mawwāj* in Lebanon, while it is *midhyā* in Egypt; in Lebanon, a computer is officially called *hāsūb*, while the official term in Egypt is 'aql 'iliktrūnī (meaning 'electronic brain'); the loanword *kumbyūtar* is also used. Finally, the impact of the academies' decisions is slow, and in many instances their recommendations remain theoretical.³²

³⁰ Hoogland, 'Lexicography: bilingual dictionaries', 27. ³¹ Shaaban, 'Language policies'.

³² Ibrahim, 'Lexical variation'.

Arabic has for many years applied a number of mechanisms for coining new words. The root–pattern system is in principle very suitable for this purpose. Both words for ‘radio’ above are examples: *mawwāj* is coined according to the pattern *fa‘āl* (meaning an intensified noun of the root) and the root *m-w-j* ‘wave’. The word *midhyā* is coined with the pattern *mif‘āl* ‘instrument’ and the root *dh-y-* ‘to emit, transmit’. Both roots and both patterns seem suitable for coining a word for ‘radio’. However, a quick corpus consultation shows the word *radiyū* is used much more in the Arab world, which perfectly illustrates the discrepancy between theory and practice.³³ In addition to the option of using an existing root and a current pattern, other options are available: the use of other patterns, the use of rare patterns, the creation of new patterns, the creation of new roots.

The language academies have created and distributed terminology lists and even dictionaries, as in the case of the *Mu‘jam al-wasīf* edited by Muṣṭafā. In addition to the local language academies there is a pan-Arabic organization called the Arab League Educational, Culture, and Sciences Organization (ALECSO) which has its head office in Tunis and an ‘Arabization Coordination Bureau’ in Rabat, Morocco. This office has published an enormous number of ‘Unified Dictionaries’ of terminology related to a number of disciplines, like the *Unified Dictionary of Human Sciences Terms* and the *Unified Dictionary of Geographical Terms*. These dictionaries contain many thousands of discipline-related terms in three languages: English, Arabic, and French. The English–Arabic part is the main part of the dictionaries, with indexes of French and Arabic terms referring to the numbers of the terms in the main part. In recent years all these terms have been stored in a freely accessible database called ArabTerm.³⁴

Compilers of specialized dictionaries which contain the terminology of specific domains or disciplines such as medicine or law also suffer from the lack of co-ordination and communication between the language academies, specialists in the fields, and lexicographers. Notwithstanding this difficult situation, remarkable dictionaries have been compiled, like Harith Suleiman Faruqi’s English–Arabic and Arabic–English law dictionaries (1962 and 1972), Ahmad al-Khatib’s English–Arabic *New Dictionary of Scientific and Technical Terms* (1971), and Yusuf Hitti and Ahmad al-Khatib’s *Hitti’s New*

³³ The word ‘*rādiyū*’ has 12,447 hits, ‘*midhyā*’ has 275 hits, and ‘*mawwāj*’ has only thirteen hits in the Arabic Gigaword Corpus.

³⁴ A list of the dictionaries, and the ArabTerm database, are accessible on the ALECSO website.

Medical Dictionary: English–Arabic (1989). Recently coined terminology in different fields can be found in the ArabTerm database.

Special Types of Dictionaries

A few special dictionaries are worth mentioning. The *Al-Mawrid Al-Mar'iyy* of Ruhi Baalbaki is a remarkable multilingual dictionary, consisting completely of detailed illustrations of all possible fields of daily life, technology, and so on, in which everything which can be named is supplied with an Arabic term, as well as English, French, and Spanish equivalents. In the field of education all kinds of dictionaries exist. Several of the above-mentioned dictionaries have special abridged student editions. A special case is the learners' frequency dictionary of Buckwalter and Parkinson, *A Frequency Dictionary of Arabic: Core Vocabulary for Learners*. In this dictionary the authors present the 5,000 most frequent words of Arabic, with one full sentence containing each word, followed by an English translation. For educational use, for some words the presence of only one example sentence is too limited, but all data in this dictionary are fully corpus-based, and the corpus, ArabiCorpus, is accessible to those who want to consult more example sentences.

Innovation

The development of electronic corpora of Arabic, databases such as ArabTerm, and dictionaries such as the online version of the *Oxford Arabic Dictionary* is part of a story of innovation which goes back to the 1990s, when dictionaries on CD-ROM became available, a pioneer being *al-Qāmūs*, produced by the Sakhr Company in Egypt. When the Internet became capable of dealing with Arabic (for instance by connecting letters correctly, and presenting text from right to left), various dictionary sites appeared, some of which have already disappeared. But some are still available, for example the website ejtaal.net, which presents an impressive number of more than thirty dictionaries. The data presented are images of pages of printed dictionaries (among them those of Lane, Wehr, Hinds and Badawi, and al-Bustānī) and the user will, after entering a lookup root, find scanned pages of all these dictionaries on the result page. So the outcome of the search action is still on the basis of printed dictionaries. This site can be compared to a shelf full of dictionaries that all open on the correct page after entering a lookup root.

The online *Oxford Arabic Dictionary* offers innovations such as the morphological analyser, which analyses any given string of Arabic characters into

its potential component lemmata and takes the user to the most likely headword page, as well as presenting them with the list of other possibilities. Even in print format, this dictionary is the product of corpus-driven research, using tools like Sketch Engine, which creates ‘word sketches’ or corpus-driven accounts of words’ grammatical and collocational behaviour.³⁵

Conclusion

This chapter can be summarized as follows. First, the long tradition of (monolingual) Arabic lexicography has been a limiting factor in embracing modern ideas and practices in lexicography. Secondly, most innovation comes from Western lexicographers of Arabic, for example the corpus-based approach in dictionary-making. Thirdly, new technology has resulted in new working methods in the compilation of dictionaries and the presentation of their content. This development will probably continue to influence the discipline in the near future.

³⁵ Arts and McNeil, ‘Corpus-based lexicography’.

Modern Hebrew

TSVI SADAN

The present chapter chronicles a mainly external history of the lexicography of Modern Hebrew.¹ It focuses on the most important general monolingual, bi- and trilingual, learners', and online dictionaries whose source language is this latest phase of the Hebrew language; rather than surveying all the bilingual and trilingual dictionaries with Hebrew as a source language, it discusses Hebrew–Russian and Hebrew–English dictionaries because of the historical and practical importance of these two target languages for Modern Hebrew. In accordance with, for example, the judgement of Uzzi Ornan, the beginning of Modern Hebrew is set in the 1880s, when Hebrew marked a true turning point sociolinguistically – hence, also linguistically – by starting to be used as a spoken language and to fulfil other functions of a modern society in Palestine and, later, in the State of Israel.² The preceding period, from the end of the eighteenth century, when Modern Hebrew literature started, until around 1880, is not, therefore, treated here. The lexicography of the earlier periods of Hebrew in the modern era is also excluded. The history of Modern Hebrew lexicography in the sense of the lexicography of Modern Hebrew is, therefore, almost equal in length to the history of the Modern Hebrew language itself.

This history of the lexicography of Modern Hebrew follows the commonly accepted periodization of Modern Hebrew: 1881–1918, 1918–48, and 1948–present.³ I divide the third period further as follows: 1948–70, 1970–2000,

¹ Other histories of (Modern) Hebrew lexicography include Segert and Sabar, 'Hebrew and Aramaic lexicography'; Goshen-Gottstein, 'Lexicography of Hebrew'; Cohen and Choueka, 'Hebrew lexicography'; Glinert, 'Modern Hebrew lexicography'; Brisman, *History and Guide to Judaic Dictionaries and Concordances*, 81–102; and Merkin, 'Lexicography: modern period'. See also Schwarzwald, 'Modern Hebrew dictionaries' and Sadan, 'Lexicography of Hebrew', for analyses of the (micro)structure of many of the dictionaries mentioned here.

² Ornan, 'Matay nolda haivrit hakhadasha'; for detailed analyses of this so-called revival of Hebrew, see Harshav, *Language in Time of Revolution*, and Morag, 'Emergence of Modern Hebrew'.

³ See, for example, Eytan et al., 'Hebrew language: modern period', 671.

and 2000–present.⁴ These five periods are characterized by the following social and cultural changes also affecting the language (and hence dictionaries, too, though with some delay). The first, 1881–1918, extends from the beginning of attempts to use Hebrew as a spoken language and for other societal functions in Palestine until the end of the First World War. The second, 1918–48, extends until the end of the British Mandate for Palestine, in which Hebrew was one of the three official languages. The third, 1948–70, extends from the independence of the State of Israel until the end of the subjugation of Hebrew to normativism, and the fourth, 1970–2000, from the beginning of the liberation of Hebrew from normativism to the advent of the Internet among the general public in Israel; the fifth period, 2000–present, is the digital age.

About twenty dictionaries or dictionary traditions are presented in the following four sections, with regard to their external history and, to a lesser extent, to internal features such as their respective lexicographic innovation, users, scientific versus practical use, normative versus descriptive nature, and lemmatization.⁵

Monolingual Dictionaries

Milon halashon haivrit hayeshana vehakhadasha (it also had titles in Latin, German, French, and English, this last being *A Complete Dictionary of Ancient and Modern Hebrew*) is the first landmark in Modern Hebrew lexicography.⁶ It was the work of Eliezer Ben-Yehuda, often cited as one of the central figures (or the central figure in the popular myth) of the revival of Hebrew.⁷ It was the first dictionary of Hebrew that included those words coined after the language started to fulfil societal functions that had been dormant for about seventeen centuries, until 1881, when Ben-Yehuda himself immigrated to Palestine and started using Hebrew as a spoken language. It was also the first historical dictionary of Hebrew to cover not only words newly coined by Ben-Yehuda himself and others through 1902, but also words from the first three periods of the Hebrew language, that is, Biblical Hebrew,

⁴ Here I follow Schwarzwald (personal communication); the lexicography-specific periodization by Merkin, 'Lexicography: modern period', 523, captures only the most important turning points in the history of *monolingual* dictionaries, however important these points might be.

⁵ See Reshef, 'Modern Hebrew grammar', on the changing normativism of researchers in Modern Hebrew.

⁶ See Merkin, 'Prakim betoldot hamilonut haivrit hakhadasha', for further details.

⁷ Fellman, *Revival of a Classical Tongue*, meticulously analyses this myth in seven areas.

Rabbinic Hebrew, and Medieval Hebrew. It is basically a monolingual dictionary, but the main senses are also translated into their equivalents in English, German, French, and Russian.

As Ben-Yehuda wrote in the introduction to his dictionary, published posthumously as a separate volume in 1940, the idea of compiling a dictionary of Hebrew occurred to him as a means of filling the lack of words for modern life, which in turn was to serve his ultimate goal of returning to the language of the forefathers as one of the two prerequisites for the Jews becoming a full-fledged nation (the other being the return to the land of their forefathers).

Ben-Yehuda's original idea was to compile a practical dictionary to help anyone wanting to speak Hebrew as a daily means of communication. As he continued to work on this self-imposed task, his plan shifted to the compilation of a historical dictionary of Hebrew. Lacking not only theoretical background in linguistics and philology but also practical knowledge of lexicography and relevant Semitic languages other than Aramaic and Arabic (for the etymology section of each headword), Ben-Yehuda did not seem fit to assume such a gigantic lexicographic project that might have been assumed by a language academy or some similar group of specialists. He was himself aware of his lack of professional qualification, but his zeal for reviving Hebrew in general and compiling a dictionary for this old new language in particular made him continue this arduous task almost single-handedly until his death in 1922.

After years of preparation, the first volume of the dictionary was published in 1908, in the middle of the first period of Modern Hebrew, by Langenscheidt in Berlin, through donations. Only the first five volumes saw the light of the day in Ben-Yehuda's lifetime, and the next three volumes, whose material he had prepared, were published after his death. The remaining eight volumes were prepared by Moshe Segal and Naftali Tur-Sinai (Torczyner), two of the most important researchers of the Hebrew language at that time – I shall return below to Segal as a bilingual lexicographer – and published by Ben-Yehuda Publications in memory of Ben-Yehuda in Jerusalem. This lexicographic project was completed with the publication of the last volume in 1959, spanning half a century since the publication of the first volume.

Although both Ben-Yehuda's contemporaries and later experts criticized this historical dictionary for its lexicographic methodology and treatment, it is a monumental work, especially considering that it was initiated by a single person, and will remain the only historical dictionary of Hebrew until *The Historical Dictionary of the Hebrew Language*, which is also supposed to cover

Modern Hebrew until one year before the establishment of the State of Israel in 1948, is published by the Academy of the Hebrew Language (see below).

With this shift in the nature of the dictionary away from being a practical dictionary, it could not meet the immediate needs of those who wanted to speak Hebrew but lacked vocabulary, for whom Ben-Yehuda originally intended his dictionary, especially in the formative and, thus, most critical period in the revival of Hebrew. It seems, however, to have been successful in making many sceptics about the sociolinguistic and linguistic enterprise to revive Hebrew consider Hebrew a language, since it now had a dictionary, which is one of the most important symbols of the *Ausbau* of a newly developed language.

One of the most remarkable lexicographic innovations of this historical dictionary by Ben-Yehuda was to arrange most of its headwords alphabetically, and not, as in all previous dictionaries of Hebrew, by their roots (if known), although verbs were still arranged this way. The dictionary is also characterized by its purism: Ben-Yehuda decided to exclude words of foreign origin, including not only Persian, Greek, and Latin but even Aramaic ones, which had become part and parcel of the Hebrew language centuries before.

It was not until 1919, around the beginning of the second period of Modern Hebrew, that the lexicographic lacuna left by Ben-Yehuda's decision not to make his dictionary a practical guide was filled, when *Hamilton haivri* ('The Hebrew dictionary') was published by Yehuda Grazovski and David Yellin. This was the first practical monolingual dictionary of (Modern) Hebrew, covering not only Modern Hebrew, which was being formed and was expanding at that time, but also Biblical Hebrew, Rabbinic Hebrew, and Medieval Hebrew; that is, it was historical as well. Grazovski continued to elaborate and update this dictionary and published the expanded edition as *Milon hasafa haivrit* ('Dictionary of the Hebrew language') between 1934 and 1936 in three volumes, significantly increasing the size and quality. He published a further updated unabridged edition as *Milon ivri* ('Hebrew dictionary') in 1946, now under his new, Hebraized name, Gur. He followed the new method of lemmatization initiated by Ben-Yehuda, but was less puristic than the latter in the selection of headwords for his three dictionaries. They were the most popular dictionaries in the second period of Modern Hebrew.

A new era in the lexicography of Modern Hebrew, which also coincides more or less with the second period of Modern Hebrew, began with the publication of *Milon khadash* ('New dictionary') by Avraham Even-Shoshan in four volumes between 1947, one year before the establishment of the State of

Israel, and 1952.⁸ It was updated, expanded, and published as *Hamilon hekhdash* ('The new dictionary') in four volumes between 1966 and 1970. Although these two dictionaries by Even-Shoshan followed roughly the same format and principles adopted by Gur, they, especially the later edition, gained popularity and dominated (monolingual) Modern Hebrew lexicography for almost five decades, so as to make the name of their author, Even-Shoshan, almost synonymous with *the* monolingual Hebrew dictionary. After Even-Shoshan passed away in 1984, Moshe Azar updated and published *Hamilon hekhdash* as a new dictionary in 2003, calling this new edition simply *Milon Even-Shoshan* ('Even-Shoshan dictionary'), thus officializing the commonly used name for Even-Shoshan's two original dictionaries.⁹

What made *Milon khadash*, then *Hamilon hekhdash*, so popular seems to have been the lack of rivals. Gur presented the language as it was in the process of formation, and characterized his two dictionaries as 'practical', by which he meant that they were meant to guide readers to an understanding of and compliance with the emerging norms. By contrast, the Hebrew which Even-Shoshan depicted in his two dictionaries was already mature enough to assume the important role as the main official language of a new independent state, and he could afford to treat it much more descriptively than Gur. However, his dictionaries were not totally descriptive. Indeed, normativism was still a sufficiently strong sociolinguistic force in the Hebrew speech community in the State of Israel in the third period of Modern Hebrew for it to be felt that the language needed 'defenders'.

The layout and principles of the three editions of Even-Shoshan's lexicographic legacy have remained about the same for nearly half a century, including the same method of lemmatization as that of Ben-Yehuda and Gur (Grazovski). The main difference between the three is the inclusion of more words and senses, the additions being on the whole newer and more colloquial, showing a decreasing normative tendency.

Otsar halashon haivrit litkufoteyha hashonot ('Thesaurus of the Hebrew language according to its historical periods') by Yaakov Knaani (Kaufmann), published between 1960 and 1989 in eighteen volumes, is a historical dictionary that is even larger in scope than Ben-Yehuda's dictionary and covers Modern Hebrew more extensively, but it is far less well known and used than his contemporary Even-Shoshan's dictionaries, not only by the general public but also by researchers. This gigantic dictionary

⁸ For Even-Shoshan's work, see Glinert, 'Leshitot hamilonaut bemilon Even-Shoshan'.

⁹ See Azar, 'Milon Even-Shoshan, revisited'.

distinguishes itself from the others in its exhaustive coverage of many words in attested Hebrew texts of all historical periods, including even nonce formations by writers (as well as by Knaani himself).

The stagnation of monolingual Modern Hebrew lexicography, which had been dominated by Even-Shoshan's dictionaries for almost five decades, for better or worse, started to dissipate with the publication of a series of new monolingual dictionaries based on different principles in 1995 and onwards, that is, from the end of the fourth period of Modern Hebrew through the beginning of its fifth (present) period.

The first of them is *Milon hahove* ('Dictionary of contemporary Hebrew') by Shoshana Bahat and Mordechai Mishor, published in 1995. Its original Hebrew title (literally 'Dictionary of the present') speaks for itself as the most conspicuous innovation of this pocket dictionary. Unlike its predecessors, it uses the (masculine singular) present as the citation form of each verb for lemmatization instead of the third-person masculine singular past. This innovation caused controversy among linguists; only time will tell whether it will be accepted as a new lexicographic custom or not.¹⁰ Bahat and Mishor identified their dictionary as following the normative decisions of the Academy of the Hebrew Language, where they worked, but, paradoxically, the dictionary also turns out to be fairly descriptive in that it describes the norms, pointing out, with cross-references, which forms and senses are normative and which are not. This dictionary is also innovative because it is one of the few *synchronic* (and not historical) monolingual dictionaries of Modern Hebrew.

A true lexicographic innovation occurred one year later, in 1996, with the publication of *Rav-Milim: hamilon hashalem laivrit hakhadasha* ('Rav-Milim: a comprehensive dictionary of modern Hebrew'), edited by Yaacov Choueka.¹¹ This dictionary remains the only comprehensive synchronic monolingual dictionary of Modern Hebrew. It was the first dictionary of Modern Hebrew to be compiled with a computer, if not fully with a corpus, and to be also available electronically, first as a CD-ROM and later online (see below). It was also the first monolingual dictionary of Modern Hebrew whose examples were invented (or sometimes adapted) rather than citations from literary sources. It is therefore regrettable and does not stand to reason

¹⁰ See Merkin, 'Al milon ivri shimushi bikhlal veal "Milon hahove" bifrat', and Mishor, 'Milon ha-Hoveh and Milon Sapir', for further details; the former is critical of the new method of lemmatization.

¹¹ See, for example, Freidkin and Freidkin, 'Rav-Milim, Hamilon hashalem: Milon ivri khadash', and Choueka, 'Rav-Milim', for further details.

that this important lexicographic work disappeared from bookstores and became unavailable in its original print form only a few years after its publication.

Milon sapir: *Milon ivri-ivri entsiklopedi beshitat hahove* ('The encyclopedic sapphire dictionary') by Eitan Avneyon, published in 1998, followed the method initiated by *Milon hahove*, mentioned above, to lemmatize verbs by their (masculine singular) present forms.¹² Unlike the latter, however, this dictionary is historical rather than synchronic. It was the first dictionary of Modern Hebrew that included encyclopedic entries, and the growing tendency is now to include them, especially in comprehensive monolingual dictionaries. Its treatment of encyclopedic entries was, however, inconsistent: only place-names were included, while personal names were excluded for the strange reason that the approach to include personal names as well seemed too far-reaching (though place- and personal names are on a par as encyclopedic entries). A second edition appeared in 2002, in which Avneyon updated some of the main entries on the basis of comments he received from readers of the first edition, as well as updating encyclopedic entries. The latest addition to the list of monolingual dictionaries of Modern Hebrew is *Milon Ariel hamakif* ('The extended dictionary Ariel') by Maya Fruchtmann and Daniel Sivan, published in 2007.¹³ This comprehensive dictionary is historical and encyclopedic as well. Its treatment of encyclopedic entries is also inconsistent, like *Milon sapir*, in that only place-names were included, but, unlike the latter, it does not explain why personal names were excluded.

Although a history of lexicography should be restricted to only those dictionaries that have already been published, at least partially, special mention must be made of *The Historical Dictionary of the Hebrew Language* which is to be published by the Academy of the Hebrew Language.¹⁴ Already very important in Modern Hebrew lexicography, it plans to cover Modern Hebrew until one year before the establishment of the State of Israel in 1948. When completed, it should supersede Ben-Yehuda's historical dictionary. The Academy of the Hebrew Language conceived of the idea of preparing a historical dictionary of the Hebrew language based on state-of-the-art methods and principles of modern lexicography as one of its main activities

¹² See Mishor, '*Milon ha-hoveh* and *Milon sapir*', for further details.

¹³ See Merkin, '*Al milon ivri moderni bikhlaal veal Milon Ariel hamakif bifrat*', for further details.

¹⁴ See Rubinstein, '*Historical Dictionary of the Hebrew Language*', and Tal, '*Historical Dictionary of the Hebrew Language* of the Academy of the Hebrew Language', for further details.

after only one year of its establishment in 1953 as the official supreme authority of the Hebrew language in the newly independent Jewish state. This national lexicographic project consists of the following phases: first, selecting primary sources spanning almost three millennia; secondly, selecting the most reliable textual editions of these sources and turning them into an annotated electronic corpus; and, thirdly, writing entries. Although the project was officially launched in 1954, it was not until 2006, that is, after five decades, that the third phase started at long last. The first phase was completed with the publication of *Sefer hamekorot* ('Book of sources') in 1963. In the second phase of the project an electronic corpus containing these primary sources, called *Ma'agarim*, was released first as a CD-ROM in 1998 and online in 2005.¹⁵ This is also the only corpus-based dictionary to date that (also) covers Modern Hebrew through 1947.

Bilingual (and Trilingual) Dictionaries

As shown in the previous section on monolingual dictionaries, Eliezer Ben-Yehuda changed his original plan to publish a practical dictionary of Modern Hebrew meant for the general public speaking the language, and it was not until 1919, almost three decades after the beginning of the revival of Hebrew, that Yehuda Grazovski, together with David Yellin, published the first dictionary which covered the vocabulary of Modern Hebrew as well as that of earlier stages of the language. The existence of a practical dictionary was crucial in the first formative decades, or the first period, of the language. Since there was no monolingual dictionary, the need for some sort of dictionary of Modern Hebrew between the 1880s and 1919 was temporarily filled by none other than Grazovski (together with Joseph Klausner) and Ben-Yehuda themselves, with their trilingual practical dictionaries titled *Milon shel kis meivrit lerusit veashkenazit umerusit leivrit veashkenazit, khelek rishon: ivri-rusi-ashkenazi* ('Pocket dictionary from Hebrew to Russian and German, part one: Hebrew–Russian–German') and *Milon ivri meturgam ashkenazit hameduberet beyn hayehudim verusit* ('Hebrew dictionary translated into the German spoken by the Jews and into Russian'), published in 1900 and 1901 in two major centres of Jewish culture in the diaspora at that time, Warsaw and Vilnius, respectively.

¹⁵ See Sadan, 'Ma'agarim', for analysis of the online version of *Ma'agarim* as both a historical corpus (in comparison with other national corpora) and a corpus query system.

Grazovski and Ben-Yehuda chose Russian and German or Yiddish as the two target languages for their pocket trilingual practical dictionaries for an apparently pragmatic reason: these three non-Hebrew languages were among the most widely used ones by the Jews in Palestine at the turn of the twentieth century, and they were the languages which the authors themselves knew best. These less well-known dictionaries were also the first to cover Modern Hebrew (as well as Biblical, Rabbinic, and Medieval Hebrew), preceding their more comprehensive monolingual counterparts, and represented Modern Hebrew in its first historical period.

Milon ivri–angli shalem / Compendious Hebrew–English Dictionary by Reuben Grossman (with contributions by Moshe Segal) and *Milon ivri–angli shimushi / A Concise Hebrew–English Dictionary* by Moshe Segal, both published in 1938, are the two most important Hebrew–English dictionaries of the second period of the language and meant mainly for receptive use by speakers of English. Although they differed in that the former covered not only Modern Hebrew but also Biblical, Rabbinic, and Medieval Hebrew, their principles were identical: they reflected more maturity of Modern Hebrew, but the language they represented was normative.

The third period of the language witnessed two important synchronic Hebrew–English dictionaries intended, unlike their predecessors, both for reception by speakers of English and for production by speakers of Hebrew. The first is *The Complete Hebrew–English Dictionary* by Reuben Alcalay, published in 1963. Its extensive coverage of Modern Hebrew vocabulary testified that Hebrew had become a full-fledged modern language capable of assuming all the societal functions of a modern society. Although more than five decades have passed since its publication, it remains the most comprehensive Hebrew–English dictionary, which does not necessarily mean that it was impeccable. This dictionary was less normative than its predecessors but, on the other hand, it was too generous in registering neologisms coined by the compiler himself, many of which were not accepted by the general public. The ‘new enlarged’ edition, published in 1990, is nothing but a combination of the original edition and two lists of neologisms appended to the end. The second is *Milon Megido hekhadish / The Megiddo Modern Dictionary Hebrew–English*, by Reuben Sivan and Edward Levenston, published two years later, in 1965. It is less comprehensive than Alcalay’s dictionary but more comprehensive than many of the later bilingual dictionaries, including those mentioned below. This dictionary is also recognized as the first among all dictionaries of Modern Hebrew then published, whether bilingual or monolingual, that arranged verbs alphabetically and not by root.

The fourth period of the language saw a plethora of synchronic Hebrew–English dictionaries, but most were pocket-sized, with only headwords in Hebrew and their equivalents in English. Three of the more comprehensive are *Milon ivri–angli/Comprehensive Modern Hebrew–English Dictionary* by Martin Wittenberg, *Milon ivri–angli/Hebrew–English Dictionary* by Menachem Dagut, and *Milon ivri–amerikani–angli–ivri lishenot haalpayim/User-Friendly Dictionary, Hebrew–English* by Shimson Inbal, published in 1977, 1986, and 1988, respectively. Although they incorporated more colloquialisms prevalent at the time of publication, all of them, as well as the two published in the third period of the language, need serious updating, since both the language and society have changed tremendously since their publication. Wittenberg characterized his lexicographic work as a learners’ dictionary but, to all intents and purposes, it is not. As a general-purpose bilingual dictionary, however, this Hebrew–English dictionary is second only to Alcalay’s in its comprehensiveness but, regrettably, is far less known and less obtainable. Dagut’s Hebrew–English dictionary is an update of Segal’s after almost three decades, with abundant colloquialisms. Unfortunately and strangely, no Hebrew–English dictionary worthy of mention here was published in the fifth period of the language. The bilingual lexicography of Modern Hebrew, therefore, is far behind the times, with all the important advances in bilingual lexicography remaining unincorporated into bilingual dictionaries of Modern Hebrew and unknown to their users. It even lags in comparison with the monolingual lexicography of Modern Hebrew, which has seen significant advances since the 1990s.

Russian was the common target language in the first two non-monolingual dictionaries of Modern Hebrew mentioned above. Hebrew–Russian dictionaries published in later periods were different in nature from these earlier dictionaries in that many of their actual and potential users were also detached from Judaism and Jewish culture. *Ivrit–russkij slovar’* (‘Hebrew–Russian dictionary’), by Feliks Shapiro, published posthumously in 1963, is unique not only among Hebrew–Russian dictionaries but among bilingual dictionaries of Modern Hebrew in general. Its compiler had to work through various forms of adversity, political, technical (especially typographical), and lexicographical. It was prepared in the former Soviet Union when any direct contact with Israelis, thus with Modern Hebrew, was forbidden.¹⁶ Not only could an appropriate Hebrew font not be found, but Shapiro had to rely on

¹⁶ See Greenbaum, ‘Status of Hebrew’, on the situation of Hebrew in the former Soviet Union.

rather outdated dictionaries of Modern Hebrew as his data sources and the number of copies published was severely curtailed. In spite of all these difficulties, which the compilers of many other dictionaries of Modern Hebrew were spared, this was more than a dictionary for many of its Jewish users. It also served as a source of knowledge about Judaism and national pride for those Jewish users who knew little or nothing about the national culture of their forefathers.¹⁷

Israel experienced two significant waves of immigration from the former Soviet Union: the first in the 1970s and the second from the late 1980s to the early 2000s. Each wave inspired the publication of a new Hebrew–Russian dictionary intended for these Russian-speaking immigrants: *Milon ivri–rusi/Evrejsko (ivrit)–russkij slovar’* (‘Hebrew–Russian dictionary’) by Michael Dror and *Ivrit–russkij slovar’* (‘Hebrew–Russian dictionary’) by Baruch Podolsky, published in 1975 and 1993 respectively. The latter was vastly expanded by the same compiler and published in 2007 as *Ivrus: novejsij ivrit–russkij slovar’* (‘Ivrus: newest Hebrew–Russian dictionary’). Unfortunately, modern Hebrew–Russian dictionaries also suffer from the same problems as their Hebrew–English counterparts, or even more serious ones, with the micro-structure of the former being far less elaborate, though they are meant first and foremost as unidirectional dictionaries, unlike the latter for speakers of the target language.

It is true that Israel is the homeland and centre of Modern Hebrew and has been absorbing immigrants from many countries, but this does not mean that it need be the only country in the world that produces monolingual dictionaries of Modern Hebrew or bilingual dictionaries with Hebrew as the source language. The fact remains, however, that other countries, including the United States, have not produced any important dictionaries of either kind.

Learners’ Dictionaries

Learners’ dictionaries are a rather new and still underdeveloped type of dictionary in the lexicography of Modern Hebrew. Modern Hebrew has been taught as a foreign language both publicly and privately in so-called *ulpanim*, intensive programmes for learning Hebrew through the medium of Hebrew. These have been provided for speakers of other languages,

¹⁷ In *Jazyk zaprashennogo jazyka* (‘Dictionary of a forbidden language’), Leah Prestina-Shapiro, the compiler’s daughter, collects often touching testimonies in Russian by people who used this historic Hebrew–Russian dictionary when they were still in the former Soviet Union (many of them are now in Israel).

including both new immigrants to Israel and foreign students in Israeli universities, to facilitate their integration into Israeli society and universities respectively, especially since the establishment of the State of Israel. If we consider the putatively large annual and cumulative numbers of such immigrants and students relative to the population of Israel and Israeli universities, a rather high demand for learners' dictionaries of Modern Hebrew must always have existed.

However, it was only in 1993 that the first learners' dictionary of Modern Hebrew, *Rav-Milon: milon didakti du-leshoni ivri-ivri-angli angli-ivri* ('Multi dictionary: bilingual learners' dictionary Hebrew-Hebrew-English/English-Hebrew'), was published by Liora Weinbach and Edna Lauden, two teachers at the Hebrew *ulpan* at Tel Aviv University, as a Hebrew-Hebrew-English semi-bilingual dictionary, filling this lacuna for the first time. This dictionary also has Hebrew-Hebrew-Russian-English and Hebrew-Hebrew-French semi-bilingual versions. The reason for the choice of Russian and French must be that they are two of the most important target languages, as reflected in the number of their respective speakers among new immigrants and foreign university students. This dictionary, in its three language versions, remains the only learners' dictionary of Modern Hebrew. There may also be enough demand for Spanish and probably also Amharic versions, since they are among the most frequently used languages after English, Russian, and French in government offices in charge of new immigrants. It is hoped that this unique learners' dictionary of Modern Hebrew will be updated and improved by incorporating recent advances in both theoretical and practical pedagogical lexicography.¹⁸

Online Dictionaries

Although a small number of two other types of electronic dictionaries – handheld hardware devices and software programmes – for Modern Hebrew already existed, only at the beginning of the twenty-first century did Modern Hebrew lexicography fully enter the digital age. This new era was heralded by the commercial launch of Yaacov Choueka's *Rav-Milim* online in 2002 as the first online version of one of the print dictionaries of Modern Hebrew mentioned above. Since then, several freely available monolingual and bilingual online dictionaries of Modern Hebrew have appeared, some

¹⁸ See Sadan, 'Tasks of the pedagogical lexicography of Modern Hebrew', for a detailed analysis of this dictionary, especially its microstructure, including the problem of following sense differentiation in two separate languages in a semi-bilingual dictionary.

converted into this new electronic medium from their original print version and others prepared electronically from scratch.

The pioneer online version of *Rav-Milim* is based on and replaces the CD-ROM version (for Hebrew Windows 95) of *Rav-Milim*, which was bundled with the print original published in 1996. Of all the new features added to the CD-ROM version and, later, to the online version of this dictionary, the morphological analysis was – and still is – revolutionary. Unlike all the other print and online dictionaries – except for *Morfis*, presented below, which is the free stripped-down version – the online version of *Rav-Milim* can accept word forms as search terms and then analyse them into lexemes. Word forms in Hebrew can include multiple lexemes, which in turn often appear in their inflected forms. Another important addition to the online version is the Hebrew–English part, making the dictionary semi-bilingual, with the concomitant problem of sense differentiation briefly mentioned above.¹⁹

Milon sapir mekuvan ('Sapphire dictionary online') is the free online version of another monolingual dictionary of Modern Hebrew, Eitan Avneyon's *Milon sapir*. Unlike *Rav-Milim*, this online dictionary has no built-in morphological analyser, so its users must split word forms into lexemes (or their inflected forms) as search terms, thus not making full use of the potential of the electronic medium. *Morfis* is a free Hebrew–English (and English–Hebrew) dictionary with no print counterpart. Since it is based on the Hebrew–English portion of the online *Rav-Milim*, it can also analyse word forms into lexemes, but sense definitions in simple glosses with no further sense disambiguation through labels or other means are the only kind of information. In spite of its minimalist nature, this online-only dictionary is quite widely used, seemingly because it is free and bilingual.

Conclusion

This historical chronicle has surveyed about twenty important dictionaries or dictionary traditions, most of which happen to have been published in Israel, in four categories: monolingual, bi- and trilingual, learners', and online dictionaries. The history of Modern Hebrew lexicography almost coincides in length with the history of the Modern Hebrew language. Dictionaries are meant to describe languages as they are used, but when it comes to Modern Hebrew, those dictionaries compiled in the formative years of Modern Hebrew, or in its first periods, from 1881 to 1917, also served to expand the

¹⁹ See Neeman and Finkel, 'Rav-Milim online'.

language lexically and partially helped raise its prestige, even in the eyes of its sceptics.

We discern two common tendencies in monolingual and bilingual dictionaries. The first ones, whether monolingual or bilingual, were historical and normative in that they also included not only Modern Hebrew but also the first three phases of the Hebrew language and excluded colloquial and sometimes foreign words and expressions. These two categories of dictionaries of Modern Hebrew, however, started to diverge from each other as the language approached its fifth period, or the digital age, around 2000. At least some of the monolingual dictionaries published around 2000 incorporated a portion of the advances made in modern lexicography. As for bilingual dictionaries of Modern Hebrew, none worthy of mention here have appeared since the end of the 1980s.

A welcome addition to Modern Hebrew lexicography is that of the only learners' dictionary of Modern Hebrew so far. Considering the number of foreign learners of Modern Hebrew and thus potential users of learners' dictionaries, this category of dictionary of Modern Hebrew can and should be developed more.

The so-called corpus revolution in lexicography that started in the 1980s has neither reached nor affected Modern Hebrew lexicography yet. The same seems to be the case with electronic lexicography, which has made significant advances in the past few decades with the exponential development of computer and Internet technology. Few dictionaries of Modern Hebrew, with the exception of *Rav-Milim* in its print and online versions (as well as the still unpublished *Historical Dictionary of the Hebrew Language*), seem to have benefited from the fruits of state-of-the-art electronic lexicography.

The Slavic and Baltic Languages

RICK DERKSEN

While the exact nature of the relationship between the Slavic and Baltic languages is a topic which cannot be addressed here, their genetic proximity and interwoven history justifies the decision to include them in one and the same chapter. I shall start with the lexicography of the Slavic languages, which is a more complex subject than the lexicography of Baltic.

Slavic

The earliest written documents in Slavic are connected with the Moravian mission of Constantine, who shortly before his death assumed the name Cyril, and Methodius. In 862/3, the two brothers from Thessalonica were selected to travel to Moravia after Duke Rostislav had requested the Byzantine emperor Michael III to send him Slavic-speaking missionaries. Apparently, Constantine and Methodius were very well acquainted with the local Slavic vernacular, which has features in common with present-day Bulgarian and Macedonian. At that time the diversity within the Slavic language area was relatively small, which meant that the liturgical language employed by the missionaries could be readily understood in regions such as Moravia. Ultimately it evolved into the *lingua sacra* and literary language Church Slavic. The documents that for linguistic and paleographic reasons are believed to belong to the oldest layer are said to be part of the Old Church Slavic canon.

The activities of Constantine and Methodius involved the creation of an alphabet suited to render the sounds of the Slavic language variety with which they were acquainted. It is generally assumed that this was the Glagolitic alphabet, which can with some difficulty be linked to cursive Greek minuscule script. The Cyrillic alphabet, on the other hand, is clearly an adaptation of the Greek uncial alphabet. While over the years the Glagolitic alphabet was marginalized (a specific form of the Glagolitic

alphabet has maintained itself in parts of Croatia, though only for liturgical purposes), Cyrillic became the alphabet of Orthodox Slavdom. Generally speaking, Slavs adhering to other Christian creeds have adopted the Latin alphabet.

It will come as no surprise that we find Church Slavic glosses in religious texts written in other languages. The Vienna Glosses and the Gregorian or Prague Glosses (eleventh–twelfth century), for instance, translate Latin text elements.¹ They are written in Roman script and constitute an example of Church Slavic of the Czech recension, which at that time had all but been ousted by Latin. In turn, the widening gap between Church Slavic and the local vernacular gave rise to Slavic glosses in Church Slavic texts. A fourteenth-century East Slavic manuscript of the *Chronicle of George Hamartolos* contains glosses that explain Bulgarian forms.²

In the West Slavic area, the Church Slavic heritage was comparatively insignificant.³ There, the local vernaculars rapidly gained prominence as literary languages, even before the Reformation. A few notable examples are the Old Czech *Chronicle of Dalimil* (early fourteenth century) and the Old Polish *Kazania Świątokrzyskie* ('Sermons of the Holy Cross', fourteenth century). It is in this era that we encounter the first Czech–Latin and Polish–Latin vocabularies. Bartoloměj z Chlumce, also known as Claretus de Solentia, was responsible for three wordlists in verse (c. 1360), translating Latin words for the benefit of readers with scientific interests. A thematically arranged Latin–Czech dictionary, dated 1409, is also attributed to him.⁴ The *Wokabularz trydencki* (1424?), is the oldest Latin–Polish wordlist. It has been suggested that it was compiled at the instigation of Alexander of Masowia, who became bishop of Trento in 1425.⁵ How the author arrived at this particular selection of about 500 lemmata is unknown.⁶ For these early wordlists in their medieval Latin context, see Chapter 13.

The sixteenth century witnessed the appearance of printed dictionaries. The first Latin–Czech dictionary, compiled by Jan Vodňanský (also known as Jan Bosák z Vodňan or Johannes Aquensis), was published in 1511. The entries are arranged according to subject matter in twelve sections. A Latin–German–Polish adaptation of the dictionary which constituted the first

¹ Kempgen et al., *Slavischen Sprachen*, II.1389–90. ² Schenker, *Dawn of Slavic*, 237.

³ Gandolfo, 'Roman Slavdom', 109.

⁴ The wordlists can be found in Flajšhans, *Klaret i jeho družina*, I; Hanka, *Sbírka nejedléjších slovníků*, 54–104, is an edition of the dictionary.

⁵ *Wokabularz trydencki* (1962) is an edition; see Gruszczyński and Saloni, 'From multilingual to monolingual', 206.

⁶ Koerner and Szwedek, *Towards a History of Linguistics in Poland*, 102.

section of the *Pappae puerorum* of the Dutch humanist Joannes Murmellius appeared in Kraków in 1528. With at least sixteen editions, its popularity stretched into the seventeenth century. Another trilingual dictionary published in Kraków that had many new editions was Franciszek Mymer's adaptation (1528) of an anonymous Latin–German–Czech dictionary from 1513. In 1544, Bartłomiej Bydgoszczy produced a handwritten dictionary by adding c. 11,000 Polish interpretamenta to Johannes Reuchlin's *Vocabularius breuiloquus* (for which see Chapter 13). An earlier manuscript by Bartłomiej (1532), containing more than 4,000 Polish words, was based on Latin dictionaries by Reuchlin, Ambrogio Calepino, and the medieval friar Joannes Marchesinus.⁷

As we have just seen, dictionaries containing more than two languages were not uncommon in the sixteenth century. These dictionaries invited the comparison of the various languages and early attempts at etymological research. With respect to the Slavic language area, dictionaries as well as grammars stressed the individual characteristics of the local vernaculars, but at the same time made it ever more transparent what they had in common, thus prompting reflections on the unity of the Slavic languages. Jan Mączyński, for instance, explicitly states in the foreword of his *Lexicon latino–polonicum* (1564) that he regards the languages of the Slavs as dialects of a common language.⁸ The same attitude may underlie the listing of Czech and South Slavic elements under a single heading in Sigismund Gelenius' *Lexicon symphonum* (1537, 1547), and is certainly present in the preface to Matouš Benešovský's *Knjžka slov českých* (1587), the first monolingual Czech dictionary, which is also, to a certain extent, a Slavic etymological dictionary.

The multilingual *Dictionarium quinque linguarum* by Faust Vrančić (1595) contains German, Italian, Hungarian, and Croatian (*lingua Dalmatica*) glosses to a Latin lexicon of 5,500 words. The author argues that Croatian is the purest of the Slavic languages. Earlier printed specimens of Croatian lexicography are a small Italian–Croatian dictionary, the *Opera nuova che insegna a parlare la lingua schiavonesca* (1527) by the merchant Pietro Valentiano and the Latin–Croatian supplement to Bartol Đurđević's *De afflictione* (1544).⁹ Vrančić's dictionary is the first major Croatian lexicographic work and was an

⁷ See Jankowiak, 'Słowniki Bartłomieja z Bydgoszczy', 11–12. An edition of the 1532 manuscript is Bydgoszczy, *Słownik łacińsko–polski* (1900); an edition of both dictionaries is Bydgoszczy, *Słownik: wersja łacińsko–polska* (1999–2012).

⁸ Stankiewicz, *Grammars and Dictionaries*, 47.

⁹ Filipović, 'Beginnings of lexicography in Croatia', 65–7; Considine, *Small Dictionaries*, 77–9.

important source for later dictionaries, such as Hieronymus Megiser's *Thesaurus polyglottus* (1603). Peter Loderecker's *Dictionarium septem diversarum linguarum* (1605) is an enlarged edition of Vrančić's dictionary, to which Czech and Polish wordlists were added.¹⁰

Though the language of the manuscript leaves called the Freising Fragments (c. 1000) is sometimes classified as Old Slovene, there are few Slovene texts that pre-date 1550.¹¹ Important steps towards a Slovene literary language were taken in the second half of the sixteenth century. In 1584, Adam Bohorič wrote the first grammar of the Slovene language, inspired by Melanchthon's grammatical works, and a quadrilingual dictionary published by Megiser in 1592 (and used as a basis for his *Thesaurus Polyglottus*) included c. 9,000 Slovene words, calling their language 'lingua Illyrica'.

In the Orthodox regions we find only a limited number of early dictionaries. As a supplement to his Church Slavic grammar (1596), Lavrentij Zyzanij published a *Leksis*, translating c. 1,000 Church Slavic and other difficult words into the East Slavic vernacular to which he referred as 'prosty Russkij dijalekt'. To be on the safe side, I shall use the designation Ruthenian for this language variety here, where others have opted for Ukrainian, Belorussian, or Russian. The same applies to the target language of the Church Slavic dictionary by Pamva Berynda (1627). Samples of spoken East Slavic, occurring in bilingual lists of words and phrases, are also attested in conversation manuals, such as the High German–Russian manual by Thomas Schroue (1546), the Low German–Russian manual by Tönnies Fenne (1607), and Mark Ridley's bidirectional Russian and English *Dictionarie of the Vulgar Russe Tongue* (1599).¹²

Croatian dictionaries from the seventeenth and eighteenth centuries may combine material from various dialect groups (the classification of Croatian into Štokavian, Čakavian, and Kajkavian dialects is based on the main forms of the interrogative pronoun meaning 'what?'). Jacov Mikalja's *Thesaurus linguae Illyricae* (1649–51) is mainly Štokavian, but contains Čakavian elements as well. The same holds true for Ardelio Della Bella's *Dizionario Italiano, Latino, Illirico* (1728), which focuses on the dialect of Dubrovnik (Ragusa). Ivan Belostenec's dictionary, which following his death was completed by

¹⁰ Stankiewicz, *Grammars and Dictionaries*, 85.

¹¹ See Mikhailov, *Frühslowenische Sprachdenkmäler*, for a survey of Early Slovene manuscripts.

¹² All three have been edited, as Schroue, *Einn Russisch Buch* (1992), 'Tönnies Fenne's Low German manual' (2008), and Ridley, *Dictionarie* (1996); for more examples, see P. Hendriks and J. Schaeken in 'Tönnies Fenne's Low German Manual', i, and Considine, *Small Dictionaries*, 89–90.

Jerolim Orlović and Andrija Mužar and finally published in 1740, contains material from a variety of regions.¹³ By contrast, Juraj Habdelić's *Dictionar* (1670) is a Kajkavian–Latin dictionary. The status of the local vernaculars with a view to the development of a unifying literary language became a particularly important issue in the nineteenth century.

Among the West Slavic literary languages that, compared with Czech and Polish, emerged at a late stage are Slovak and Sorbian. In the case of Slovak, its late literary emergence is due to the long-standing dominance of Czech. The earliest dictionary containing Slovak material is the *Verborum in institutione grammatica contentorum in Ungaricum et Sclavonicum translatio* ('Translation into Hungarian and Slavonic of the words contained in the Introduction to Grammar', 1648), also known as the Trnava dictionary. The first attested instance of Sorbian consists of fragments from a glossed Psalter called the 'Magdeburger Glossen' (twelfth century). Extensive written texts begin to appear in the sixteenth century. Megiser's *Thesaurus polyglottus* also contains Sorbian words.¹⁴ The first printed dictionary is Jurij Hawštyn Swětlik's *Vocabularium Latino–Serbicum* (1721). Polabian is a West Slavic language that was sparsely recorded in the period 1650–1750, when extinction was imminent. The most important document is the 'Vocabularium Venedicum' by the pastor Christian Hennig, which remained in manuscript until the twentieth century.¹⁵

The so-called Illyrian movement (1835–48), which was inspired by pan-Slavism, had a considerable impact on the linguistic situation in the South Slavic area.¹⁶ Initially, its members, such as the Croat Ljudevit Gaj, strove for a supra-national Illyrian literary language that would integrate elements from all the South Slavic vernaculars. This endeavour proved fruitless, mainly because the national consciousness of the Slovenes and Serbs had recently gained strength. On the other hand, the fact that the protagonists of the movement had switched from Kajkavian to Štokavian paved the way for what was to become known as Serbo-Croatian. The Vienna Literary Agreement (1850) stipulated that for Serbs and Croats alike the Ijekavian variant of Štokavian should serve as a literary language. Standardizing an existing vernacular was deemed preferable to creating an artificial literary standard. This is in line with the views of Vuk Karadžić, the author of a highly

¹³ See Vončina, 'Belostenčeva hibridni jezik'.

¹⁴ According to Schaarschmidt, *Historical Phonology*, 24, these forms originate from a dialect spoken around the Oder.

¹⁵ See Polański and Sehnert, *Polabian–English Dictionary*, 15–18, and Rost, *Sprachreste*, 87–176.

¹⁶ See Greenberg, 'Illyrian movement'.

influential Serbian–German–Latin dictionary (¹1818, ²1852), who in his linguistic reforms gave priority to the living language. His adaptation of the Cyrillic alphabet, for instance, was an attempt to bring the orthography closer to everyday speech.

The Enlightenment proved to be a stimulus for Slovene lexicography, but not all manuscripts found their way into print, for example the still unpublished German–Slovene–Latin dictionary by Valentin Vodnik, to which the author had made his last addition by 1817.¹⁷ A printed lexicon from the first half of the nineteenth century is Anton Murko's bidirectional dictionary of German and Slovene (1833). This publication still employed the orthography devised by Bohorič. It was subsequently replaced by an orthography involving what is basically the alphabet that was devised by Gaj for Croatian, which corresponded neatly with the variant of Cyrillic that Vuk Karadžić had designed for Serbian. Under the editorship of Matej Cigale, a German–Slovene dictionary appeared, which contained the material gathered by Vodnik. The Slovene–German part of the project was published by Maks Pleteršnik in 1894–5. This comprehensive dictionary was very influential, among other things in matters of orthography, and of great significance to historical linguistics. Pleteršnik, by the way, had studied in Vienna, where he contributed to the monumental Church Slavic dictionary (1862–5) of the Slovene scholar Franc Miklošič.

In the Bulgarian and Macedonian language area, the Old Church Slavic period was followed by the Middle Bulgarian period, the beginning of which may be dated to the second quarter of the twelfth century.¹⁸ Due to complex political circumstances, the rise of the modern literary languages was a cumbersome affair. It is telling that the first description of Bulgarian was provided by the Serb Karadžić in 1822.¹⁹ (Karadžić's research was inspired by the Slovene linguist Jernej Kopitar, who held the view that Bulgarian was an independent language rather than a Serbian dialect, as had been assumed by others.) A Bulgarian vocabulary recommended by Karadžić, the Bulgarian part of the four-language dictionary of Daniil of Moscopole, actually reflects the Macedonian vernacular of the Bitola region.²⁰ The first comprehensive Bulgarian dictionary was compiled by Najden Gerov over a period of fifty years, and published between 1895 and 1904, the last two volumes posthumously, with a supplement by Todor Pančev in 1908.

¹⁷ Legan Ravnikar, 'Characteristics', 80.

¹⁸ Schaeken and Birnbaum, *Altkirchenslavische Schriftkultur*, 54. ¹⁹ Karadžić, *Dodatak*.

²⁰ The dictionary is part of Daniil's *Eisagōgikē didaskalia*.

The Romantic nationalism of the nineteenth century stimulated the study of Slavic vernaculars and folklore. As we have seen, this produced dictionaries of languages that hitherto had not received much attention. At the same time, the nineteenth century was the era of multi-volume dictionaries, some of them monolingual, which often incorporated earlier lexicography. The first monolingual Polish dictionary was Samuel Linde's (c. 60,000 entries in 6 volumes). Linde, a native speaker of German, provided many quotations from both literary and non-literary sources. Apart from his own era, he covered the Polish literature of the sixteenth century particularly well.²¹ The so-called Vilnius Dictionary (1861), edited by Aleksander Zdanowicz and others, is a monolingual Polish dictionary that was intended as a practical tool. Hence, the number of quotations is kept to a minimum, while comparisons with other languages are rare. Though the dictionary comprises merely two volumes, the number of entries is 110,000. Josef Jungmann's Czech–German dictionary in five volumes (1834–9) includes not only the lexicon of Czech literature from the fourteenth and fifteenth centuries onwards, but also neologisms and borrowings from other Slavic languages. In this way Jungmann tried to establish a full-fledged Czech lexicon while reducing the number of German borrowings. By contrast, attempts by Slovaks at codifying their language were only moderately successful. The Catholic priest Anton Bernolák propagated a literary language that was based on the West Slovak dialects. In spite of its size, his Slovak/Czech–Latin–German–Hungarian dictionary, published posthumously in six volumes (1825–7), never gained authoritative status but, then again, such may not have been the author's intention.²² Subsequent efforts to establish a Slovak standard language were headed by Protestants like Ľudovít Štúr, who used the central dialects as a basis. These activities did not yield any dictionaries of particular significance.

Following the example of the *Dictionnaire de l'académie française* (for which see Chapter 14), the Imperial Russian Academy, which was founded in 1783 by Catherine the Great, commissioned the compilation of a normative dictionary, *Slovar' Akademii rossijskoj*, which was published in six volumes (1789–94). One of the academy's objectives was to diminish the use of foreign words, but at the same time the lexicon was far removed from the vernacular. This was a consequence of the fact that the Church Slavic element was strongly

²¹ Piotrowski, 'Development', 187; Gruszczyński and Saloni, 'From multilingual to monolingual', 214.

²² According to Kamusella, *Politics of Language*, 534, Bernolák wished to further the knowledge of Hungarian among the speakers of Slavic in Upper Hungary.

represented in the Russian literary language as established by Mikhail Lomonosov, as well as in the Bible.²³ The conservative character of the *Slovar' Akademii Rossijskoj* was maintained in the subsequent three editions, the fourth edition appearing in 1867–8.²⁴ A completely different type of dictionary was compiled by Vladimir Dal', who tirelessly collected words and sayings on his travels through the Russian empire. His dictionary in four volumes (1863–6) contains a wealth of dialect material and reflects the ethnographic interests of the author. The third and fourth editions were prepared by Jan Baudouin de Courtenay, who introduced words that had previously been deemed unfit for publication. Soviet editions, which contained their own instances of censorship, were invariably based on the second edition.

The political developments of the twentieth century obviously had far-reaching consequences for the standardization of the various Slavic languages. In Czechoslovakia, for instance, the Slovak language finally acquired official status. In practice, however, the language played second fiddle to Czech and attempts to stress its independent character were looked upon with suspicion.²⁵ It was not until well after the Second World War that the first authoritative Slovak dictionary was published, edited by Štefan Peciar, in six volumes (1959–68). A similar Czech dictionary is the nine-volume *Průruční slovník jazyka českého* (1935–57). Worth mentioning is the comprehensive monolingual Czech dictionary by František Trávníček and Pavel Váša (1937).

In Poland, the first quarter of the twentieth century saw the publication of another large monolingual dictionary, *Słownik języka polskiego* (1900–27), edited by Jan Aleksander Ludwik Karłowicz, Adam Antoni Kryński, and Władysław Marcin Niedźwiedzki. With c. 280,000 entries, this dictionary, which is commonly referred to as the Warsaw Dictionary, superseded all previous Polish dictionaries in size. This results from the fact that it includes dialect material – Karłowicz was also the compiler of an important multi-volume dialect dictionary, *Słownik gwar polskich* (1900–11) – and lexicographic material from earlier periods. The eleven-volume post-war dictionary edited by Witold Doroszewski, on the other hand, starts from the Polish vocabulary

²³ Among Lomonosov's many publications on a wide variety of subjects were *Rossijskaja grammatika*, an influential Russian grammar, and 'Predislovie o pol'ze knig cerkovnyx v rossijskom jazyke', a treatise in which he unfolded his theories of the three styles, in which Church Slavic elements belonged to the 'high style' and, to a lesser extent, to the 'middle style'. See also Considine, *Academy Dictionaries*, 156, and Offord, Argent, and Rjéoutski, 'French and Russian in Catherine's Russia', 38–9.

²⁴ Farina, 'Marrism and Soviet lexicography', 161–2.

²⁵ See Kamusella, *Politics of Language*, 879–81.

of the second half of the eighteenth century.²⁶ Unlike its predecessors, it was intended to be normative, which obviously had an effect on the selection of the entries. That does not detract from the fact that Doroszewski's dictionary (c. 128,000 entries based on 6.5 million slips) is fairly complete and excels at providing well-documented citations.²⁷ The vocabulary of the older periods is being covered by separate dictionary projects: *Słownik staropolski*, an Old Polish dictionary (1953–2002), and dictionaries of the Polish of the sixteenth century, and of the seventeenth and the first half of the eighteenth centuries.

The Russian Academy dictionary survived into the twentieth century but, after the understandably chaotic publication of the sixth edition between 1897 and 1929, there never was to be another, in spite of preparations for a new edition in fifty-six volumes.²⁸ Instead, the Leningrad Academy of Sciences compiled the seventeen-volume *Slovar' sovremennogo russkogo literaturnogo jazyka* ('Dictionary of modern standard Russian', 1948–65), which was normative and confined itself to the Russian lexicon from the times of Pushkin onwards. Meanwhile Dmitrij Ušakov had managed to publish *Tolkovyj slovar' russkogo jazyka* ('Explanatory dictionary of Russian') in four volumes (1935–40). It served as a basis for Sergej Ivanovič Ožegov's well-known one-volume dictionary, which was first published in 1949.

Throughout the eighteenth and nineteenth centuries, Ukrainian struggled to be recognized as a language rather than a dialect of Russian. In 1876, the tsarist regime imposed a ban on publications in Ukrainian, which was mitigated in 1881 and eventually lifted in 1905. In Galicia, which had become part of the Austro-Hungarian empire, speakers of Ukrainian enjoyed relative freedom. In the second half of the nineteenth century, there were several unsatisfactory attempts at codifying the Ukrainian lexicon.²⁹ However, this was also the period in which the foundations were laid for an important Ukrainian–Russian dictionary. Its compiler, Borys Hrinčenko, tried his best to include the living language, disregarding words that had been coined by lexicographers. As to the excerpting of literary sources, however, the journal *Kievskaja starina*, which played a role in the publication of the dictionary, obliged Hrinčenko to limit himself to the period 1798–1870.³⁰ Despite its limitations, Hrinčenko's dictionary was of the utmost importance for

²⁶ Gruszczyński and Saloni, 'From multilingual to monolingual', 218.

²⁷ Piotrowski, 'Development', 190–1; Gruszczyński and Saloni, 'From multilingual to monolingual', 217–18.

²⁸ Farina, 'Marrism and Soviet lexicography', 164.

²⁹ Ptashnyk, 'Wörterbuch der ukrainischen Sprache', 352.

³⁰ Ptashnyk, 'Wörterbuch der ukrainischen Sprache', 355.

Slovnyk ukrajins'koji movy, the Ukrainian Academy dictionary, edited by I. K. Bilodid (1970–80).

Following the Sorbian revival of the nineteenth century, a number of dictionaries were published, such as the Upper Sorbian–German dictionaries of Christian Traugott Pfuhl (1866) and Filip Jakubaš (1954). A remarkable work is Arnošt Muka's comprehensive dictionary of Lower Sorbian and its dialects, published first in St Petersburg (or, as it became in the course of publication, Petrograd) and then in Prague between 1911 and 1928.

As we have seen, the Vienna Literary Agreement was a step towards a common Serbo-Croatian literary language. The first of the twenty-three volumes of the monumental *Rječnik hrvatskoga ili srpskoga jezika* appeared as early as 1880. In 1959, the Serbian Academy of Sciences and Arts published the first volume of the equally ambitious *Rečnik srpskohrvatskog književnog i narodnog jezika*. In spite of the fact that nineteen volumes have already been published, this project is not even near completion. Since the dissolution of Yugoslavia there has been a growing tendency to treat Serbian and Croatian as separate languages. Furthermore, Bosnian and Montenegrin have gained independent status. The Slovene Academy dictionary, the *Slovar slovenskega knjižnega jezika*, was finished in 1991, the year in which Slovenian independence was proclaimed. As part of the codification of Macedonian, which had become an official language in Yugoslavia, the lexicon was recorded in the *Rečnik na makedonskiot jazik* (1961–6).

Baltic

The earliest Baltic linguistic monument is the German–Old Prussian Elbing Vocabulary, a thematically arranged list of 802 lexical items. The vocabulary was part of a manuscript called the Codex Neumannianus, which was lost in the Second World War. While the codex dated from c. 1400, the vocabulary may have been a copy of an older original. A later wordlist (early sixteenth century) is Simon Grunau's Vocabulary, which contains 100 Old Prussian words with German translations. There are eight copies, exhibiting various degrees of corruption. Leaving aside a few minor documents, the only surviving texts written in Old Prussian are three catechisms from the sixteenth century.³¹ As Old Prussian is the only West Baltic language that has

³¹ Mažiulis, *Prūsų kalbos paminklai*, is an edition; see also the *Database of the Old Prussian Linguistic Legacy*.

been recorded to some extent, I shall in the following pages confine myself to the lexicography of the East Baltic languages Lithuanian and Latvian.

The first written records of the East Baltic languages date from the sixteenth century. Printed books from the sixteenth and seventeenth centuries are generally connected with either Catholicism or the Reformation. For the development of the Lithuanian language, the fact that a substantial number of speakers lived outside the Duchy of Lithuania, in the secularized Duchy of Prussia (subsequently the province of East Prussia), proved to be important, because books could be published more freely there. The eighteenth century was the starting point of a tradition of grammars and dictionaries from Prussian Lithuania or, as it is also called, Lithuania Minor.

The first Lithuanian dictionary, however, appeared in the Duchy of Lithuania. It was compiled by Konstantinas Sirvydas and is one of the first printed texts in the eastern variant of the Lithuanian language. Since the title page of the only extant copy of the first edition is missing, the year of publication is uncertain, but it has been argued that it must have been before 1620. An expanded edition was apparently published in 1631 (there are no surviving copies), followed by posthumous editions in 1642, 1677, and 1713. Sirvydas' dictionary was trilingual, Polish–Latin–Lithuanian.

In the Duchy of Prussia, Daniel Klein, a pastor in Tilsit (now Sovetsk, in the Kaliningrad Oblast of Russia), published two influential Lithuanian grammars (1653, 1654), but regrettably his Lithuanian dictionary remained unpublished. The same holds true for several other seventeenth-century dictionaries, most of them anonymous. The first printed dictionary intended for Prussian Lithuania was the *Vocabularium* of Friedrich Wilhelm Haack (1730), which consisted of a Lithuanian–German part and a smaller German–Lithuanian one. Haack was a teacher at the Lithuanian Seminar of the University of Halle, which had been founded to supply preachers for the Lithuanian inhabitants of East Prussia. The dictionary was primarily based on the language of printed texts, particularly the New Testament.³² It was superseded by Philipp Ruhig's bidirectional Lithuanian and German dictionary (1747). Ruhig, who is also known as an author of religious works, had an interest in folk literature, which is reflected in his dictionary.³³ Ruhig's dictionary provided the basis for the one published by Christian Gottlieb Mielcke in 1800. Among Mielcke's additional sources were several dictionaries in manuscript, of which the most important was compiled by Jacob

³² Schiller, 'Litauischen Seminare', 390.

³³ Sabaliauskas, *Lietuvių kalbos tyrinėjimo istorija*, I.28.

Brodowski, and included forms heard in everyday speech and proper nouns, as well as religious vocabulary.³⁴

In the second half of the nineteenth century, scholars of comparative Indo-European linguistics displayed a growing interest in the Baltic languages. August Schleicher's Lithuanian grammar and reader (1856–7) may be considered a milestone in the study of Lithuanian. G. H. F. Nesselmann's Lithuanian–German dictionary, which had been published in 1851, was for several decades the most important lexical source for Indo-Europeanists. It was subsequently replaced by the German–Lithuanian and Lithuanian–German dictionaries compiled by Friedrich Kurschat, which appeared between 1870 and 1883. Unlike Nesselmann, Kurschat was a native speaker of Lithuanian, who did not have to rely on fieldwork done by others and, among other things, was able to perceive tonal contrasts on monophthongs. Kurschat's dictionaries, as well as his *Grammatik der littauischen Sprache*, constitute a faithful description of the literary language of Lithuania Minor, at a time when the process of Germanization had been sped up.

In tsarist Russia, which ruled the territories of the former Duchy of Lithuania from 1795 onwards, the position of Lithuanian was by then even worse. In spite of a ban on the use of the Latin alphabet, an intended project of Russification, which in some areas was rivalled by Polonization, ultimately failed. The centre of the Lithuanian revivalist movement was located in Suvalkija, in the south-west of Lithuania. For those propagators of the Lithuanian language who were speakers of the local dialects it was completely natural to turn to the publications of Schleicher and Kurschat, as they described a variety of Lithuanian that was close to their own. Following the independence of Lithuania in 1918, the codification of the literary language became the task of linguists such as Jonas Jablonskis. This included the reduction of foreign elements and the creation of neologisms.

An interesting dictionary, which through the help of Russian linguists was excepted from the ban on printing Lithuanian in the Latin alphabet, was compiled by the priest Antanas Juška. It was remarkable for being based on the spoken language and containing many illustrative phrases. Though it suffers from a number of deficiencies, of which the unreliable accent marking is probably the most important, its significance is beyond doubt. The publication of the dictionary was a laborious process. Between 1897 and 1922, three volumes appeared, containing only a part of the material gathered by

³⁴ Brodowski, *Lexicon Germanico–Lithvanicvm* (2009), is an edition.

Juška, who had died in 1880.³⁵ Among the many notable linguists involved in the editing process was Kazimieras Būga, who as early as 1902 had started his own collection of card files.³⁶ Būga initiated a project that was to result in the most comprehensive dictionary of the Lithuanian language, the monolingual *Lietuvių kalbos žodynas* (Dictionary of the Lithuanian language; 1941–2002). Sadly, Būga only managed to complete two fascicles, of which the second was published posthumously. Būga's untimely death in 1924 was a devastating blow for Lithuanian linguistics. The dictionary project was interrupted until 1930, when the linguist Juozas Balčikonis became editor. It then took a while before the aims and the scope of the project were defined. After the war several reorientations took place, some of them ideological. The first two volumes, published in 1941 and 1947, were re-edited in 1968 and 1969, respectively.

Probably the most important bilingual dictionary is the *Wörterbuch der litauischen Schriftsprache, Litauisch–Deutsch* by Max Niedermann, Alfred Senn, Franz Brender, and Antanas Salys. The original editors, Niedermann and Senn, had agreed to team up with Būga, but their plans were thwarted by the latter's death.³⁷ The team was first joined by Brender and later by Salys, who is known for creating the first Lithuanian dialect map. The posthumously published Lithuanian–German dictionary of Alexander Kurschat is wide-ranging, but for the greater part a compilation of earlier dictionaries, including those by his uncle Friedrich Kurschat.

The first Latvian dictionary, *Lettus*, published in 1638 by the Lutheran theologian Georg Mancelius, consists of an alphabetical German–Latvian lexicon, with about 6,000 headwords, with a second part called *Phraseologia lettica*, a thematically arranged German–Latvian phrasebook. It was apparently intended to facilitate communication between German clergymen or land-owners and Latvians. Another seventeenth-century dictionary, compiled by the Jesuit Georg Elger, contained approximately 14,000 headwords and was published in Vilnius. Apparently, the third edition of Sirvydas' Lithuanian dictionary served as a model for this Polish–Latin–Latvian dictionary.³⁸ This is the first dictionary in which Latin type (rather than Gothic) is used to render the Latvian forms. One of several dictionaries that survive only in manuscript is Christoph Fürecker's *Lettisches und Teutsches Wörterbuch*.³⁹ Fürecker, who had

³⁵ Sabaliauskas, *Lietuvių kalbos tyrinėjimo istorija*, I. 163–6; cf. Zinkevičius, *Lietuvių kalbos istorija*, IV.218.

³⁶ Būga, in *Lietuvių kalbos žodynas*, fascicle 1, v.

³⁷ Sabaliauskas, *Lietuvių kalbos tyrinėjimo istorija*, I. 230. ³⁸ Veisbergs, 'Overview', 310.

³⁹ Fürecker, *Lettisches und Teutsches Wörterbuch* (1997), and Fürecker, *Lettisches und Teutsches Wörterbuch* (1998), are editions of the surviving manuscripts (neither of which is authorial); *Manuale Lettico–Germanicum* (2001) is an edition of a work derived from Fürecker's.

an excellent knowledge of Latvian and, for instance, translated hymns into Latvian, recorded words he had heard in everyday speech, in addition to incorporating material from Mancelius' dictionary.⁴⁰

The most important eighteenth-century Latvian dictionaries were compiled by Jacob Lange (1777) and G. F. Stender (1789). Lange's dictionary is the first to consist of a Latvian–German as well as a German–Latvian part, presenting about 10,000 and 15,000 lemmata respectively. Lange did not merely regard his dictionary as a tool for improving the reader's competence in Latvian, but also intended it to be a source of encyclopedic knowledge.⁴¹ Stender's *Lettisches Lexikon* was preceded by *Entwurf eines Lettischen Lexici*, the first printed Latvian–German dictionary, which was a supplement to his Latvian grammar of 1761. The dictionary of 1789 consisted of two volumes, a Latvian–German volume that surpassed the *Entwurf* in extent (7,000 entries as opposed to 4,000), and a German–Latvian volume (c. 14,000 entries); A. G. Wellig's *Beiträge zur lettischen Sprachkunde* (1828) may be viewed as a supplement to it. In Stender's lexicological works, as well as in his grammar, the Latvian forms were printed in Latin type, while Lange used Gothic type throughout.

In the nineteenth century dictionaries of the traditional type were still appearing. A late example is Karl Ulmann's Latvian–German dictionary (1872). Ulmann was a professor of theology who managed to become a Lutheran bishop in St Petersburg. His dictionary contained c. 20,000 entries and was primarily intended for a German readership.⁴² A German–Latvian second part was completed by Gustav Brasche. A different, but nonetheless familiar, type of dictionary is represented by Jan Kurmin's trilingual dictionary (1858), which was published for the benefit of Catholic priests. It is based on Elger's Polish–Latin–Latvian dictionary and the fifth edition of Sirvydas' *Dictionarium trium linguarum* (1713).⁴³ The Latvian component reflects the language of Latgale, which was a predominantly Catholic region.

Krišjānis Valdemārs, one of the leading members of the movement known as First Awakening (*Pirmā atmoda*), was also the editor of a Russian–Latvian–German dictionary (1872). The preface makes it clear that the dictionary was meant for Latvians wishing to learn Russian and vice versa: it seems that the German element was introduced at a later stage and was indeed dropped again, for later editions of the dictionary are bilingual.⁴⁴ As the so-called

⁴⁰ Zemzare, *Latviešu vārdnīcas*, 73; Karpinska, *English–Latvian Lexicographic Tradition*, 97.

⁴¹ Balode, *Deutsch–lettische Lexikographie*, 59–60. ⁴² Veisbergs, 'Overview', 317.

⁴³ Veisbergs, 'Overview', 316. ⁴⁴ Karpinska, *English–Latvian Lexicographic Tradition*, 101.

Young Latvians strived to reduce the influence of the German Balts and the German language, the dictionary, which has about 37,000 Russian head-words, contains neologisms.

The most important Latvian dictionary was initiated by Karl Mühlenbach, who started working on it in the last decades of the nineteenth century. After Mühlenbach's death in 1916 his work was continued by Latvia's greatest linguist, Jānis Endzelīns. The dictionary appeared between 1923 and 1932 and contained 77,175 entries. It was followed by a sizeable supplement (55,543 entries) compiled by Endzelīns and Edīte Hauzenberga (1934–46).⁴⁵ Endzelīns focused on genuine Latvian forms, removing rarely used borrowings and neologisms from Mühlenbach's manuscript.⁴⁶ The material was gathered from written sources, such as older dictionaries, as well as from the spoken language, and the dictionary contained archaic and regional forms. The language of the *dainas* (folk songs) occupies an important position. Thus, we are dealing with a comprehensive representation of the Latvian lexical stock, with an emphasis on the inherited lexicon. It is worth noting that the Latvian public was involved in the project on a remarkable scale (a predecessor in this respect is Wellig's *Beiträge zur lettischen Sprachkunde*).

In the dictionary of Mühlenbach and Endzelīns as well as in its supplement, dialect forms were normalized. In addition to marking the tone (or tones) of a given form, Endzelīns employed the sign ² to indicate a conflation of two tones, which information is crucial to the interpretation of the tone from a historical point of view. The etymology of the lemmata, a field which Mühlenbach gladly left to Endzelīns, represents the state of the art at that time. With respect to its importance for comparative linguistics, we may observe that the dictionary is not monolingual but that its Latvian–German format enhances its accessibility. Latvian monolingual dictionaries appear relatively late. The *Latviešu valodas literārās vārdnīca* (Dictionary of the Latvian literary language), which has 80,000 entries, is essentially normative and suffers from various forms of censorship.

⁴⁵ Entry count from Veisbergs, 'Translators' tools', 139.

⁴⁶ J. Endzelīns in Mühlenbach and Endzelīns, *Lettisch-deutsches Wörterbuch*, I.4.

The Germanic Languages Other than English from c. 1700

ULRIKE HAB

The Changing Role of Lexicography Within a Society

There are different sorts of dictionaries used within a language community. Monolingual dictionaries usually play the most significant role in the standardization of a (young) language, while bilingual ones mostly foster interlingual communication following situations of language contact. Monolingual dictionaries are usually large; they are aimed at the standard variety of a language and are usually available in different editions for different users; they are almost always written by native speakers. Bilingual wordlists at an earlier stage in history are written by persons not belonging to the language community. Later on, when the monolingual standard dictionary has been established, bilingual lexicography often yields another type of dictionary. These are written from inside the language community and are meant to promote comprehension in situations of language contact with other communities.

Thus, the focus of this chapter will be on monolingual lexicography, because this sort of lexicography interacts with the process of language standardization. But bilingual lexicography will be treated, shedding light on language contact. The question of which languages are linked together by bilingual or multilingual wordlists is strongly related to the historical conditions of the respective language communities.

German Dictionaries in the Age of Enlightenment

The role of language changed when absolutist sovereigns recognized the interdependence of scientific development and economic benefits. The wish to know more about the real world had to involve large parts of the

population, not only scholars, so that Latin was no longer sufficient to foster progress. Craftsmen and manufacturers needed to be informed of the results of scientific investigations. According to masterminds such as Gottfried Wilhelm Leibniz, the German language needed some improvement to play that role. Poets also rose to speak up and express their interests. In 1697, Leibniz wrote his 'Unvorgreiffliche Gedancken' ('Fundamental considerations'), in which he proposed different sorts of dictionaries, among them an etymological dictionary, a dictionary of German dialects, a thesaurus, and a dictionary of language for special purposes. In 1711, the king of Prussia, Friedrich Wilhelm I, commissioned the Berlin Society of Science (later to become the Academy of Sciences) to publish a complete German dictionary.¹ But things went differently than planned, so that the three important dictionaries described in the following paragraphs indeed refer to the project, but cannot be considered as its fulfilment. Lexicographers had their own ideas and were individually influenced by time, money, and social circumstances.

One central aspect of the long discussion was the definition of an overarching standard language by means of lexicography. But the process of standardization lasted, as we know today, until the beginning of the nineteenth century, so that the lexicographers of the eighteenth century only tried to impose one or another regional variety as the standard.

Christoph Ernst Steinbach published his *Vollständiges Deutsches Wörter-Buch* in 1734.² He introduced two major innovations. First, he marked words according to four language varieties, one of which was a sort of pre-form of standard German while the others were dialectal, old-fashioned words, and words which were unusual in the written 'standard'. He used little stars and crosses for marking. Because he copied material from much older dictionaries, there were many old and unusual words in his dictionary. Secondly, he was the first lexicographer in the German-language tradition to give literary citations, though not at all systematically. Before compiling his dictionary, he had edited the works of Johann Christian Günther, a Silesian poet, and he was also interested in other Silesian poets. So he took examples from them and gave the author's name in order to provide authority for the 'good' language, but he did not give precise references.

Johann Leonhard Frisch published his *Teutsch-Lateinisches Wörter-Buch* in 1741. As the son of a Protestant theologian from Bavaria, he belonged to the

¹ The lexicographical programme and its discussion are described in Reichmann, 'Geschichte lexikographischer Programme'.

² For more details, see introduction in Steinbach, *Vollständiges Deutsches Wörter-Buch* (1793), and Wiegand, 'Historische Lexikographie', 659–60.

middle classes. He studied a broad range of subjects, worked as a farm administrator, a teacher, and a translator, and travelled around Europe before he settled down near Berlin and wrote about diverse subjects: natural sciences as well as historical and philological subjects. His interests were more practical than academic.³ The innovative characteristics of his dictionary are as follows. First, in order to collect and array his words, he used older principles of German word formation but avoided any artificial forms; he focused on contemporary usage. Secondly, he gathered old and specialized words from texts of several crafts, hunting, fishing, shipping, mining, salt making, forestry, housekeeping and home economics, pharmacy, medicine, and so on. Thus, he was the first lexicographer to practise excerpting in a systematic manner, which he did with texts from the Middle Ages and the Renaissance. This older language material was integrated with the contemporary language that Frisch knew, for which he did not need citations. He only mentioned the names of a few special and well-known authors, namely Martin Luther and Leibniz, and he often cited from the famous German encyclopedia *Zedlers Universal-Lexicon* (1732–54). His intention in giving old words and expressions was to close the widening gap between the contemporary language and what was recognized as the pure original German language, to enrich the modern language from its roots, and to make sure that the old texts could be understood if the dictionary was at hand. Finally, he gave all kinds of information in his entries, about form (spelling, word formation, inflection, etymology of form) as well as about actual and historical meaning, usage, and encyclopedic information. Among the latter categories, there were invented examples as well as commentaries on the meaning and usage of the word. For example, an entry such as *aber* ('but') is like a window opening onto the everyday life and everyday language in the early eighteenth century. Frisch even distinguished between social varieties, for example the different words for 'supper' used by the nobility and by the common people.

No dictionary has been the object of more philological research than Johann Christoph Adelung's *Grammatisch-kritisches Wörterbuch der hoch-deutschen Mundart*, in four volumes.⁴ Indeed, it is still fascinating today. It still gives an impression of modernity, maybe because it interrupted the series of more or less historical dictionaries which had been predominant in the German lexicographical tradition – even the work of Frisch, who represents

³ For more details on Frisch, see Wiegand, 'Historische Lexikographie', 660–2.

⁴ The most recent volume about Adelung is Kämper et al., *Aufklärer, Sprachgelehrter, Didaktiker*.

the early Enlightenment, was still part of this series. The title of Adelung's dictionary might be misunderstood; it means that words belonging to the standard variety of high cultural prestige (*hochdeutsche Mundart*) are analysed (*kritisch*) by linguistic aspects (*grammatisch*). There is little historical or etymological perspective on words. In fact, around the year 1800 the process of language standardization almost reached its peak, and Adelung's dictionary played an important role in this.

Adelung studied grammar and philology and was later librarian to the elector of Saxony at Dresden. He was asked by a publisher to complete what the poet Johann Christoph Gottsched could not achieve: a new German dictionary following the plans of the Berlin Society of Science mentioned above. His interests were orthography, German dialects and linguistics. His main innovations in lexicography were as follows. First, the structure of the entry is very clear, and for the first time an entry looks a bit like what we are used to nowadays. It has a semantic focus. Secondly, one of Adelung's main intentions was the explicit description and explanation of word meaning. Etymology is meant to support the understanding of meaning, not to prove the word's value by its heritage. Definition followed the classical scheme of *genus proximum* and *differentiae specifica*. Thirdly, his style of commenting is reasonable and descriptive, sometimes even debating, not at all prescriptive. Fourthly, he created a systematic and principle-based way of ordering the different senses of an entry word; he called this classification of meanings *Linnäisch*, after Carl Linnaeus, the Swedish biologist, who invented the taxonomy for the naming of organisms, a comparison that expressed Adelung's rational and reasonable point of view. His system, called the 'ladder of senses', was based on categories such as *figürlich* ('figuratively') that were borrowed from rhetorical principles and still are in line with modern lexicographic ideas about sense disambiguation. Fifthly, Adelung invented labels for five levels of style marking the distance of words from the 'good' standard language. The fifth, lowest level is completely omitted in his dictionary and only a few words belonging to the fourth level are included. Adelung may have taken the idea from Steinbach, but he developed it further, comprising social varieties of language and giving names to each level that are still in use today: *dichterisch* (poetic), *gehoben* (elevated), *vulgar* (vulgar), *obszön* (obscene). The default level in between elevated and vulgar was not labelled explicitly. Sixthly, the system of marking words was extended to foreign words, if they were 'good' enough to be included. Seventhly, and finally, he extended the presentation of literary examples accompanied by the name of a good author.

What is less noted in the research on Adelung is the fact that he took information, especially concerning special domains (crafts, arts, and the science of his time) from an economic-technical encyclopedia of his time, the *Oeconomische Encyclopädie* founded by Johann Georg Krünitz, whereas the encyclopedia's authors used Adelung's dictionary. Thus, dictionaries and encyclopedic lexicographs were not, in this case – and there are others – as strictly separated in nature as might be suggested by the different terms which refer to them.

Adelung's work was a milestone, stimulating further discussion about one of the most important questions of German lexicography: whether lexicography should describe the usage or standardize the language according to somebody's rules. In fact, Adelung achieved both.

German Lexicography, Patriotism, and Nationalism

Two factors shaped the lexicography of German between the end of Napoleon's regime and the beginning of the First World War. The first was the rise of the *Bildungsbürgertum* ('educated middle class') in the nineteenth century. Dictionaries never ceased to belong to the history of education, but now significantly more people from the middle classes wished to read, to learn, and to participate in political discussions and, if possible, in political life, while the ideas of democratization permeated all countries. The market for reference works (dictionaries and encyclopedias) promising education independently from institutions increased significantly. The *Konversationslexika* were especially appreciated by women and the German-Jewish population, who were often not admitted into universities and schools of higher education.⁵ The second factor was the construction of a fundamental ideology for the German language in conjunction with German history, especially medieval history. The two were equally instrumental in the creation of German identity. It was the period of nationalism throughout Europe, but in Germany there was no unified state, and nationalism had to focus on culture (*Kulturnation*), above all on history and language.

The more and more educated middle classes thought that the only national unifying factors were their language (by now well on its way to standardization) and their common history. Ideologies always simplify and

⁵ Clear statements are impossible because 'Germany' existed as more than thirty sovereign states, each of which had its own judicial and educational system.

distort reality. Thus, the Old High German and the Middle High German languages were far more prestigious than the contemporary language, which was judged to be declining and mixed with unnecessary foreign words borrowed from 'the enemy'. Also, there was the stereotype of good old brave Germans fighting first against the Romans and later, repeatedly, against the French. Consequently, the lexicographical treatment of the German language had to be strongly historically oriented, while at the same time ignoring all the evidence of language contact.

In this social context, the notion of *Fremdwort* ('foreign word' or 'loan-word') was created (there is no expression for 'hard word' in German). It always had negative connotations, and most Germans would insist that foreign words should be distinguished from language which was (thought to be) of pure German origin. So, the dictionary for foreign words was, and still is, a German speciality; this type of dictionary constitutes a class of its own.

Science and the humanities reflected public opinion: the birth of German philology as an academic discipline dates from the beginning of the nineteenth century; of course it was historically oriented. The discovery that the Indo-European languages are a family and that there must have been a common parent language made it possible for the first time to establish serious etymologies. One of the many scientists involved in these linguistic discoveries, besides William Jones and Rasmus Rask, who came from other European countries, big or small, was Jacob Grimm.

Jacob and Wilhelm Grimm's *Deutsches Wörterbuch*

The idea of a new and complete dictionary of German as a national symbol appeared and developed in the 1820s and 1830s, and eventually two of the most famous personifications of German *Kultur*, Jacob and Wilhelm Grimm, accepted the task of lexicography.⁶

Their idea of the *Deutsches Wörterbuch* (DWB) deliberately fostered the symbolizing role of the dictionary, but they also had academic objectives. They wanted to test their method of etymological deductions on a large scale, with hundreds and thousands of words, and demonstrate its validity. The second objective, perhaps less urgent, was to provide an orthography which might be adopted as a unifying standard. Of course, their orthographic

⁶ For more on the Grimms' dictionary, see Haß-Zumkehr, *Deutsche Wörterbücher*, 119–42, and Kirkness, 'Deutsches Wörterbuch'.

principles obeyed historical models instead of phonetic ones. In his preface, Jacob Grimm clearly distanced himself from Adelung, dismissing his dictionary, especially his definitions, as boring. For Grimm, people know the meaning of words intuitively, by themselves, and it is unnecessary to define them, except for special terms.

Wilhelm Grimm died in 1859, and when Jacob died in 1863 their *Deutsches Wörterbuch*, of which the first fascicle had appeared in 1852, was far from complete; it was at the entry *frucht* ('fruit'). Its completion took a long time, until 1961, and that time was filled with many intriguing episodes, but the dictionary never lost its exceptional prestige; it was funded by the government and administered by the Prussian Academy of Science and its successors.

It seems that a certain number of dictionary projects in the nineteenth and early twentieth century in Europe followed the model of the Grimm brothers to some extent, but they did not always simply copy it. Rather, they usually acquired their own profile by critically discussing what Jacob and Wilhelm Grimm did.

The outstanding characteristics and innovations of the *Deutsches Wörterbuch* over the whole period from the beginning of the plan (around 1837) until the publication of the last volume containing the references (1971) included the selection of entry words according to what is old and therefore 'good' German, excluding many, especially recent, loanwords; the provision of substantial academic etymological discussions (this was reduced in the later period of the compilation); the ordering of senses according to the date of first occurrence; and the introduction of a new and decisive historical orthography, since orthography was not standardized in Germany before 1904 (only a few of their proposals have survived in today's orthography). Citations and examples played a new role in the dictionary. While Adelung gave citations of authoritative writers in order to present examples of 'best practice' and to illustrate what was said more generally in definitions, the Grimm brothers wished to popularize many pieces of good, old poetry in order to convince the users of their beauty and make them proud of their language. Metaphors such as *Blumenmeer* ('a sea of flowers') were used by Jacob Grimm to illustrate the role of citations. The publisher was worried at first, because the estimated length of the work was quickly extended.

Finally, an important feature was the division of the work. The Grimm brothers, being professors, asked several colleagues to contribute to their national project by excerpting from relevant literary works. But not many did, and the brothers complained that the contributions of those who sent in

material were not of the quality they had expected. At the beginning of the project, the Grimm brothers did not have enough material, but they still rejected words and citations sent by people they did not know, such as Daniel Sanders (see below), who hoped to be allowed to take part in the national project. They insisted on controlling the selection criteria according to their idea of what the German language ought to be; accordingly, they decided to exclude modern authors – as well as all contemporary language material. Newspapers were unacceptable to them. After the brothers had died, the publisher paid students to sort out the citation slips and engaged a group of editors to continue working; these editors had finished their studies but were not professors. Later they were paid by the Prussian Academy of Science. During the dictionary's history there were different models of organization – with and without central editorial control. Only some of the problems caused by this can be mentioned here. From a modern point of view, the Grimms' project lacked overall organization and management. That is hardly surprising, as it was the first dictionary to be compiled by more than one person. The rules for lexicographical practice were not written down before the 1930s, when the work was charged to a team of editors instructed by one chief person.

The innovations of the Grimm brothers were accompanied by remarkable conservatism. For example, information on meaning was very brief, only Latin equivalents being provided, almost like the dictionaries published 300 years before. Information about grammar, word formation, and usage was given only if it was of historical relevance. Sometimes Jacob Grimm gave prescriptive comments; for example he criticized the famous poet Schiller for some of the expressions he used.⁷ Because of the uncontrolled growth of the dictionary, systematic cross-references between the numerous elements of a word family were missing, and the word families themselves were not even completely covered by the entries.

The Grimms' lexicography predominated in the nineteenth century, so that competitors and alternative concepts of what the German nation's dictionary ought to be are hard to identify. Daniel Sanders, a teacher belonging to the Jewish minority, had clearly different lexicographical ideas and eventually put them into practice with his three-volume *Wörterbuch der deutschen Sprache* (1859–65).⁸ When the excerpts he had offered to Jacob Grimm as a young man were rejected, he published a detailed and polemical

⁷ Haß-Zumkehr, 'Alle welt erwartet', 11.

⁸ Cf. Haß-Zumkehr, *Deutsche Wörterbücher*, 143–61; Haß, 'Daniel Sanders'.

critique of the Grimms' lexicographical method and published his own plan which was carried out when it was accepted by a publisher. The main characteristics of Sanders' alternative idea of how a national dictionary of German should be were as follows. He called for a focus on usage instead of history, updating Adelung's ideas of contemporary German language use in the age of democratization and industrialization. Usage was to be established according to the majority of serious writers and newspapers; this was the first time that newspapers were used in the material corpus of a dictionary and for citations. The forms and meanings of a word were to be described instead of prescribing a certain norm; he invented a sort of usage notes, more or less integrated in information on meaning. Semantics was to be re-installed by precise definitions and only describing what was thought to be 'good usage'. For Sanders, 'good' meant clear, intelligible and in line with modern authorities – historical explanations and etymologies were given only when they supported the user's comprehension; a word's value based on heritage was no longer of relevance. Common loanwords were to be included; no foreign words were excluded because of their origin. Alphabetical order was combined systematically with principles regarding German word formation, so that main entries were simple morphemes followed by derivations and compositions using the morpheme as a base. For example, the lemma *Zug* contains the sublemmata *Anzug*, *Bezug*, *Entzug*, and so on to *Verzug* and *Zuzug*. Orthography was to be established according to good and common usage. Besides the Grimms' dictionary, Sanders referred to Emile Littré, the French lexicographer, and to Franz Passow, who between 1812 and 1823 published his ideas about the way a dictionary of ancient Greek should be planned. A few years before Sanders died, his dictionary became the basis of the largest bidirectional dictionary of German and English ever, the *Muret-Sanders* (published by Langenscheidt), which is still in print today, and has been available online as part of the *Langenscheidt Online-Wörterbuch* since 2014.⁹

After the foundation of the Deutsches Reich, or unified German nation, in 1871, the government wished to set up standards for weights and measures, and also for orthography, because there were almost as many different systems as there had been small independent sovereign states before. So, during the second half of the nineteenth century, many spelling dictionaries were published, providing answers to the many questions asked by people who wanted to write well. Prussia's domination in cultural and educational

⁹ See Langenscheidt Digital, 'Der Große Muret-Sanders jetzt online!'

policy in the Deutsches Reich made it possible for a single standard of spelling to be imposed, and this was done in the teacher Konrad Duden's orthographic dictionary, which became the official standard after a conference on orthography in 1904. This was the birth of the Duden brand that was maintained by various publishing houses and is known today as a central authority concerning any question about the German language – although Duden is now the name of just a small part of one company.

German Lexicography in the Twentieth and Twenty-First Centuries

Around the year 1900 the situation of German-language lexicography was characterized by efforts to carry on the big project of the *Deutsches Wörterbuch*. But because it was still incomplete and too expensive for normal users, a certain number of smaller, usage-oriented, and more or less historical dictionaries of different types, like those of Daniel Sanders, succeeded on the market. They have all been marginalized in the research on the history of lexicography, and no sound statements about them can be made.¹⁰ The public was mainly interested in orthography rather than in etymology – many people, even today, find etymological statements puzzling, and even confuse them with statements about the current meanings of words – and it was interested in bi- or multilingual dictionaries for practical purposes linking German with other modern languages.

One outstanding dictionary, the *Deutsches Wörterbuch* (1897), written by the philologist and linguist Hermann Paul, must be mentioned. Paul developed a highly sophisticated scientific concept of lexicography that was published before the dictionary itself was realized. He found his sources within the tradition of German lexicography, accepting and adapting some of its practices and refusing others. He took the best from Adelung and Sanders on the one hand and from Grimm on the other. He combined rational definitions with historical explanations, and he integrated recent results of linguistic inquiry about semantic change. He took part in an international academic discourse.¹¹

After the First World War and the economic crisis of the 1920s, the production of dictionaries stagnated or was even reduced. The social class that had supported a flourishing lexicographical activity, the *Bildungsbürgertum*, declined. The Nazi regime came up with fundamental changes in the

¹⁰ Haß-Zumkehr, *Deutsche Wörterbücher*, 162–82. ¹¹ For details, see Kilian, 'Hermann Paul'.

educational system, and parts of the intelligentsia emigrated. New, smallish, and cheap dictionaries as well as encyclopedias were compiled as instruments of propaganda.¹² After the Second World War, the political separation of Germany made the idea of *one* new dictionary of German difficult.

It was Wolfgang Steinitz of the Academy of Sciences in the German Democratic Republic who in 1952 proposed a plan to create a new dictionary of contemporary German to show that the national literary language belonged to the working class. The *Wörterbuch der deutschen Gegenwartssprache* (WDG) was to start from scratch, which meant assembling a corpus of citation slips instead of borrowing from older dictionaries. The realization of this project was taken over by two sisters, Ruth Klappenbach and Helene Malige-Klappenbach. For the first time in the German-speaking world, women were publicly recognized as lexicographers, although they had been involved as 'helpers' much earlier.¹³ The dictionary was published in six volumes between 1961 and 1977, with parallel revisions from the 1970s until 1985. Its innovations were the following. The language to be described was defined as modern and widely used, including well-known special words (for instance, scientific and technical words in contexts from Goethe and Schiller to actual newspapers), and of course it included loans and foreign words. Old-fashioned and dialectal words were omitted, and neologisms were included. The focus was on semantics rather than on (etymological) form. The dictionary was conceived with reference to linguistic structuralism in its Russian and Czech variant: that means that a word sense is understood as a bundle of semantic features and as part of a hierarchy of hyperonyms, (co)hyponyms, antonyms, and so on; a system of levels of style was adopted which was more differentiated and usage-oriented than the style labelling of earlier dictionaries. Politically relevant words were marked as belonging to the German Democratic Republic and 'socialism', or to the Federal Republic of Germany and 'capitalism'. The ever-present task of handling German word formation was carried out by means that were as systematic as they were user-friendly: the WDG clustered the relevant compounds at the end of the entry for the simplex, whether the entry word was on the left or on the right of the compound. It used a style that was a compromise between compact information and user-friendliness. Cited examples were carefully selected in order to illustrate, as Adelung and

¹² Haß-Zumkehr, 'Propagandainstrument Wörterbuch'.

¹³ The first trace dates from the eighteenth century; Luise Gottsched, the wife of the poet whose material Adelung used for his wordlists and preliminary work, is known as a female lexicographer.

Sanders had done before, but not Grimm. This dictionary is still incorporated, revised, and amended, within the *Digitales Wörterbuch der deutschen Sprache*.

The WDG was hard to buy in the Federal Republic of Germany, where the Duden publishing company also planned a new dictionary in six volumes which was to provide more than information on orthography. This, the six-volume *Duden: das große Wörterbuch der deutschen Sprache*, started to be published in 1976, when the WDG had just been completed. F. J. Hausmann has shown that the western dictionary was ‘scandalously dependent’ on its socialist model.¹⁴ This concerned mainly the definitions, the usage labels, and the style categories, for the material base was different. *Duden: das große Wörterbuch* gathered citations from more modern authors, including popular books, and was in consequence more representative of common speech slightly below the standard. In fact, it was a product of the era of the student revolution in 1968, when the language used in public life was becoming less formal. In later editions, this tendency was reduced in favour of more classical and literary language. In time, Duden published a one-volume edition that was often revised, *Duden: Deutsches Universalwörterbuch* (¹1983, ⁸2015), with increasing use of digital text corpora.

One competitor of the *Großes Wörterbuch* must be mentioned: the Bertelsmann publishing house and its dictionary called *Wahrig* from the name of its first lexicographer, Gerhard Wahrig.¹⁵ Although the *Brockhaus–Wahrig* dictionary was published in a single six-volume edition like Duden’s, it became well-known in the form of a one-volume edition (*Wahrig Deutsches Wörterbuch*) that was continuously revised and was the first dictionary of German to be produced by means of a computer, before 1980. At the end of the 1990s, the Wahrig dictionary programme assembled its own electronic text corpus. In Germany, corpora owned by publishing houses are not available to the public. The Wahrig programme focused its attention on information about inflection and grammar, which is appreciated especially by foreign-language learners. Its eighth edition, edited by Renate Wahrig-Burfeind, was published in 2010.¹⁶ Bertelsmann’s dictionary publishing division, Wissenmedia, closed in 2014.¹⁷

¹⁴ Hausmann, ‘Wörterbuch und Wahrheit’.

¹⁵ Wahrig, *Das große deutsche Wörterbuch*, had already appeared in 1966; in later editions, his name came to be presented as the first element of the title of the dictionary.

¹⁶ Wahrig-Burfeind, ‘Wörterbuch als Datenbank’; Krome, ‘Das Wörterbuchprogramm WAHRIG’.

¹⁷ Verlagsteam Wissenmedia, ‘Auf Wiedersehen’.

Concluding Remarks on German Lexicography Since 1700

Since the 1970s, ‘theory of lexicography’ and ‘metalexicography’ have been established as subdisciplines in German universities. This had a serious impact on ongoing and new dictionary projects; it contributed to their improvement and it fostered some innovations. But the digital age has prevailed and changed almost everything within lexicography. On the one hand, popular dictionaries in the Duden and Wahrig series are losing their share in the market drastically, although there are digital versions of the bestsellers in these series. But people get used to obtaining information at no charge, often ignoring the quality of such information. On the other hand, scientific institutes and academies in Germany are developing new and sometimes experimental forms of word information, tending to integrate numerous and diverse monolingual dictionaries into one internet portal such as the *Digitales Wörterbuch der deutschen Sprache* and *Online-Wortschatz-Informationssystem Deutsch (OWID)*. But the acceptance of these innovative portals by dictionary users is not yet evident. German is a language with a long past, which is being documented today for the sake of contemporary dictionary users – and of dictionary users in future generations.

Lexicography in the Scandinavian Countries

At the beginning of the eighteenth century, neither of the Scandinavian kingdoms – Denmark-Norway, which was at the time one state, and Sweden, which at times ruled other Baltic regions – were at all monolingual. Instead, many bi- and multilingual constellations throughout regions and throughout classes were usual: Low and High German, even Latin, and Swedish dialects, as well as Finnish, Norwegian, Icelandic, and Faroese were used, and both Danish and Swedish took on many loanwords. Even French and Italian were imported by young Swedish noblemen when studying in southern Europe. Appropriate to this situation were numerous, rather small bi- or multilingual glossaries and dictionaries. Further, from the end of the seventeenth century and throughout the eighteenth, a debate on language standardization took place in Sweden as in Denmark, focusing on orthography, spelling, and inflection rather than on words.¹⁸

¹⁸ Teleman, ‘Role of language cultivators and grammarians’, 1384–94.

By the end of the eighteenth century, the representatives of a Scandinavian Enlightenment aimed for 'a cultivated, concise and clear form of language'.¹⁹ Church and state wanted everybody to be able to read. The plain language which supported literacy was also a goal of Romantic nationalists, who were opposed to foreign, and especially French, impact on their respective languages. A plain national language was called for by the middle of the eighteenth century in Denmark and by the beginning of the nineteenth century in Sweden. In both countries within the same period (around the middle of the eighteenth century), academies of science were founded. One of their objectives was the cultivation of language and literature. The proper means to achieve this goal was thought to be the creation of a new and great dictionary. But compared to the German situation, where one finds lexicography of Enlightenment and lexicography of Romanticism as separate concepts, these concepts more or less merged in Sweden as in Denmark while the national monolingual dictionary of each country was established. In the nineteenth and twentieth centuries there was a movement known as pan-Scandinavianism which focused on the common language origin of all Nordic languages, but the movement and its goals remained within intellectual circles, though it addressed some orthographic and terminological questions.²⁰ It is hard to say whether the pan-Scandinavian language policy has made a strong impact on lexicography. In the 1970s, the Nordic Language Council and Secretariat was founded, aiming at more and stronger inter-Scandinavian language learning; it has achieved, among other things, the production of a Norwegian-Swedish dictionary in the 1990s and wordlists dealing with the smaller Nordic languages, and even bringing together the non-Indo-European languages Finnish and Estonian.²¹ Of course, these activities have taken place against a background of discussion about the Anglo-American language influence on the Scandinavian languages.

Numerous bilingual wordlists have been produced for special purposes, and there are wordlists for learners and for particular kinds of words such as neologisms.²² There are also several large and more or less historical dictionary projects which are at their final stages or on their way into internet versions. The most important of these are characterized as follows.

¹⁹ Olsson, 'Historical and sociocultural preconditions', 1241.

²⁰ Venås, 'General tendencies', 2014–15. ²¹ Venås, 'General tendencies', 2021.

²² For more details, see Malmgren, 'Scandinavian languages', 707–10.

Danish

The Kingdom of Denmark established early and strong lexicographic efforts. Three great monolingual dictionaries are to be selected here from many other wordlists. First, the *Dansk Ordbog, udgiven under Videnskabernes Selskabs Bestyrelse* ('Danish dictionary published under the auspices of the Royal Danish Academy of Science and Letters'). It started collecting material in 1693 and was published from 1793 to 1905 in eight volumes. Its history has been called 'depressing', because it never managed to catch up with contemporary language or users' needs.²³ But it is an impressive example of merging a rational perspective of 'enlightened absolutism' with philology by the historical principles of the nineteenth century.²⁴ It was meant to be the national dictionary for Danish, but failed in this.

The *Ordbog over det Danske Sprog* (ODS) was published in a paper version of twenty-eight volumes, over a rather short time span: 1919–56. A five-volume supplement was published subsequently. It is a non-normative and ahistorical documentation of the Danish lexicon from around 1700 until 1950 and has become quite popular, although it is a 'rather uneven work in several respects'.²⁵ But unevenness may be attributed to perhaps all of the typical nineteenth-century great national dictionaries. The ODS and its conceptual frame are strongly affected by its Swedish counterpart, the *Ordbok över Svenska språket utgiven av Svenska Akademien*, discussed below.²⁶

In the early 1990s, preparation of a new, monolingual defining dictionary of Danish in six volumes, *Den Danske Ordbog*, started, and it started at once with an appropriate corpus, an adequate search engine, and a recognition of real users' needs. Addressing practical needs results in a public debate about how normative or liberal a dictionary should be and what makes it 'useful' and for whom.²⁷ Public reflection on lexicography of this sort is rather a new thing (but compare the debate between Jacob Grimm and Daniel Sanders in the 1850s); its impact on lexicographical work is growing, as dictionaries find their way onto the internet.

Swedish

Several Swedish kings before and after 1700 fostered plans for creating great national dictionaries, generally without success.²⁸ There have been quite

²³ Hjorth, 'Danish lexicography', 1913.

²⁴ Bergenholtz and Pálfi, 'Videnskabernes Selskabs Ordbog'.

²⁵ Hjorth, 'Danish lexicography', 1915.

²⁶ Malmgren, 'Das schwedische Akademiewörterbuch', 305–13.

²⁷ Malmgren, 'Scandinavian languages', 706–7.

²⁸ Holm and Jonsson, 'Swedish lexicography', 1934–7.

a number of mono- and bilingual dictionaries with Swedish, covering technical terms as well as slang words and others, most of them being rather small and handsome. Documentation of the Old Swedish lexicon was necessary in order to understand old Swedish text material, and this took place before and after 1900.²⁹ The most important wordlist is *Ordbok över Svenska språget utgiven av Svenska Akademien* (SAOB), Dictionary of the Swedish Academy, with a projected forty volumes, which started in 1890, and may be finished in 2020.³⁰ This dictionary addresses linguists and teachers first, but has gained the interest of a broader public, too.³¹ It documents the Swedish lexicon from 1521 until today, and it is obviously the Swedish equivalent of the *Oxford English Dictionary*, the *Deutsches Wörterbuch*, and the Dutch *Woordenboek Der Nederlandsche Taal* (see below). The SAOB is very precise and detailed in distinguishing senses and gives a lot of information on meaning and semantic nuances, and it does so while treating contemporary usage as well as historical use over several centuries.³²

The SAOB is supposed to be among the earliest dictionaries to be based on a digital corpus of texts. Its digitized version started in the 1980s. As in many other countries and languages, a Swedish dictionary has been linked with an encyclopedic lexicon.³³

Norwegian

From the sixteenth century until 1814, Norway belonged to the Danish kingdom, where the written standard was Danish. Thus the Norwegian language existed in spoken dialects, and only dialect words were collected in rather small wordlists by authors mostly aiming at Danish official persons who did not understand the Norwegian vernacular. This kind of lexicography lasted until the middle of the nineteenth century and made its contribution to the written language in the Danish-Norwegian kingdom.³⁴

Once the country became independent from Denmark in 1814, Norwegian was the subject of language planning activities, which included lexicography.³⁵ Around 1850 a new Norwegian standard language based on rural dialects, Nynorsk, was created, and at the end of the nineteenth century

²⁹ Holm and Jonsson, 'Swedish lexicography', 1934.

³⁰ Malmgren, 'Das schwedische Akademiewörterbuch', 303.

³¹ Malmgren, 'Das schwedische Akademiewörterbuch', 302.

³² Holm and Jonsson, 'Swedish lexicography', 1941.

³³ Nationalencyklopedin, 'Uppslagsverket'. ³⁴ Gundersen, 'Norwegian lexicography', 1924.

³⁵ Venås, 'General tendencies', 2014.

it was given rights equal to the traditional Dano-Norwegian standard. Both varieties had to be cultivated, but later they came closer, for example by spelling reformations.³⁶

Independence combined with national Romanticism created fertile soil for lexicography. Probably the most important lexicographer and linguist of this period was Ivar Aasen, who published *Ordbog over det norske folkesprok* ('Dictionary of the Norwegian vernacular') in 1850 and *Norsk Ordbog met dansk forklaring* ('Norwegian dictionary with Danish explanations') in 1873. Bilingual dictionaries and those for special purposes and word classes focused more and more on Nynorsk.

Outstanding is the *Norsk Ordbok* ('Dictionary of Norwegian', i.e., Nynorsk), started in the 1930s and completed in 2016. As an ongoing project it has been digitized and transferred to the University of Bergen, which maintains the website and describes the project as follows:

Norsk Ordbok: Dictionary of Norwegian dialects and the written language Nynorsk gives an exhaustive account of the vocabulary of Norwegian dialects and Nynorsk (the dialect-derived variety of written Norwegian). Entries in *Norsk Ordbok* include information on word meanings, dialect forms, grammatical features and etymology.³⁷

As usual in modern lexicography, resources of citations are directly linked to the text within the corpus.

Faroese and Norn

The first Faroese dictionary was completed in 1773 but remained in manuscript until the twentieth century; a printed glossary of 1891 was followed by a twentieth-century dictionary tradition.³⁸ The relics of Norn, the Nordic variety once spoken on the Orkneys and Shetlands, were sampled by a wordlist at the end of the nineteenth century.³⁹

Icelandic

In the age of national Romanticism and as a result of searching for the common origin of all Nordic languages, Icelandic or, as a synonym, Old

³⁶ Gundersen, 'Norwegian lexicography', 1923.

³⁷ 'About Norsk Ordbok 2014'; see also Grønvik, 'Saga of Norsk Ordbok'.

³⁸ Considine, *Small Dictionaries*, 217–21; Barnes, 'Language cultivation', 2010.

³⁹ Pétursson, 'Inselnordische Lexikographie', 1930–1.

Norse became the model, as its characteristics have been studied by linguists from all over Scandinavia. Thus, the nineteenth century was the heyday of Icelandic lexicography. It started in 1814 with an Icelandic–Latin–Danish lexicon which demonstrated that Icelandic shared its concepts with the dominant languages of the tradition. After 1850 a wordlist of Old Icelandic poetry appeared and fostered the study of old mythological texts. Towards the end of the nineteenth century a great bilingual Icelandic–English dictionary was established and has been supplemented several times up to today; it is still the main reference work.⁴⁰ Recently, the central lexicographic institute of Iceland has been occupied with a complete documentation of the Icelandic literary tradition from its beginnings. Independent of this, there are electronic dictionaries bridging Icelandic with English and Danish, searchable through the website orðabók.is. A more scientific lexicographic project, the Icelandic–Scandinavian online dictionary *Islex*, combines Icelandic with five other Nordic languages.

Dutch

Since the nineteenth century, the Dutch language has been documented by bilingual as well as monolingual, by diachronic as well as synchronic wordlists, and by small scholarly dictionaries as well as one big scientific dictionary. Words of Dutch dialects have been sampled since the end of the eighteenth century.⁴¹ Bilingual lexicography with French has played a significant role since the late Middle Ages (see Chapter 14) because of the multilingual situation in the area now called the Benelux countries. But many more bilingual wordlists with Dutch appeared after 1700, exploding in quantity after 1800, and covering a wide variety of languages – Malay as well as English, Russian as well as Spanish and Yiddish, and so on.⁴² This was a result of the fact that the Netherlands has been a major trading nation since early modern times. Dutch bilingual lexicography developed to include Turkish and Polish in the 1970s, reflecting working migration. Since the 1990s, the Dutch government has aimed to foster bilingual lexicography for the use of linguistic minorities living in the Netherlands, for example with Sranan Tongo, because dictionaries in a multilingual and

⁴⁰ For details and bibliography, see Ottosson, 'Research in Icelandic language history', 97, 100.

⁴¹ Heestermans, 'Niederländische Lexikographie und Lexikographie des Afrikaans', 2018–19.

⁴² Osselton, 'Bilingual lexicography with Dutch'.

mobile society are as essential, as Willy Martin has put it, as road infrastructure.⁴³

Surprisingly, lexicography between Dutch and German started rather late, with Matthias Kramer's *Het koninglyk neder-hoog-duitsch en hoog-neder-duitsch Dictionnaire* (1719).⁴⁴ Together with its later editions this dictionary shows that Dutch and German were still near to intercomprehensibility, but yet were drifting apart as two languages of two nations. From 1700 on both were undergoing a standardization process which led to communication barriers lexicography aimed to overcome.

Beside the communication-oriented bi- and multilingual wordlists, forming a vivid tradition over the years, Dutch took part in the older tradition, where the lexicography of any Germanic vernacular started with Latin and bilingual Latin-vernacular wordlists (see Chapters 13 and 14). This heritage accumulated some practical knowledge in lexicography which was common throughout Europe, and which became essential for the monolingual dictionaries of Dutch.⁴⁵

The *Woordenboek Der Nederlandsche Taal* (WNT)

'The world's biggest dictionary' it was called when forty volumes had been completed in 1998, and one has to add the three supplementary volumes published in 2001. It is certainly the case that the WNT has more volumes than the *Oxford English Dictionary* and the *Deutsches Wörterbuch*, which are monuments of the same kind and with a similar significance for their language communities as the WNT has for Dutch. The WNT has 95,000 entries covering about 400,000 words. Starting with preparatory work in 1852 and with the first fascicle published in 1864, it took about 150 years to be completed; it has been available online since 2007, managed by the Instituut voor Nederlandse Lexicologie.⁴⁶

Regardless of the fact that the characteristics of a dictionary are bound to change if it is composed over decades, one may fix a portrait of the WNT as follows. It covers a time span from 1500 until 1976 and describes the present Dutch language diachronically. Like the *OED* and the *DWB*, the WNT originated from a pan-European movement, the Romantic notion of language and history; this was at its climax in the first half of the nineteenth

⁴³ Martin, 'New developments', 824–5.

⁴⁴ Osselton, 'Bilingual lexicography with Dutch', 3037.

⁴⁵ Haß, 'In search of the European dimension'.

⁴⁶ Eickmans, '*Woordenboek der Nederlandsche Taal*', 271.

century.⁴⁷ But each of these three dictionaries played a special role within its political context. Since 1830 there have been two countries – the Netherlands and Belgium – with regions of different languages (Dutch, French) and different confessions of the Christian religion. In this time of severe conflicts some people, particularly philologists, engaged in promoting the Flemish (Dutch) language in different ways. The most ambitious of these was to call for a dictionary.

The *WNT* had to overcome a lot of problems. Though there was a plan laid down and discussed before work started, the first volume took too long a time and treated too few words to be a guide for the rest of the alphabet. The vast number of compound words, which is a feature of Dutch as of German, needs regulation, as do archaisms, loanwords, and dialect words. One needs concise guidelines for such questions and for an institutional and financial frame to be successful in writing a voluminous dictionary. It took some time to solve such problems.

Despite the fact that the *WNT* and the *DWB* originated from the same ideology of linguistic Romanticism and that Matthias de Vries, head of the *WNT*, was full of admiration for Jacob Grimm, the great European philologist of his time, the *WNT* is not a kind of copy of the *DWB*. De Vries departed from the Grimms' example in four ways. The definitions of the *WNT* are given in Dutch (instead of Latin) and are carefully made. Etymological information is reduced to well-established basics (instead of discussions or even speculation) and is thus much shorter than in the *DWB*. The *WNT* is unique in giving information about the semantic interrelations of words such as synonyms and antonyms; to do so was not at all easy before computer-aided storage of lexical data. Finally, the *WNT* presented the language of literature as the guideline for any other variety of Dutch; it wanted public language use to follow literature.⁴⁸ In that perspective the *WNT* is normative within description. The lexicographers even set up spelling rules for Dutch in 1863, which was for some time the official spelling of Dutch. But when spelling changed in the twentieth century, the *WNT* kept its way of spelling; nowadays one ought to write *Nederlandse* instead of *Nederlandsche Taal* as in the name of the dictionary.⁴⁹ In contrast to the *DWB*, the *WNT* is nearer to present-day language, explaining the change of words through history without being a 'museum' for the old language. The *WNT* eliminates the need for another multi-volume dictionary of the standard

⁴⁷ Haß, 'In search of the European dimension'.

⁴⁸ Eickmans, 'Woordenboek der Nederlandsche Taal', 278.

⁴⁹ Eickmans, 'Woordenboek der Nederlandsche Taal', 274.

language, although there exist some smaller dictionaries of present-day usage with slightly different focuses.⁵⁰

The newer Dutch lexicography beyond the *WNT* has three noteworthy features. More encyclopedic information is added; more attention is paid to collocations of words; and more attention is paid to semantic relations between words.⁵¹ Additionally, as one might expect, the digitization of dictionaries and the building of digital infrastructure and research for better user-friendliness of lexicographic resources are in progress.⁵²

Afrikaans

The lexicography of Afrikaans shared some features with that of many other languages on the African continent (see Chapter 31). First, there were bilingual wordlists written by foreigners, for instance colonial administrators, who needed them for communication with the people. Secondly, the language was described by indigenous speakers for their own community, giving lexicography a different role.⁵³

The first bilingual wordlist of Dutch with Afrikaans, A. N. E. Changuion's *Proeve van Kaapsch Taaleigen* ('Specimen of the language of the [South African] Cape'), dates from 1844 and was meant as a warning, as if to say 'Beware of these disgusting words used by uncivilized people.' But shortly afterwards, Afrikaans-speaking people became proud of their language and developed a new attitude. Afrikaans was no longer a depreciated variety of Dutch, but acknowledged as a language of its own, Cape Dutch, of which people in the Netherlands should be aware. The *Patriot woordenboek/Patriot Dictionary* (1902–4), was the first bilingual dictionary written by one of the language community's insiders, though it linked Afrikaans to English, which became important as a result of the Anglo-Boer war (1899–1902). It is an example of the use of lexicography as an instrument to promote social harmony between linguistic groups. But there is sometimes a direct development from the standardization process of a language to promoting the identity of its speakers and finally to language and political nationalism, which is the opposite of social harmony.⁵⁴

Between 1914 and 1925, Afrikaans became one of the official languages of South Africa used in schools, administration, and church. In 1926, work on

⁵⁰ Heestermans, 'Niederländische Lexikographie und Lexikographie des Afrikaans', 2012–15.

⁵¹ Martin, 'New developments', 817–19. ⁵² Martin, 'New developments', 820–3.

⁵³ Gouws, 'Aspects of Afrikaans lexicography', 827–8.

⁵⁴ Gouws, 'Aspects of Afrikaans lexicography', 830.

the official Afrikaans dictionary, *Woordeboek van die Afrikaanse taal*, started, but published volumes only started to appear in the 1950s, because the language itself went through a process of standardization during that period. Wordlists for translators and for special purposes resulted, literature in Afrikaans and a Bible translation came into existence – all these sources and their use of the language were to be considered by the lexicographers of the standard dictionary of Afrikaans.⁵⁵

Since 1994, Afrikaans has been one of eleven official languages of the Republic of South Africa: along with English there are four Nguni and three Sotho-Tswana languages as well as Tsonga and Venda. The fairly developed lexicography of Afrikaans, itself being influenced by Dutch, English, and German lexicographic practice and insights, has become important for establishing lexicography for the African languages named above, some of which have never been documented by a wordlist.⁵⁶ Thus, the case of Afrikaans is a good example to illustrate the strong interrelations between language and politics tied up within lexicography.

Frisian

Sampling of older words of one of the numerous Frisian dialects started with the Romanticism of the nineteenth century.⁵⁷ Many enthusiasts produced wordlists in a more or less professional way. Outstanding, and to some extent the Frisian equivalent to the Grimms' dictionary, was the *Lexicon Frisicum* by J. H. Halbertsma, published by his son in 1872.⁵⁸

Financial support was hard to find due to the fragmented situation of the language, which is spoken in an area which extends not only over two countries, but also includes many offshore islands. From the twentieth century onwards, there have been two institutes for the Frisian language in the Netherlands (one in Groningen and one in Leeuwarden) and one in Germany (at the University of Kiel). All of them have tried to achieve some bigger Frisian dictionaries. The modern, scholarly dictionary of the West Frisian variety from 1800 to 1975, with Dutch as a metalanguage, is the *Wurdboek fan de Fryske Taal*, a project of the Fryske Akademy, published from 1984 onwards.⁵⁹ The variety of Northern Frisian is characterized by

⁵⁵ Heestermans, 'Niederländische Lexikographie und Lexikographie des Afrikaans', 2019.

⁵⁶ Gouws, 'Aspects of Afrikaans lexicography', 833.

⁵⁷ Århammer, 'Friesische Lexikographie'.

⁵⁸ Dykstra, 'Lexicography of Modern West Frisian', 148–9.

⁵⁹ Dykstra, 'Lexicography of Modern West Frisian', 150–2.

relatively strong regional dialects; thus different wordlists concerning the dialects of a number of more or less well-known islands such as Helgoland, Sylt, and Föhr have been elaborated. From the beginning of the twentieth century, some efforts were made to establish an overall dictionary of the Northern Frisian language.⁶⁰ But, because of the dominance of Northern Frisian regional varieties, the task of documentation is nowadays shifting towards a database.⁶¹

From a global perspective, there are four challenges or functions the lexicography of Frisian has to fulfil: first, it aims to document the dialects somewhat like a museum would and keep them from getting lost. Secondly, it tries to help people, especially the many tourists visiting the regions and islands in the North Sea, to learn some words, and hence to learn about the way of life of the old Frisian people which was characterized by the sea, by fishing, and by catching birds. These wordlists are bi- or even multilingual, joining Frisian to Dutch, to German, to Danish, and to English. Thirdly, the historical lexicography of Frisian contributes to language comparison within the family of Germanic languages. Frisian is often overlooked by etymologists. Thus, etymology in general may have significant benefit from Frisian lexicography. The fourth function concerns the making of a Frisian standard variety which is hard to define because of the remarkable regional fragmentation of the Frisian language. But this is necessary to maintain the language at least within family and private communication. Dictionaries of Frisian standard language always try to incorporate words of modern life in order to enhance language development.

Yiddish

Bilingual wordlists with Yiddish have existed since the sixteenth century. They relate one of the different Yiddish varieties to one of numerous vernaculars surrounding the Yiddish community, and to Biblical Hebrew. Some of them deal with general and others with specialized vocabulary. Wordlists addressing special purposes contributed to an overall Yiddish standard which was and is of interest, because Yiddish-speakers in the Jewish diaspora might welcome a common variety. There never was anything like a national standard of Yiddish.

⁶⁰ Wilts, 'Nordfriesische Lexikographie', 353–63.

⁶¹ Wilts, 'Nordfriesische Lexikographie', 363.

Monolingual dictionaries of Yiddish have been published only since the 1950s, partly enriched by glosses in Modern Hebrew, English, or both. This work is still in progress. The central institution fostering Yiddish language, culture, and lexicography is the YIVO Institute for Jewish Research in New York, which holds more than one hundred Yiddish dictionaries in its library.⁶²

A special challenge for Yiddish lexicography when linked to other European languages is that Yiddish is written from right to left, but the equivalents in English, German, or Polish are written from left to right, and they are used in different writing systems. Thanks to advances in printing technology this problem could be solved since 1869; thus, equivalent information is situated within the same line. Before, only rough columns could be placed one beside another.⁶³

Luxemburgish

The case of Luxemburg, with its three languages French, German, and Luxemburgish, is a good example of the fact that lexicography plays its role in the identity of a nation no matter how small it is. This is true especially for the period of the late nineteenth and early twentieth centuries. Between 1924 and 1977, the *Luxemburger Wörterbuch* was published in four volumes plus one supplement. The introduction to the first volume lists all the preceding wordlists sampled by laymen and with the help of volunteers, and it shows the efforts to define a kind of standard Luxemburgish distinguished from several dialects. The commission needed six months to fix an overall spelling system, which did not yet exist.⁶⁴ Concerning the lexicon, the question of the inclusion of many French and German loanwords had to be answered. The *Luxemburger Wörterbuch* covers the language of literary and lexical books, especially of the twentieth century.⁶⁵ The commission decided to use German as a metalanguage, and thus the title of the whole work is German.⁶⁶ This shows that the dictionary wanted both to address scientific objectives and to be a reference book for the 'interessierte[n] Publikum' (interested public).⁶⁷

⁶² 'Yiddish dictionaries' is a list of those held by the YIVO library; for a characterization of some lexicographic milestones, see Fishman, 'Lexicography of Yiddish', 2249–51.

⁶³ Fishman, 'Lexicography of Yiddish'. ⁶⁴ *Luxemburger Wörterbuch*, I.xlv.

⁶⁵ *Luxemburger Wörterbuch*, I, unpaginated preface. ⁶⁶ *Luxemburger Wörterbuch*, I.xlv.

⁶⁷ *Luxemburger Wörterbuch*, I.xlv.

Standard Varieties of English from c. 1700

CHARLOTTE BREWER

This chapter sets out a history of the main developments in the lexicography of English as a standard language since around 1700.¹ In practice, this means focusing first on Samuel Johnson's *Dictionary of the English Language* (1755), enormously influential on English-language lexicography well into the nineteenth century, and secondly on the *Oxford English Dictionary* (*OED*), similarly influential on subsequent publications, whether in its first edition (*OED1*, 1884–1928), its main twentieth-century Supplement, or its gradually emerging reincarnation in revised form online (*OED3*, 2000–). Many dictionaries play an important constitutive or subsidiary role in the story, however, ranging from the dictionaries preceding Johnson, through those for English-language learners of the late twentieth and early twenty-first centuries (most notably *Cobuild*, 1987), to those of non-English varieties of the standard such as Scots, Canadian, Australian, and others. This chapter discusses dictionaries produced in the UK and (to a lesser extent) the United States, reflecting the pervasive historical and cultural influence of England and the UK on English lexicography to date.

Before we begin, it will be useful to reflect on the term 'standard' as applied to English, the composite term 'Standard English' generally designating the supra-regional dialect or dialects of English associated with educated written usage – a relatively small proportion of English usage overall. When Johnson wrote his dictionary there was no such term, or concept, but there was certainly a sense of polite and impolite language, correlating largely with wealth and class, and geographically associated with London as a political and cultural centre. Johnson thus designated some usages 'low', 'barbarous', 'Scots', and so on, to indicate that the vocabulary he predominantly recorded was none of these things. Since the *OED*'s revolutionary attempt to cover

¹ Two overviews of the subject are Béjoint, *Lexicography of English*, and Cowie, *Oxford History of English Lexicography*.

English inclusively and descriptively, and especially from the mid twentieth century onwards, the scope and range of vocabulary considered to be 'Standard English' have broadened. UK and US dictionaries have correspondingly expanded their wordlists, but have at the same time found it difficult to be consistent and dispassionate in decisions on labelling and/or defining contentious and culturally sensitive words and uses – owing partly to changes in cultural mores but also to popular (and occasionally lexicographic) associations of the standard with propriety and 'correctness', associations clearly detectable in Johnson's day and more or less continuously present ever since.² The notion of the standard has extended still further with the rise of English as a world language, with differing national and regional varieties of the language recorded in many different dictionaries around the globe. While linguists acknowledge that 'Standard Englishes' represent culturally prevailing dialects of English which coexist, in a variety of ways, with many others, the general public occasionally takes a different view, looking to dictionaries to tell them what is 'correct'. It remains a tricky and sometimes contested task, therefore, for individual lexicographers to tackle and determine the term's scope.³

Samuel Johnson

A monument to English cultural and literary heritage as well as to the English language, Johnson's *Dictionary* shaped retrospective understanding of that heritage and preserved it as a record for future generations. Chapter 14 sketched the beginnings of the monolingual English dictionary tradition in the seventeenth century; by the early eighteenth, these dictionaries had developed into larger works, covering a much more comprehensive range of vocabulary. Their definitions were often brief and unsatisfactory, however, and the other sorts of information we now routinely expect from dictionaries – etymologies, usage labels, information on pronunciation and spelling variants, consistent treatment of headwords and derivational forms, illustrative quotations or examples – were either absent altogether or inadequately and variably supplied.

In particular, none of these word books supplied the need, increasingly articulated from the late seventeenth century onwards, for a regulatory dictionary of the English language which would establish a standard of usage. A number of commentators had identified the desirability of such

² See Crowley, *Proper English*. ³ See Bex and Watts, *Standard English*.

a work, as a way of checking the flow of new words into the language and preserving those which were falling out of use – and a generally increased awareness of, and self-consciousness about, language usage was bound up with a host of social and historical factors. Increases in literacy had gone hand in hand with a vast expansion of material flowing from printing presses ever since the abundance of polemical pamphlets published during the Civil War, and discussions of political, religious, and philosophical matters were now routinely conducted in English rather than Latin. Major cultural shifts of a similar type were being experienced across Europe, and several countries had responded to resulting changes in their vernacular languages by creating academies, intended to regulate and codify usage, thus halting variation and the proliferation of new vocabulary (see Chapter 14). It was with an eye to these European developments that the English writers Defoe and Swift (along with many others, including Dryden, Addison, Pope, and members of the Royal Society) had pronounced on the need for regulation and codification of the language, whether of grammar and syntax (the eighteenth century is often called the age of grammars) or of vocabulary. Defoe, writing in 1698, thought the purpose of an English Academy (or ‘Society’, as he called it) ‘shou’d be to encourage Polite Learning, to polish and refine the *English Tongue*, and advance the so much neglected Faculty of Correct Language, to establish Purity and Propriety of Stile, and to purge it from all the Irregular Additions that Ignorance and Affectation have introduc’d’, while Swift in 1712 described how ‘our Language is extremely imperfect . . . its daily Improvements are by no means in proportion to its daily Corruptions; . . . the Pretenders to polish and refine it, have chiefly multiplied Abuses and Absurdities’.⁴

To many of Johnson’s readers, both at the time and subsequently, his dictionary appeared to answer these concerns – and certainly he himself began by conceiving of his task as adjudicatory and regulatory. His *Plan of a Dictionary of the English Language*, published in 1747, announced its exclusionary purpose from the outset, thereby breaking continuity with the tendency of his English-language predecessors to accrete vocabulary rather than apply discriminatory principles of selection. ‘It was not easy to determine by what rule of distinction the words of this dictionary were to be chosen’, Johnson wrote, explaining that, ‘The chief intent of it is to preserve the purity, and ascertain [i.e., “render certain, fix”] the meaning of our English idiom.’⁵ Even before the *Dictionary* itself was published, in 1755,

⁴ Considine, *Academy Dictionaries*, 100–6; texts in Bolton, *English Language*, 93, 107–8.

⁵ *Johnson on the English Language*, 29.

Johnson was gratefully hailed as a ‘dictator’ on usage – proscribing bad words and licensing good – and his reputation as *arbiter elegantiae* continued into the nineteenth century and beyond; it is reported that the historian Lord Macaulay preserved a copy on his desk, more than a hundred years after it was first published, ‘to keep his diction up to the classical standard, and to prevent himself from slipping into spurious modernisms’.⁶

Johnson himself, however, having written his *Dictionary*, changed his mind on the feasibility of preserving purity and ascertaining meaning, recognizing that no dictionary could ‘embalm . . . language, and secure it from corruption and decay’ and correspondingly that the lexicographer’s task was ‘not [to] form, but register the language . . . not [to] teach men how they should think, but relate how they have hitherto expressed their thoughts’.⁷ But it is not always easy to determine where his *Dictionary* sits between what we would now recognize as descriptive versus prescriptive lexicographical practices. He certainly had a didactic bent – ‘It is not enough that a dictionary delights the critic, unless at the same time it instructs the learner’, he wrote in his *Plan*; he was instinctively resistant to change (‘All change is of itself an evil’); and he supplied comments on usage far more extensively than his predecessors.⁸ But when he said that words were ‘low’ (e.g. *abominable*, or *adorer*, or *bang*, or *coax*) or ‘barbarous’ (the adverb *wondrous*), or commented that a form was ‘corruptly’, or ‘unskilfully’ used, was he proscribing these usages, or by contrast recording nuances of register widely recognized by his contemporaries? Evidently, the authority accorded Johnson’s dictionary by later generations of users had a tendency to transform codification into prescription, whether or not the latter was intended: dictionaries come to be regarded as prescriptive even when they explicitly set out to describe how language is, not determine what it should be. And Johnson himself was clear that his *Dictionary*’s purpose was to provide the same authority for the English language that European academy dictionaries had done for theirs: ‘I have devoted this book, the labour of years, to the honour of my country, that we may no longer yield the palm of philology to the nations of the continent.’⁹

Far more significant than Johnson’s usage comments, however, or the degree to which he was or was not a prescriptivist, was the transformation he wrought in English monolingual lexicography by basing his dictionary on quotations of actual historical usage. This fundamental change is the direct antecedent of our present-day use of corpora to construct dictionaries, the

⁶ Coleridge, ‘Observations’, 155. ⁷ *Johnson on the English Language*, 105, 102.

⁸ *Johnson on the English Language*, 29, 36. ⁹ *Johnson on the English Language*, 109.

central tenet of modern dictionary-making and one which is grounded in the recognition that a dictionary can only be as good as the evidence feeding into it. This momentous innovation was not a new practice in lexicography more generally: in this respect as in others, Johnson was following the example set by the European academy dictionaries earlier referred to, some of which were illustrated by quotations from canonical authors. The idea was to establish a model for good writers of the future, consistent with the long-standing and widely held belief that great writers have a unique role in the transmission of the national language and the maintenance of high standards of written discourse.

While we would now demur at this belief, it was Johnson's use of quotations from historical sources to illustrate the meanings of words that drove many of the improvements he was able to make on his predecessors. Where definitions were concerned, this is something aptly recognized by his greatest successor, James Murray, editor of the *Oxford English Dictionary* (OED), who explained that 'the special new feature' which Johnson's dictionary 'contributed to the evolution of the modern dictionary was the illustration of the use of each word by a selection of literary quotations, and the more delicate appreciation and discrimination of senses which this involved and rendered possible'.¹⁰ Looking at how each word he recorded had been used in a succession of real historical contexts opened Johnson's eyes to the variety and proliferation of shades of meaning that it could convey. One of the most visible results of this is the number of senses he was thus able to identify for individual words, and the consequent length of his entries – since naturally every sense was illustrated by one or more quotations. Phrasal verbs are an oft-quoted example here. Johnson found sixty-six different senses of the verb *take* used on its own, together with a further fifty-odd senses of the same verb combined with a preposition or used idiomatically. By comparison, Johnson's two main rivals, Nathan Bailey's *Dictionarium Britannicum* (1730) and Benjamin Martin's *Lingua Britannica reformata* (1749), had respectively found only three and seventeen senses altogether.

Johnson's attention to detail in discriminating the various possible senses of words, by analysing them in contextual examples of actual use, looks a familiar technique to us today: he appears to be basing definitions on an inductive approach to the linguistic data, in other words, on the quotations he has gathered as evidence. The significant question to ask, however, is how Johnson chose those quotations. Johnson's own account (and subsequent

¹⁰ Murray, *Evolution*, 38–9.

practice) makes it clear that he was guided in this respect by the desire not to register usage more generally but to select only (what he considered) the best. Acknowledging in his *Plan* that 'the credit of every part of this work must depend' on what he called 'the authorities' – the quotation sources – he explained that in choosing these 'it will be proper to observe some obvious rules'. Writers 'of the first reputation' were to be preferred to those 'of an inferior rank' and, wherever possible, Johnson would select quotations from their works which – besides their immediate use in illustrating the meaning of a word – might 'give pleasure or instruction by conveying some elegance of language, or some precept of prudence, or piety'.¹¹ In the event, this Horatian ideal – the notion that the highest aim of writing is to instruct, which is best achieved by pleasing the reader – was hard to keep to. As he later put it in his Preface, 'words must be sought where they are used; and in what pages, eminent for purity, can terms of manufacture or agriculture be found?'¹² Most remarkably, however, in his desire to draw only from 'the wells of English undefiled' – English before it had been corrupted by increasing French influences – Johnson chose his quotations where possible, he said, from the Elizabethan period up to 1660 (the return of the monarchy).¹³ This extraordinary aim – representing the English language with illustrative quotations a hundred years old and more – has been insufficiently remarked on, and points to a considerable difference between the *Dictionary* and the actual usage of the day.

Modern electronic analysis of the *Dictionary* sheds some interesting light on these claims. Just seven sources furnish nearly half the quotations in the 1755 edition of the *Dictionary*: Shakespeare, Dryden, Milton, Bacon, the Bible, Addison, and Pope, while most of the other writers quoted in significant numbers were born between 1552 and 1668.¹⁴ Despite his many outlying sources, therefore, Johnson overwhelmingly favoured a highly select and time-restricted canon. There can be no doubt that the scope and variety of the language in his *Dictionary* were correspondingly limited.

Johnson's use of quotations established the cultural importance of his book and marked a crucial stage forward in English-language lexicography; and in other ways, too, his work established lexicographical standards. Although he did not seek to establish the oldest use of a word, or to illustrate the historical development of senses, he nevertheless printed the quotations (albeit undated) in chronological order, and was thus able to represent the historical

¹¹ *Johnson on the English Language*, 55. ¹² *Johnson on the English Language*, 94.

¹³ *Johnson on the English Language*, 95. ¹⁴ Schreyer, 'Illustrations of authority', 62–8.

character of the language – and his use of the editorial label ‘obsolete’ where appropriate achieved the same result. Besides his Preface, which set out the function and purpose of the work and explained how to use it, he also provided a ‘History of the Language’ and a ‘Grammar’; consistently with the latter, each of his headwords was identified by grammatical category such as noun or adverb. He provided fewer headwords (c. 43,000) than Bailey’s (60,000), however, and Bailey’s dictionary along with others continued to sell and to be reissued during the later eighteenth century. Johnson’s brief provision of etymologies reflected the comparatively meagre state of knowledge of the time, which was to be transformed in the following century by far more thorough historical research, while for the first proper guide to pronunciation readers had to wait till William Kenrick’s *New Dictionary* of 1773, followed by William Perry’s dictionary in 1775, Thomas Sheridan’s in 1780, and especially John Walker’s in 1791; the latter’s dominance was such that some popular dictionaries of the early nineteenth century claimed to have been compiled with, for instance, ‘the Pronunciation on the basis of Walker’.¹⁵ There is no doubt that smaller dictionaries such as these, and such as the abridgements of Johnson which were published from 1756 onwards, were handier and easier to consult. Johnson’s two magnificent folios – each c. 18 cm long, 25.5 cm wide, and 9 cm thick – represented a larger investment, and the profusion of quotations on display pointed to its different cultural value: a monument to English literary history, summoning up the associations of language with nation and heritage, and the first dictionary to succeed in establishing a widely acknowledged standard of national usage.

After Johnson

Johnson’s two most important successors were Noah Webster (*American Dictionary of the English Language*, 1828) and Charles Richardson (*New Dictionary of the English Language*, 1836–7).¹⁶ Webster deserves special mention as the first lexicographer to publish a comprehensive standard of non-English English claiming equivalence with that of Britain. Already established as the author of a number of smaller dictionaries and wordbooks, Webster was determined to produce a wordbook reflecting the characteristic features and

¹⁵ This form of words appeared in editions of *Johnson’s English Dictionary in Miniature* from 1836 onwards; it was remembered at the end of the century by James Murray (*Evolution*, 43).

¹⁶ For others, see Simpson, ‘English lexicography after Johnson’.

independent status of the United States. Correspondingly his dictionary featured encyclopedic references to specific political, cultural, and geographical phenomena of the newly established country – as in entries for *congress*, *senate*, *tomahawk*, for example, or mention of the St Lawrence River and the Great Lakes of America s.v. *source*. In addition, though condemning Johnson's over-supply (as he saw it) of quotations, he borrowed from them copiously, and skilfully adapted this aspect of Johnson's *Dictionary* for his own cultural purposes, 'with pride and satisfaction' citing Franklin, Washington, and other distinguished American writers on the same page as English authorities such as Hooker, Milton, and Dryden.¹⁷ Subsequent abridged versions of Webster's work dominated US dictionary publishing over the next few decades, rival inheritors of the tradition vying with each other in a series of 'dictionary wars', with Merriam (the publishing firm which had secured rights to Webster's text) producing the first US dictionary with pictorial illustrations in 1859 and finally achieving dominance with the so-called *Webster's Unabridged* of 1864, which increased the 1828 vocabulary count from 70,000 to 114,000 words.¹⁸

As we have seen, it was Johnson's scrutiny of words in their contexts that had informed the detail and nuance of his definitions. Richardson by contrast subscribed to an entirely different theory of meaning, namely that true meaning properly resided in a single primitive etymological original. Correspondingly, he grouped derivative and variant forms of a word all together, under the same headword, simply listing his quotations underneath, without indicating how they instanced different senses of a word and/or different grammatical and syntactical uses. In this way he dispensed so far as possible with what he dismissed as Johnson's purely contextual definitions, regarding them as irrelevant and supererogatory. Benighted as this etymological theory was, Richardson was sufficiently convinced of the importance of historical research to extend the chronological range of his quotations back to the medieval period and forward to Byron. In this way, paradoxically, his dictionary contributed to the historical lexicographical method with which his etymological theory conflicted.¹⁹ And certainly he shared the view typical of major dictionaries of standard English, that his work was of worldwide significance, helping to perpetuate and extend the reach of both the language and its culture. His Preface closed with a vision of his dictionary finding 'a resting place upon the tables of an English Settler on the banks of La Plata',

¹⁷ Webster, *American Dictionary* (1828), 'Preface', sig. A2r.

¹⁸ Landau, 'Major American dictionaries', 195–200.

¹⁹ Zgusta, *Lexicography Then and Now*, 19–26.

relieving ‘the languor of military inaction at the Mess of Abednuggar’, and engaging ‘the acuteness of nearly a century of critics in the United States of North America’.²⁰

The *OED* and the Historical Turn

Well before Richardson’s *Dictionary* was published, linguistic studies in Europe were undergoing transformation. Philosophical speculation on language origin and theories of meaning was gradually being displaced by research into historical texts of European vernacular languages by Grimm, Bopp, Rask, and others; in turn this research illuminated links and relationships between these languages and pointed to their shared origin in an earlier Indo-European tongue. Parallel intellectual developments in the study of the natural and life sciences – culminating in Darwin’s *Origin of Species*, 1859 – were similarly rooted in painstaking empirical observation of surviving geological and biological phenomena such as sedimentary layers in rocks, fossils, or the characteristics of living organisms. The feature uniting all these disciplinary shifts was the enumeration and classification of actual physical evidence – whether words, rocks, fossils, or other life forms – in constructing histories of past eras and theories of origin. In their reliance on data rather than conjecture, these new disciplinary approaches were (as Murray later claimed of the *OED*) ‘permeated . . . through and through with the scientific method of the [nineteenth] century’.²¹

Where historical lexicography was concerned, ‘data’ took the form of excerpts from historical texts, and ‘scientific method’ required that data be sourced as widely as possible. Both Johnson and Richardson had gathered quotation evidence from past texts, but what the newly conceived historical dictionaries of the nineteenth century set out to supply was historical coverage, so that present-day vocabulary could be seen and explained in the context of changes and developments over time. This principle was first fully articulated by the German classicist Franz Passow in a passage written in 1825:

The dictionary should . . . set out . . . the life story of each single word in a conveniently ordered overview; it should state where and when each one was (as far as we know, of course) first hit upon, in which directions it developed . . . and finally, at what period it disappeared from use.²²

²⁰ Richardson, *New Dictionary of the English Language*, I.61. ²¹ Murray, *Evolution*, 49.

²² Quoted from Considine, ‘John Jamieson, Franz Passow’, 262.

Independently, a lexicographer of Scots had already hit upon the same idea: John Jamieson, whose *Etymological Dictionary of the Scottish Language* appeared in 1808, with a supplement in 1825. Other such dictionaries of European languages, based on quotations from actual usage, were soon embarked on elsewhere (if not so soon completed), notably in Germany (the Grimm brothers' *Deutsches Wörterbuch*, 1852–1961, with a bibliography in 1971: see Chapter 22); France (Littré's *Dictionnaire de la langue française*, 1863–72, with a supplement in 1877: see Chapter 25); and the Netherlands (*Woordenboek der Nederlandsche Taal*, 1864–1998, with a supplement in 2001: see Chapter 22).²³ This was the intellectual context in which the *OED* was conceived, and its editors sought to improve decisively on past lexicographical practice in English by gathering quotations much more widely and thoroughly than before, not (like Johnson) for purposes of pleasure and instruction but to create the evidential basis for a comprehensive history of vocabulary from 1100 to the present day. The importance of this disciplinary shift was fully recognized by the lexicographers themselves. Reflecting on their achievement in the Preface to the reissue of the first edition of the *OED* in 1933, the editors described its 'basis' as 'a collection of some five millions of excerpts from English literature of every period', forming 'the only possible foundation for the historical treatment of every word and idiom which is the *raison d'être* of the work. It is a fact everywhere recognized that the consistent pursuit of this method has worked a revolution in the art of lexicography.'²⁴ This ambitious project was without precedent in English and continues, in its present-day manifestation, to have a claim to be the single most extensive dictionary project of its kind, not least owing to 'the sheer length of its continuous documentation from the earliest records . . . down to the very latest'.²⁵

The effort involved in assembling then processing these vast quantities of quotation evidence was considerable, not to say heroic. Years later, reviewing the second edition, the Nobel laureate and novelist William Golding described how, 'in the high days of Queen Victoria a dictionary was conceived, not to say dared, which matched her iron bridges, her vast ships and engines'.²⁶ The comparison is apt: the construction of the *OED* was not unlike a magnificent and massive feat of engineering, contributed to by many different hands but formed by the organizing intelligence of a small number of men, themselves driven by a combination of brilliant vision, minute

²³ For these, see Osselton, 'Murray and his European counterparts'. ²⁴ *OED*1 (1933), i.v.

²⁵ Osselton, 'Murray and his European counterparts', 73.

²⁶ Quoted in Brewer, *Treasure-House*, 229.

attention to detail, and sheer doggedness. First proposed in 1859 (the same year as the publication of Darwin's great work), the dictionary did not appear in print until 1884, working through the alphabet instalment by instalment until completion forty-four years later in 1928. The finished work provided full information on the etymology, spelling, meaning, and evidence of use of words and senses from first occurrence to last, with editorial notes and also indication of pronunciation (based on Murray's own system).²⁷

Along the way many setbacks were encountered and overcome. For example, extending the range of quotations back to the medieval period in any consistent way was made possible only by the founding, by one of the early editors, of the Early English Text Society in 1864, which made available a gradually increasing stream of editions of works otherwise inaccessibly preserved in hard-to-read manuscript form. For excerpting of quotations both from these texts and from all others, the editors were crucially reliant on an army of volunteers, rustled up by a succession of public appeals, who needed to be trained in some minimal but adequate way, and whose contributions needed to be harvested and checked by teams of more experienced and qualified editors and subeditors. Over the years many contributors and editors died or moved on, which made continuity and consistency in editorial practice virtually impossible. It was not until the late 1870s that the dictionary achieved institutional and editorial stability, when the project (originally conceived and managed by members of the London Philological Society) was taken over by Oxford University Press (OUP) and the chief editor, James Murray, appointed. From then onwards, however, the lexicographers – principally Murray himself, but also his co-editors in chief Henry Bradley, William Craigie, and C. T. Onions – engaged in constant struggle with the publishers, the former determined to produce as full and perfect a dictionary as possible, and the latter seeking to bring the work to conclusion, so as to minimize the burdens of salary and administration. Contemplating what he regarded as 'the natural dilatoriness of lexicographers', the OUP publisher Kenneth Sisam observed, 'It is the exception for any huge dictionary to be finished . . . Have you ever found a reason why a sane man should start on one of these enterprises unless he is comfortably paid and housed? Or why, if he is comfortably provided for, he should ever finish it?'²⁸

²⁷ Different features of the dictionary are analysed in contributions to Muggleston, *Lexicography and the OED*; for the history of its making, see Gilliver, *Making of the Oxford English Dictionary*, and the older biographical account by K. Murray, *Caught in the Web of Words*.

²⁸ Brewer, *Treasure-House*, 18.

Finished it was, however, and received deserved acclaim both from linguists and from the general public as a definitive account of the history of English vocabulary – and, by extension, of the history and culture of the nation too. The latter response was largely due, as with Johnson's dictionary, to its munificent display of quotations; reviewers have described this 'vast storehouse of the words and phrases that constitute the vocabulary of the English-speaking people' as 'a history of English speech and thought from its infancy to the present day', or 'a history of thought and civilization'.²⁹ This combination of comprehensiveness, manifest cultural value, size, and cost – to the editors and publishers rather than to the buying public, though the dictionary was never cheap to buy – meant that the *OED* has dominated English-language lexicography ever since.

The *OED*: Limitations and Room for Improvements

Here we should pause and ask ourselves how well this new dictionary achieved the lexicographical ideals which Johnson and his successors fell short of. As we have seen, Murray himself was clear that Johnson had established the value of close scrutiny of quotations – evidence of real usage – as a way of identifying and discriminating between a word's possible senses. There can be no doubt that his own dictionary equalled and exceeded Johnson's in this respect. As is evident on every page, and usefully described by the lexicographers themselves, the *OED* method of analysing evidence from quotations to distinguish and delineate semantic relationships and developments over time was consistently applied and extraordinarily productive, especially given the vast chronological range: describing this method, Murray went even further, claiming that the *OED* sought 'not merely to record every word that has been used in the language for the last 800 years, with its written form and signification, and the pronunciation of the current words, but to furnish a biography of each word'.³⁰

But complete inclusivity had been aimed at by no previous English dictionary, partly because it was an impossible ideal to realize – how could one ever be sure of including every word ever used in the language, and how might the resulting dictionary ever be completed? – and partly because to do so would have offended against both politeness and (in some instances) obscenity law. On the contrary, as necessary in assembling any linguistic corpus, the *OED* applied criteria for exclusion of data, and the editorial

²⁹ Brewer, *Treasure-House*, 249. ³⁰ Murray, *Evolution*, 47.

process then as now involved careful assessment of words to check that they qualified for entry. Infrequency of attestation was an obvious reason to omit a word, though ‘nonce words’ (i.e. one-off usages or *hapax legomena*) were much more likely to get in if used by well-known writers in the English literary tradition (as understood by the Victorian and early Edwardian lexicographers). And Murray and his editors excluded or only partially defined many words on grounds of decency (four-letter words such as *fuck* and *cunt*, and many others relating to sex and the body). The selectiveness and bias of *OED*₁ in this respect are insufficiently understood – and one of the long-overdue tasks of the twenty-first-century revision currently underway in Oxford has been to add words and senses from a wider range of English (including non-UK varieties) and rewrite the definitions and editorial labels we would now see as homophobic, racist, euphemistic, or in other ways representing social and cultural attitudes that are now outdated and/or offensive.³¹

Another lexicographical principle enunciated by the editors of *OED*₁ was that of descriptiveness, not prescriptiveness. The lexicographer, as R. C. Trench (one of the founders of the *OED*) put it in 1857, ‘is an historian of [the language], not a critic’; while years later William Craigie observed, ‘Some of our predecessors in the science of lexicography thought it was part of their duty to improve the English language. We have got beyond that stage, and consider that if it is to be improved it is not our business to do so, but record it as it was and as it is.’³² Nevertheless, the first edition – and also the main twentieth-century Supplement (see below) – included a number of judgements, expressions of opinion, and recommendations on language which were clearly prescriptive, not descriptive. The most obvious indication of this was the use of a special symbol, the paragraph mark (¶), to indicate what the editors judged to be ‘catachrestic and erroneous uses, confusions, and the like’, despite quoting evidence of usage that defied this judgement. Prescriptive attitudes can also be detected in a number of editorial comments and labels; for example, under the entry for *ps-*, Murray described the practice of dropping the initial *p* in pronouncing words of Greek origin such as *psyche*, *pseudonym*, *pneumonia* etc. as ‘irretrievably mutilated by popular use’ – a remarkable judgement for a descriptive lexicographer to make, amounting to a contradiction in terms.³³

³¹ Mugglestone, *Lost for Words*; Brewer, *Treasure-House*, 110–22; Brewer, ‘*OED* Online re-launched’.

³² Trench, *On Some Deficiencies*, 5; Craigie, ‘Response’, 26. ³³ Brewer, ‘Pronouncing the P’.

A further insufficiently understood limitation of *OED*₁ is its marked literary bias. Notwithstanding the use of a much wider range of sources than any preceding dictionary (works relating to commerce and trade, arts, skills, and crafts, as well as newspapers, periodicals, and pamphlets on a myriad topics, printed letters and diaries, and more), the *OED* is dominated by quotations from Shakespeare, the Bible, Walter Scott, Milton, Chaucer, Dryden, and Dickens, along with the many other male authors (often poets, not prose writers) regarded by most educated Victorians as constituting the literary canon. This was partly due to the continuing prevalence, in the nineteenth and early twentieth centuries as well as the eighteenth, of the assumption that great works of literature play a tutelary as well as exemplary role in the history, development, and usage of the language: that (as the American linguist and lexicographer W. D. Whitney put it in 1867), 'A great body of literary works of acknowledged merit and authority, in the midst of a people proud and fond of it, is an agent in the preservation and transmission of any tongue, the importance of which cannot easily be over-estimated.'³⁴ In this cultural context, it is unsurprising that *OED*₁ was often regarded and praised as a repository of the nation's great writers of the past, with this element seen as its keystone. Accordingly, the publishers' press release of 1928 announcing completion described it as 'a Dictionary not of our English, but of all English: the English of Chaucer, of the Bible, and of Shakespeare is unfolded in it with the same wealth of illustration as is devoted to the most modern authors'.³⁵

This now-anachronistic privileging of literary sources, in constructing a history of the language, is additionally explained by their ready availability to the non-specialist readers on whose quotation-gathering labour *OED*₁ was crucially reliant. Naturally, these individuals (educated volunteers with sufficient leisure to perform such a task) had readiest access to the books on their shelves or in their libraries, collections in turn reflecting the literary and cultural biases of the day – very different from the 'social documents', often unprinted, to which today's historical linguists routinely turn, such as judicial records, inventories, wills, diaries and journals, and letters. Welcome as this cornucopia of literary quotations was to its readers, it does not, by today's linguistic standards, constitute the best basis for representing the diachronic or synchronic use of words – especially given that a significant portion of *OED*₁'s literary quotations were for unique or eccentric usage. By contrast, modern lexicographers now seek to balance the range of sources when

³⁴ Whitney, *Language*, 23. ³⁵ Quoted in Brewer, *Treasure-House*, 126.

constructing a corpus, so that fiction and poetry, however selected, are counterpoised by a much larger quantity of non-literary works.

After the *OED*: *Webster's Third*

As with Johnson's dictionary, abridgements were more practicable for many users than large-scale works. The *Shorter Oxford English Dictionary*, first published in 1933, was the only true abridgement, a two-volume selection with quotations and other supporting material pruned so as to give 'a quintessence' of the original at a much lower price.³⁶ It was certainly a hit with the dictionary-buying public, selling 40,000 copies in two years to those with insufficiently long shelves, or deep pockets, for the thirteen-volume original. OUP addressed the more immediate needs of the non-scholarly English-speaking community with a series of smaller publications, beginning with the *Concise Oxford Dictionary* (1911), which condensed the most important elements of current English in *OED*₁ and supplied some of the gaps in coverage (such as words for aviation, which had proliferated over the First World War) already apparent. The *Concise* was another astonishingly successful publication, continuously in print since its first appearance, swiftly followed by the *Pocket Oxford Dictionary* and *Little Oxford Dictionary*. By the mid 1930s and beyond, these derivative Oxford dictionaries, regularly re-edited, were established as leaders in their respective fields where the UK market was concerned; in this way, the press recouped some of its vast outlay of costs on the *OED*, and kept its dictionary branding prominently in public view, while forestalling (or, at least, competing with) the predatory forays other contemporary dictionaries, published by Routledge, Longman, Merriam-Webster, and others, were making on the lexicographical treasures in the *OED* itself. Unlike the *OED* and the *Shorter Oxford English Dictionary*, these smaller dictionaries focused on the current senses of words, tended to include derivatives and other forms of a word all in one entry, and moved etymological information (suitably condensed) to the end of an entry. These measures all saved space, reducing the size and cost of the resulting volumes.

The enormous historical and social changes over the course of the twentieth century, especially after the Second World War, continued to bring with them corresponding changes in language usage. UK dictionaries were faster to react to this than US ones. In North America, Merriam-Webster's various publications, descending from Noah Webster's 1828 *American Dictionary of the*

³⁶ For 'quintessence', see Brewer, *Treasure-House*, 77.

English Language, had long been influential; indeed, Murray had taken the 1864 edition of *Webster's* (the 'Unabridged') as a model in determining the scale of *OED*'s coverage of individual stretches of the alphabet.³⁷ The 1864 edition had been replaced on a grander scale by *Webster's International Dictionary* in 1890, whose title made clear the increasingly global reach of the English language – and of the American version of English. *Webster's Third New International Dictionary of the English Language* (1961) marked a watershed in the lexicography of Standard English on both sides of the Atlantic.³⁸ Its editor, Philip Gove, knew from his extensive quotation collections that vast numbers of colloquialisms in regular standard usage were omitted from dictionary record, despite occurring in the same sorts of discourse – whether presidential speeches or respectable newspaper sources – from which preceding editions of *Webster's* had routinely quoted examples of unstigmatized words or senses. He included abundant quantities of such vocabulary in the *Third New International*, therefore – 'all the way from breezeway and split-level to fringe benefit and sit-in, from airlift and no-show to deceleration and astronaut, from beatnik and den mother to wage dividend and zen', as one review had it.³⁹ Perhaps most notorious was the treatment of *ain't*, described as 'used orally in most parts of the U.S. by many cultivated speakers'. Up and down the country, reviewers published outraged notices of this transgressive new edition, marketed as 'the greatest vocabulary explosion in history', but there was no going back lexicographically from such commonsense policy, based on the facts of usage – and incidentally entirely in keeping with the groundbreaking descriptive principles (if not consistent practice) of the *OED*, whose first founders had been clear that 'It is no task of the maker of [a dictionary] to select the *good* words of a language. If he fancies that it is so, and begins to pick and choose, to leave this and to take that, he will at once go astray.'⁴⁰

Gove had ushered in a new regime of openness, and other dictionaries swiftly followed suit. In the United States, this inclusiveness led to another valuable new development, namely much more careful and nuanced application of usage labels, as lexicographers, appalled by the *Webster's Third* controversy, struggled with the issue of what was and was not Standard English usage. By the 1970s and 1980s, American desk dictionaries were, on the whole,

³⁷ Gilliver, *Making of the Oxford English Dictionary*, 133.

³⁸ For the dictionary, and the controversy over it, see Morton, *Story of Webster's Third*.

³⁹ Pei, 'Ain't is in'.

⁴⁰ For 'greatest vocabulary explosion', see Pei, 'Ain't is in'; 'It is no task . . . ' is from Trench, *On Some Deficiencies*, 4.

more sophisticated than English ones in identifying the various sorts of offence that certain widely used locutions could cause, and valuably extended this awareness to issues of race, a particularly sensitive issue given recent changes in US legislation and increasing public awareness of the objectionable prejudices inherent in certain terms.

Supplementing the *OED*

The influence of *Webster's Third* was most strikingly felt in the UK through the unlikely conduit of the *OED* – or rather its main twentieth-century *Supplement*, an endeavour underway since 1957 on the appointment of a new editor, R. W. Burchfield.⁴¹ *OED*₁, completed in 1928 and followed by a short catch-up supplement in 1933, had by the 1950s been recognized by the publishers to be in significant danger of turning into a 'magnificent fossil' (in the later words of one of its staunchest twentieth-century volunteers, Marghanita Laski).⁴² Revision had been considered but rejected as too expensive, and OUP settled instead for a more substantial (but still one-volume) supplement which would incorporate its 1933 predecessor and in addition address the floods of new words and senses that had since entered the language. But on seeing *Webster's Third*, as Burchfield himself described, 'the sheer quantity of words included . . . made it apparent at once that I had seriously underestimated the task of collecting modern English vocabulary wherever it occurred. The whole editorial process had to be delayed – in the event by several years – until my editorial assistants and outside readers had assembled evidence on this majestic scale.'⁴³ As Burchfield expanded his reading lists to include a much wider range of newspapers, periodicals, journals, and other sources of informal English, his projected single volume swelled to four, and the inclusion of so many colloquialisms, slang usages, and examples of 'common language' marked a major change in editorial policy in the *OED*, reflected in dictionaries of English more widely. One notable category of newly included vocabulary – and more explicit treatment in general – was that relating to sex and the body; Burchfield was proud to be the first to record *fuck* and *cunt* in the *OED*, supplied with quotations and etymologies dating back centuries that had long been on file in the *OED* offices (the new climate of liberality in publishing enabled by the 'Not Guilty' verdict in the *Lady Chatterley's Lover* trial in 1960 had made it possible to

⁴¹ See further Brewer, *Treasure-House*, 152–229, on which this section draws.

⁴² Quotation from Brewer, *Treasure-House*, 257.

⁴³ Quotation from Brewer, *Treasure-House*, 168.

publish such material without fear of prosecution). Like his US counterparts, Burchfield was also forced to consider issues of labelling such vocabulary. The wider a dictionary cast its nets, the more necessary it became to indicate variations of register and social acceptability in the words included – and, again as in the United States, changes in cultural mores meant that usages offensive to minorities (e.g., derogatory senses of *Jew* and its derivatives), formerly included without notice in dictionaries, needed to be suitably identified and explained (many of these changes had already begun to be made in the more frequently updated smaller dictionaries of standard English published both by Oxford and its competitors).

More surprising is Burchfield's initiative on 'correctness'. From Volume 1 of the *OED Supplement* onwards he had reinstated the paragraph mark, dropped by the 1933 supplement, and in the Preface to Volume 3, published in 1982, he noted the recent prescriptivist backlash to the 'markedly [*sic*] linguistic descriptivism of the post-war years' and commented, 'One small legacy of these great debates is that here and there in the present volume I have found myself adding my own opinions about the acceptability of certain words or meanings in educated use.' Burchfield is thus the only *OED* editor to have consciously and explicitly flouted the descriptivist principles of the first edition. As is not infrequently the case with prescriptivists, he found it hard to be consistent; for example the *OED Supplement* entries for the words he instanced in his Preface as arousing dissenting voices, *ongoing*, *relevant*, and *viable*, do not in fact alert the reader to any issues of usage. And while he chastised usages which his quotation evidence revealed to be commonplace, such as the use of *media* as a singular noun, it is easy to find other unlabelled entries for words or senses in his supplement routinely disparaged by prescriptivists writing at the same time.

Two other main areas in which the *OED Supplement* expanded the range of the parent dictionary – scientific and technical vocabulary, and World Englishes – once again reflected the increasing scope of what was now recognized as Standard English. Here too consistency was difficult in covering such extensive and various fields. Burchfield himself drew attention to variations in documentation between the Supplement volumes: for instance, 'loanwords and loan translations' of Chinese origin, such as *pipa*, *putonghua*, *Little Red Book*, *running dog*, and *scorched earth*, increased threefold after his own visit to China in 1979 and the publication of Volume 3.

Where another category of vocabulary was concerned, however, namely literary language, Burchfield retreated to a markedly conservative linguistic position, insisting on the role of the *OED* as 'literary instrument', and

decrying descriptive linguists who turned to general, non-literary usage for evidence of how language worked. Burchfield's literary canon favoured pre-war writing and was both selective and unevenly represented, with large quantities of quotations – many for eccentric or unique usages – from Joyce, D. H. Lawrence, Auden, and other modernist writers.

While diverging from the lexicographical mainstream in this respect, Burchfield's respect for literary usage struck a chord with reviewers and users devoted to *OED* traditions and in particular to the association between language and high literary culture. Completed in 1986, the Supplement was soon afterwards (in 1989) typographically merged with *OED1*, to create a publication described by OUP as a 'second edition' – though virtually no changes had been made to the original dictionary, based on long-outdated scholarship, which constituted by far the greater part of the work (twelve volumes compared with Burchfield's four). *OED2* was greeted ecstatically as 'the living scriptures . . . [making] the whole past a common treasure', while OUP's own publicity material drew attention to the cultural value enshrined in its quotations from great writers, printing cameo pictures of some of the best known, including Shakespeare, Austen, and Dickens, and reprinting cultural accolades to the first edition (for instance as 'a history of English speech and thought from its infancy to the present day', quoted above) now sixty years out of date.⁴⁴

Corpus-Based Lexicography

Elsewhere, however, linguistics and lexicography were moving in a radically different direction. From the late 1950s onwards, linguists had begun to harness computers – then in their early stages of development – to compile language corpora, understanding that systematic analysis of continuous large tracts of language, both spoken and written, chosen from a balanced and representative range of sources, would shed light on patterns of language use and on the conditions in which linguistic variation occurred. The two pioneers, Randolph Quirk's *Survey of English Usage* in the UK and the Brown Corpus in the United States, spawned many successors and generated huge quantities of research on grammar, but the potential for dictionary-making was also soon spotted. *The American Heritage Dictionary* (1969), based on specially prepared citations from the Brown Corpus and related information on word frequencies, was the first to break the new ground – although

⁴⁴ Quotations from Brewer, *Treasure-House*, 217–18, 249.

ironically, given the descriptive ideology that underlies the making and use of corpora, it combined this material with often prescriptive comments provided by a usage panel. In the UK a much more revolutionary dictionary appeared, the *Collins COBUILD English Language Dictionary* (1987), based on the Collins Corpus, then around twenty million words. As its editor, John Sinclair, explained in the Introduction, 'For the first time, a dictionary has been compiled by the thorough examination of a representative group of English texts': books, magazines, pamphlets, leaflets, conversations, radio and television broadcasts (all fully listed), comprising 'a fair representation of contemporary English'. Consequently, 'in addition to all the tools of the conventional dictionary makers – wide reading and experience of English, other dictionaries and of course eyes and ears – this dictionary is based on hard, measurable evidence'. *Cobuild* constitutes the first major development in lexicography since Johnson and the *OED*, taking their reliance on examples of real language use several steps further, first by ensuring that these excerpts were representative of usage more generally, and secondly by harvesting the information that could be gathered by systematic examination of a word's contextual features – its typical collocations, syntax and grammar, and relative frequency.⁴⁵

The advantages of language corpora to lexicographers of current Standard English were self-evident: dictionary-makers could now discriminate between core and peripheral vocabulary, make consistent and evidence-based decisions on declining or increasing usage, and improve definitions and editorial labels. They could also more readily identify when all these needed updating, one of the trickiest jobs for regularly reissued dictionaries. Within a decade or so, all publishers of major English dictionaries had followed suit, many of them establishing in-house corpora feeding works of different sizes and for different markets.

Cobuild, conceived in the first place as a dictionary for language learners, had also introduced a radically different definitional style. Up to this point, dictionaries had generally offered synonyms, or definitions on the long-standing analytic (Aristotelian) model, identifying first the superordinate category of the *definiendum* (item being defined) and secondly the delimiting features distinguishing it from other members of that category (for instance, *square*: 'rectangle whose sides are of equal length') – the whole often rendered in hard-to-decode 'lexicographese'. Instead, *Cobuild* provided explanatory sentences informed by corpus-supplied evidence of typical collocations and

⁴⁵ For the *Cobuild* project, see Sinclair, *Looking Up*.

other contextual information, including syntax and grammar.⁴⁶ Its definition of *down-to-earth*, for example, read ‘Something or someone who is **down-to-earth** is concerned with doing practical things and solving problems in a practical way, rather than with abstract theories’, compared with the 1982 *Concise Oxford Dictionary*’s laconic ‘practical, realistic’. Although criticized for wordiness, *Cobuild*’s method wins hands-down in communicating meaning clearly and fully and in representing the grammatical context and register in which a word might typically be used. Definitions of this type have, however, been slow to be adopted – even in a more concise form – by dictionaries for native speakers, whose editors know (consciously or unconsciously) that they can rely on their users’ intuitions, rooted in long-standing experience of the language. They have been more enthusiastically absorbed in learners’ dictionaries, a lucrative market in which dictionary publishers hotly compete, thus driving forward a number of lexicographical improvements and changes – whether in dictionary design or the more consistent and thoughtful application of usage labels – where dictionaries for native speakers have limped behind.

But despite the enormously enhanced resources now available to lexicographers – corpora amounting to many millions of words, and sophisticated tools for analysing their contents, enabling, one might have thought, roughly similar judgements to be made on which items to include and how to treat them – dictionaries of Standard English continue to vary widely. Some include proper nouns and/or more encyclopedic information than others (a traditional practice in US dictionaries), some more obsolete or unusual words (the *Collins* and *Chambers* dictionaries, useful to crossword and Scrabble enthusiasts respectively), some more colloquial usage. And a survey of contemporary works, even those covering the same regional variety of the standard (UK, in the following examples) readily turns up differing views on the use of words in context, particularly noticeable on matters such as sexism or correctness. The *Concise Oxford Dictionary* (2011), the *Collins Concise English Dictionary* (2008), the *Collins Cobuild Learner’s Dictionary, Concise Edition* (2003), the *Shorter Oxford English Dictionary* (2007), and the *Oxford Dictionary of English* (2005), for example, variously disagree on whether and how to identify the potential offensiveness of the terms *chick* to mean ‘attractive young woman’ and *girl* to mean ‘woman’. The same works inconsistently identify why and how certain terms offend traditionalists, for example *refute* to mean ‘deny’, *decimate* to mean ‘destroy a large part of’, or *disinterested* to mean ‘bored’.

⁴⁶ See Hanks, ‘Definitions and explanations’.

Such variation indicates that subjective judgement continues to play a role in determining issues of usage: notwithstanding the ‘hard, measurable evidence’ derived from corpora, it remains a tricky matter to decide when an apparently non-standard form reaches a sufficient level of widespread attestation to merit being acceptable as part of the standard variety. Occasionally the lexicographical struggle is visible to the user, as when the 1998 *New Oxford Dictionary of English* condemned *infer* to mean ‘imply’ as a straightforward ‘error’, at the same time acknowledging that it was ‘common enough for some dictionaries to record it as a more or less standard use’.

In this respect the *OED* has an interesting advantage, since it records usage over time. Many of the lexical shibboleths to which readers are alerted in contemporary dictionaries turn out to be without historical authority: this sense of *infer*, as *OED1*’s entry told us in 1900, has been continuously attested in perfectly respectable sources since the early sixteenth century. Even with the benefit of this historical perspective, however, it can sometimes seem that Oxford dictionaries cling to a popular notion of right and wrong which other publishing houses, and certainly linguists, have moved on from. There is a thought-provoking discrepancy between the definitions of ‘Standard English’ on the *Oxford Dictionaries* and *Merriam-Webster* websites at the time of writing (July 2017). *Oxford Dictionaries* speaks of a single form and associates this with ‘correctness’: ‘The form of the English language widely accepted as the usual correct form’; *Merriam-Webster* recognizes variations both of region and of register, invokes ‘usage’, and substitutes ‘acceptable’ for ‘correct’: ‘English that . . . is substantially uniform though not devoid of regional differences . . . well established by usage in the formal and informal speech and writing of the educated . . . widely recognized as acceptable wherever English is spoken and understood’.

*OED*₃

Apart from *Cobuild* itself, the *OED* has been boldest and most prescient among English dictionaries in grasping the opportunities afforded by digitization. Merging Burchfield’s *OED Supplement* with *OED1* in 1989 to produce the so-called second edition turned out to be a byproduct of a much more ambitious undertaking, in which every constituent element in every entry of the dictionary was electronically tagged. For the first time, users could search the *OED*’s phenomenally rich content by any category they chose – date of first citation, etymology, quotation source, definitional content, editorial labelling, or anything else – thus escaping the limitations of the

printed copy. It is difficult to overstate the significance of this electronic revolution, which opened up virtually endless possibilities for lexical and other kinds of research. Digitization – or rather, the analyses of the *OED*'s entries it was now possible to conduct – also exposed the faultlines in construction identified above: inconsistencies in excluding or including vocabulary, cultural biases in the treatment of areas such as gender and race, occasional prescriptiveness prevailing over the evidence of usage, and preferences for some quotation sources over others. Similarly, however, this revolution – and its exploitation in all forms of information technology, such as the creation of databases of contemporary and historical texts – afforded radically improved methods of overhauling and recompiling the *OED* (as it had smaller dictionaries).

In the late 1990s, Oxford University Press took the long-overdue decision to revise the entire dictionary from scratch, establishing a large editorial team who since 2000 have been publishing their slowly emerging results – the third edition of the dictionary, *OED3* – in online instalments: revised stretches of existing entries on the one hand, and altogether new entries for recent vocabulary on the other. By the end of 2011, 102,133 entries (or 37 per cent of the dictionary entries then available) were new or fully revised.⁴⁷ Where new entries are concerned, access to OUP's own Oxford English Corpus (currently one of the largest such collections of twenty-first-century English in existence), and to work in progress on OUP's many published dictionaries (including those for World Englishes), should help ensure that *OED*'s updated representation of Standard English – a much more capacious entity than its Victorian equivalent – is evidentially based. Revisions of entries for older vocabulary are now researched from databases such as *Eighteenth Century Collections Online* and *Early English Books Online*, and many other non-literary historical digitized sources such as newspapers and judicial records, as well as from the wealth of historical linguistic scholarship published over the past hundred years. Resources have been steadily added since the inception of the online publication of the dictionary, including links to OUP's own *Dictionary of National Biography* and *Oxford Dictionaries* websites, the University of Michigan's *Middle English Dictionary*, and the University of Toronto's *Dictionary of Old English*. And the online format is equipped with an impressive battery of tools, permitting users to search by category of word (subject, usage, region, or origin), by dates or date ranges at which words are recorded as entering the language, or by quotation source.

⁴⁷ Simpson, '100,000 entries published'.

All this would appear to deliver on the *OED Online*'s strapline description of itself as 'The definitive record of the English language'. Unfortunately, this is far from true. The website seamlessly combines old with revised material, so that users are continually brought up short by entries incompatible with modern lexicographical standards, such as the definition of *slang*, unrevised as of July 2017, as 'The special vocabulary used by any set of persons of a low or disreputable character; language of a low and vulgar type'. Worse, assessing the new content in any systematic way is impossible, as is using it to conduct new research. Since searches do not differentiate between old and revised material, one can neither isolate the latter to examine its contents nor compare it with previous versions (themselves not electronically searchable). One of the most obvious ways to assess *OED3*'s presentation of Standard English, for example, would be to investigate its new quotations; indeed, the front page of the website encourages this, inviting one to click and 'Explore the top 1,000 authors and works quoted in the OED.' But the results – an amalgam of revised and unrevised material, in proportions one can only guess at – continue to be dominated by the (Victorian) canonical writers who similarly dominated the first edition, newly interspersed with newspapers and journals. Either the revision is according these literary sources the same distorted significance as before, or it has yet to amass sufficient quantities of new evidence to tip the balance – we cannot know, and the website does not tell us.

Examined individually, *OED3* entries certainly look promising: the cultural biases of *OED1* are in the process of being eliminated, so that racist or homophobic or other now-unacceptable definitions and labels are disappearing; citations from more culturally and geographically diverse sources appear to be on the up (e.g., from female and/or non-English writers), if not yet rivalling the 'great writers' already mentioned; antedatings of first-attested words and senses are occurring in their hundreds (probably thousands); and editorial notes of a variety of different kinds now supply more nuanced and descriptive information (for instance on the lexical shibboleths discussed above) than were found in preceding versions of the *OED*. A handful of provisional analyses, or accounts, of *OED3*'s new material have been published to date, but proper assessment of this central endeavour in English-language lexicography will have to wait until OUP improves the dictionary website – or until the revision is complete, in perhaps two or three decades.⁴⁸

⁴⁸ See, e.g., Simpson et al., 'The *Oxford English Dictionary* today'; Brewer, 'Shakespeare, word-coining, and the *OED*'; Brewer, '*OED Online* re-launched'; and Brewer, 'That reliance on the ordinary'.

Conclusion

Meanwhile, as increasingly evident in *OED3*'s new entries, varieties of Standard English around the world continue to develop and to be recorded lexicographically, with new research extending the historical record back decades and in some cases centuries.⁴⁹ OUP itself publishes (among others) *The New Zealand Oxford Dictionary* and *The Canadian Oxford Dictionary*, along with versions of the tenth edition of the *Concise Oxford Dictionary* which have been customized for different countries, such as the *South African Concise Oxford Dictionary*. Numerous other successful commercial dictionaries, such as Australia's *Macquarie Dictionary* and the long-established *Webster's*, flourish and thrive both online and in a variety of printed forms. Such works typically combine coverage of Standard English (i.e. a version sufficiently widespread to be regarded as international, not just current in the UK) with extensive entries for vocabulary items peculiar to or characteristic of the relevant national varieties.

As all these works compete for buyers, subscribers, and online traffic, and their websites have expanded to include blogs, videos, and articles on many different aspects of English, notably new words and senses, usage variation, and the history of words, while versions of well-known commercial dictionaries have (allegedly) been used to provide dictionaries for Apple's and Google's electronic devices and services. All this indicates a cheerfully buoyant public appetite for information on Standard English vocabulary. The future of dictionaries of these varieties of English, realized in whatever technological form, seems secure.

⁴⁹ Background in Trudgill and Hannah, *International English*; a case study is Salazar, 'Towards improved coverage of Southeast Asian Englishes'.

Regional Varieties of English

MICHAEL ADAMS

For centuries now, English has been the language of villages and nations throughout the world, and it varies, we say, according to something we call ‘region’, though we could hardly come up with a vaguer term. The most familiar dictionaries of English describe a standard variety, though different regions may develop different standards over time – *Webster’s Third New International Dictionary of the English Language* accounts for Standard American English, while *The Macquarie Dictionary* does the same for Standard Australian English, and so on (for both, see Chapter 23). Some dictionaries of regional English represent a nation’s distinctive lexical features, while others represent local features, with dictionaries of every imaginable scope in between those extremes. ‘No nation is of a piece’, writes F. G. Cassidy – editor of both the *Dictionary of American Regional English* and the *Dictionary of Jamaican English* – ‘It is no accident therefore that language, which reflects conditions in the society, is nowhere all of a piece either.’¹ Regional dictionaries assemble non-standard pieces of English into complex pictures of regional language, history, and culture.

Yet regional dictionaries are problematic. For a nation, which words, senses of words, and word forms belong there and nowhere else? Within a nation, where does one draw the line between one region and another, or between a locality and other pieces of an associated region? All dictionaries are instruments of ideology – they look scientific but, in fact, in their contents and structure and methods, they ascribe some values to the language under scrutiny and not others. Ideology is especially marked in regional dictionaries, because both their lexicographers and their readers see national, regional, or local heritage and identity inscribed therein. Codification asserts significance. Regional lexicography proposes that significance both to those who speak the variety or dialect in question and those who do not.

¹ Cassidy, ‘Meaning of “regional”’, 284.

As problematic as they are, regional dictionaries are also necessary, once a standard variety has been selected, elaborated, and itself codified in grammars and dictionaries. Otherwise, our sense of the language is misproportioned.² Besides informing us, dictionaries of regional English reflect our curiosity about human behaviour – what do all those other people say and where do the pieces of English I speak fit into the puzzle? Beside the political, cultural, and identity issues already mentioned, historical dictionaries of regional English are also aesthetically interesting texts, not just because one can find beauty in dialectal speech, but because the quotations in historical dictionaries often place words in aesthetically interesting contexts.

Regional Lexicography of the United Kingdom

The United Kingdom's first notable regional dictionary was John Jamieson's *Etymological Dictionary of the Scottish Language* (1808). Dictionaries don't come from nowhere, and Jamieson's dictionary certainly had its antecedents, especially Andrew Duncan's Latin–Scots glossary (1595) and Thomas Ruddiman's glossary to the 1710 edition of Gavin Douglas' *Aeneis*, two early modern expressions of interest in and celebration of a markedly non-standard variety of English. Such works rode a tide of antiquarianism that rose throughout the early modern period to the very end of the nineteenth century. Jamieson's dictionary was widely recognized not only as an outstanding work of scholarship but also as an innovative dictionary – the earliest dictionary compiled on historical principles as well as the first dictionary devoted to a regional variety of English – and it had considerable influence on English lexicography as it entered a triumphal period, that of the *New English Dictionary* before it became known as the *Oxford English Dictionary* (OED) and Joseph Wright's *English Dialect Dictionary* (EDD, 1898–1905), though, in the case of the former, perhaps not directly.³

Provincial or dialect words first received lexicographical attention from the antiquary Laurence Nowell: in his manuscript dictionary of Old English, the 'Vocabularium Saxonicum' (1566), he indicated which Old English words were, to his knowledge, preserved in one or another sixteenth-century dialect, claims subsequently confirmed in EDD.⁴ One finds similar intimations

² Cassidy, 'Meaning of "regional"', 283.

³ For the reception of Jamieson, see Rennie, *Jamieson's Dictionary*, 156–62, and, for Jamieson and the genesis of OED, see Gilliver, *Making of the Oxford English Dictionary*, 4.

⁴ Nowell, *Vocabularium Saxonicum* (1952), is an edition; see also Marckwardt, 'Unnoted source', 177, 181.

of dialect in other works of the sixteenth and seventeenth centuries.⁵ In 1674, John Ray, the botanist, published *A Collection of English Words Not Generally Used*. The first free-standing glossary or dictionary of English provincialisms, it was reprinted and influential in the eighteenth century.

While it set protocols and standards for big dictionary projects of the nineteenth century, Jamieson's dictionary did not immediately prompt other regional dictionaries. Joseph Hunter's *The Hallamshire Glossary* (1829) is a prominent exception. Hallamshire is a historical region in south Yorkshire, covering Sheffield, Ecclesfield, and Bradfield, and the surrounding suburbs and villages. Hunter represents some 800 words from the region in entries of various lengths and little method, with a strong interest in the supposed archaisms which he sees as preserved in Hallamshire dialect: 'The rustic and the mechanic will speak as his father spoke before him . . . Hence amongst them may be found fragments of our antient tongue, relics of what, three or four centuries ago, constituted the language not of the common people only, but of all ranks from the king to the peasant.'⁶ The words Hunter collected were not archaic to the Hallamshire folk he consulted from 1790 to 1810, who spoke them and explained them to him. Hunter's irrational tendency to see dialect as preserved archaism frequently infects dialect dictionaries, not only in his time and place. Readers too often assume that the language therein is quaint – lexical folklore, but not their language – and lexicographers too often present it as such.

Turning to the entry for *a-gate-wards* – according to Hunter 'pronounced AGATERDS' – one encounters Hunter's typical method and presentation:

This is a very common, and, I may add, very remarkable expression. To go a-gate-wards with anyone is to accompany him part of his way home. *Gate* is the public high-way; *wards* denotes direction, as in *home-wards*, *to-wards*, &c. To go *a-gate-wards*, was therefore to conduct a guest to-wards the high-road, the last office of hospitality, necessary both for guidance and protection, when the high-way lay across an uninclosed and almost trackless country, amidst woods and morasses.

Jamieson was Hunter's professed model, but Hunter's entries fall somewhat short of Jamieson's rigour and reticence.⁷ They are informative and entertaining, but Hunter's volubility sometimes leads him astray. For instance, the *Dictionary of the Older Scottish Tongue* (DOST, 1931–2002), the *Concise Scots*

⁵ Penhallurick, 'Dialect', 291–4; Wakelin, 'Dialect', 157–63.

⁶ Hunter, *Hallamshire Glossary*, xiii–xiv.

⁷ Hunter, *Hallamshire Glossary*, xviii–xix; for Jamieson's style, see Rennie, *Jamieson's Dictionary*, 122–57.

Dictionary (CSD, 1985), and *OED* agree that later use of *a-gatewards* – of about Hunter's time – implies a homewards direction – but also suggest that Hunter's definition is both too narrow and, in its details, overdetermined.

The most influential mid-nineteenth-century compilations were James Orchard Halliwell-Phillipps' *Dictionary of Archaisms and Provincial Words* (1847) – which ran through eleven editions – and Thomas Wright's *Dictionary of Obsolete and Provincial English* (1852). Both focused primarily on early English literature and treated obsolete terms in texts they were recuperating through the Early English Text Society, the Percy Society (Wright), or the Shakespeare Society (Halliwell-Phillipps). Wright did not have the best reputation for care and accuracy, but Halliwell-Phillipps' dictionary was often the only source on which the early *OED* could rely in recording dialect words, and it is cited frequently.⁸

Their literary focus partly missed the point of interest in provincial language, which was the speech of people as well as that of old texts. W. W. Skeat founded the English Dialect Society (EDS) in 1873, the very purpose of which was to produce a comprehensive dialect dictionary.⁹ He saw the need to supplement *OED*, which could not register all provincial words. James A. H. Murray, chief editor of *OED*, eventually served on EDS's executive committee. By 1877, EDS had 350 members.¹⁰ These local experts collected material in the field on which Wright's *EDD* would depend. When Wright became editor of *EDD* in 1887, he inherited more than a million slips recording evidence of dialect words, which he doubled by the time he had finished the dictionary.¹¹

Some members collected their data into glossaries or grammars and published them under the auspices of EDS. In all, EDS published eighty volumes, including *A Glossary of Words in Use in the Wapentakes of Manley and Corringham, Lincolnshire* (1877) by Edward Peacock, a Fellow of the Society of Antiquaries who was also a long-time, productive contributor to *OED*; *Rutland Words* (1891), 'Collected by the Rev. Christopher Wordsworth, M. A., Rector of Tyneham, Dorset, and late of Glaston, Rutland', who knew first-hand of what he wrote, following a tradition of fieldwork established by Jamieson and carried forward by Hunter; and *Sheffield Words* (1888) by Sidney Oldall Addy.¹² Skeat edited many of the older extant glossaries of provincial

⁸ For Wright, see Benzie, *Furnivall*, 92, 150, 176; for Halliwell-Phillipps and *OED*, see Gilliver, *Making of the Oxford English Dictionary*, 69, 71.

⁹ Pennhallurick, 'Dialect', 302. ¹⁰ Pennhallurick, 'Dialect', 302.

¹¹ Wright, *Life of Joseph Wright*, II.355.

¹² For all three volumes, see English Dialect Society, *Complete List of Publications*, 4–5; for Jamieson and fieldwork, see Rennie, *Jamieson's Dictionary*, 141–9.

words, including Ray's *Collection of English Words Not Generally Used, from the Edition of 1691, Together with Thoresby's [1703] Letter to Ray* (1874) – the letter includes a glossary of 529 dialect words, and is also appended to *The Hallamshire Glossary*.¹³

EDD is brilliant, durable lexicography, which is why a digital version was published in 2016, but it has its faults. The glossary evidence is hidden behind abbreviations, never quoted. One finds reference to Hunter's glossary – designated 'w.Yks.'¹⁴ – but encounters none of Hunter's explanations. Thus, the references are actually cross-references, and the dictionary less self-sufficient than one might like, especially if one lacks access to old and often rare glossaries, a problem somewhat mitigated by digital access. By today's standards, *EDD*'s spoken evidence is unhelpfully imprecise: in England, quotations are identified as coming from a speaker in a county – no more specific locality – while all of Scotland is labelled 'Sc.' and all of Ulster 'Uls.' Unlike *OED* citations, examples collected in the field are not precisely dated. Unlike the *Dictionary of American Regional English* (see below), *EDD* provides no information on informants, such as age, sex, and level of education.

EDS disbanded in 1896, by which time dialect antiquarianism – having nourished both *OED* and *EDD* – seemed to have served its immediate purpose, and the hitherto rising tide ebbed. Wright essentially absorbed Hunter's glossary into *EDD*, and Addy may have influenced Wright's confidence in Hunter's work.¹⁴ From Thoresby to Ray to Skeat and from Jamieson to Hunter to Addy to Wright, *EDS* and *EDD* established something of a lexicographical tradition in the treatment of regionalisms, of which *EDD* – at some 70,000 entries – would be the ultimate artifact. Linguists and other scholars have not stopped studying dialect, of course, but there has been no other big dialect dictionary of British English since *EDD*.

In 1919, not long after *EDD* was completed, one of *OED*'s editors, William A. Craigie, proposed a series of period dictionaries to amplify *OED*'s coverage – dictionaries of Old, Middle, Early Modern, and Late Modern English – to the Philological Society in London.¹⁵ Two of them – dictionaries of older Scottish and American English – were regional as well as period dictionaries, and both of these were originally under Craigie's direction and were eventually published. Like so many historical dictionary projects, *DOST* took longer to complete than its original editor proposed. Craigie had conceived

¹³ Hunter, *Hallamshire Glossary*, 101–29.

¹⁴ Beal, 'Contribution of the Rev. Joseph Hunter's *Hallamshire Glossary*', esp. 45–6.

¹⁵ Craigie, 'New dictionary schemes'.

it by 1916.¹⁶ He saw it through to 1956, when he was succeeded by his assistant, A. J. Aitken, who became one of the twentieth century's pre-eminent lexicographers of English in his own right. It was finally completed under the direction of Margaret Dareau in 2002. *DOST* includes roughly 38,000 main entries, with nearly 600,000 illustrative quotations.¹⁷ Craigie expected to publish *DOST* at 3,000 pages, but, in the end, it ran to 8,000.¹⁸

Like some other historical dictionaries – the *Middle English Dictionary*, for instance – the focus and method of *DOST* changed somewhat over time, in the hands of different editors.¹⁹ Craigie stuck with easily available texts, but Aitken expanded *DOST*'s bibliography of print and manuscript sources considerably.²⁰ Whereas Craigie entered forms on the basis of phonological variation – what might be regarded as the same word appeared in several entries if its sound structure had varied in time or space – Margaret Dareau and Harry Watson insisted that semantic development was the only relevant factor in an entry per se or in entry structure.²¹ *DOST*'s later editors illustrated the changes in editorial approach by re-editing one of Craigie's entries, that for *gif* 'give', an exercise that 'highlighted the deficiencies of, and potential for error in, the earlier method'.²²

Craigie was a Scot, and while *DOST* was a component of the comprehensive treatment of English vocabulary he proposed in 1919 – note that he had *DOST* in mind well before that date – it is also an example of identity lexicography and attempts to construct a Scots heritage on which to found regional identity. The *Scottish National Dictionary* (*SND*, 1931–76) picked up where *DOST* left off, covering Modern rather than Older Scots. *SND* treats roughly 50,000 words in 20,000 main entries, with 300,000 citations as evidence of use, but only about 160,000 quotations.²³ The one-volume *CSD*, edited principally by Mairi Robinson, digests *SND* and *DOST*.²⁴ A second edition of *CSD* appeared in 2017, revised by a team too large and complex to itemize here. Like *CSD*, the digital *Dictionary of the Scots Language* is published by Scottish Language Dictionaries Limited; it 'brought *DOST* and *SND* together, thus allowing the reuniting of the two parts of the language, formerly severed at 1700'.²⁵

¹⁶ Dareau and MacLeod, 'Dictionaries of Scots', 307. ¹⁷ Aitken, 'Period dictionaries', 106–7.

¹⁸ Dareau and MacLeod, 'Dictionaries of Scots', 310.

¹⁹ For the changes to the *MED*, see Adams, 'Phantom dictionaries'.

²⁰ Dareau and MacLeod, 'Dictionaries of Scots', 309–10.

²¹ Dareau and MacLeod, 'Dictionaries of Scots', 312.

²² Staff of *DOST*, 'Re-editing of GIF', 27. ²³ Aitken, 'Period dictionaries', 106–7.

²⁴ Dareau and MacLeod, 'Dictionaries of Scots', 323.

²⁵ Dareau, 'DOST: its history and completion', 230.

The Scottish lexicographical approach to Scots has been organized and coherent. Lexicography of Anglo-Irish, on the other hand, has not. There are several worthy recent dictionaries of it, including Terence P. Dolan's *Dictionary of Hiberno-English* (1988, ²2004), James Fenton's *The Hamely Tongue* (1995, ⁴2014), Caroline Macafee's *Concise Ulster Dictionary* (1996), and Diarmaid Ó Muirithe's *A Dictionary of Anglo-Irish* (1996), each focusing on different elements of the vocabulary, for instance, the Gaelic element in Ó Muirithe's case. These and many other dictionaries differ significantly, but in common eschew principled collection and suffer from weak coverage and anecdotal analysis.²⁶ Dictionaries of Anglo-Irish face a conceptual problem, however: 'Is Irish an extraterritorial variety?' asks Michael Montgomery.²⁷ The question is answered partly by Montgomery's own work of identity lexicography, *From Ulster to America: The Scotch-Irish Heritage of American English* (2006).

The regional lexicography of Great Britain took a new direction in the twentieth century upon the founding of the English Place-Name Society in 1923. From that point forward, the society has focused on a survey of English place-names in many volumes, beginning with *The Place-Names of Buckinghamshire* (1925), by Allen Mawer and F. M. Stenton, and reaching to, most recently, *The Place-Names of Leicestershire*, in seven parts (1998–2016), by Barrie Cox. The society has initiated a series of popular place-name dictionaries, such as Keith Briggs and Kelly Kilpatrick's *Dictionary of Suffolk Place-Names* (2016). Such lexicography establishes what we know intuitively, that names as well as words express regionality and participate in regional heritage and identity, but recursively also that modern heritage and identity in some measure require lexicography.

Many think of dialects as natural phenomena, an illusion fortified by association of language with geographical regions, but both are socially constructed. Dialects exist in so far as people perceive them and perhaps most importantly talk about what they perceive as marking one form of speech from another – 'talk about talk', as the linguists sometimes put it – and thus one group of people from another. Linguists now refer to this process as 'enregisterment', and folk dictionaries – a type of talk about talk – participate in it. Returning to Hallamshire, Joan Beal discusses examples of this dictionary genre – represented, for instance, by Scott Dobson's *The Geordie Dictionary* (1974) and Derek Whomersley's *Sheffieldish: A Beginner's Phrase-*

²⁶ Montgomery, 'Core or periphery?', 220. ²⁷ Montgomery, 'Core or periphery?', 222.

book (1981) – and its role in the construction of Sheffield and Geordie dialects.²⁸

Regional lexicography in the United Kingdom has been written at various scales, from thin tourist dictionaries to massive tomes such as *EDD* and *DOST*, but the data within these works, whatever their size, are as granular as any one speaker's verbal attachment to a regional culture and identity.

Regional Lexicography in North America

North America has engendered dictionaries of English abundantly – standard and regional – not least because the United Kingdom established its various Anglophone countries as colonies centuries ago. British colonies later joined together in the United States and Canada, and the Caribbean islands remained under British rule well into the twentieth century. The United Kingdom provided the founding varieties of several North American Englishes, and these later diversified into regional and local dialects within the national varieties. Lexical distinction from the parent variety is an existential matter, at whatever scale.

Because of contact with indigenous American and African languages, among others, North American English diverged almost immediately from British English in some lexical features, a repertoire that expanded considerably during the colonial period. In the soon-to-be United States, differences were so apparent by the late eighteenth and early nineteenth centuries that writers compiled glossaries of Americanisms. Earliest among them is the Revd Jonathan Boucher's list of thirty-eight American words, included in the posthumous edition of his incomplete *Glossary of Archaic and Provincial Words* (1832), edited by none other than the Revd Joseph Hunter. Attentive to American English, Boucher acknowledged its inevitability yet deplored its departures from British English.²⁹

David Humphreys, a popular Connecticut writer, included a glossary to the printed edition of his play *The Yankey in England* (1815?), extending to 275 items.³⁰ Many of them indicate American pronunciation rather than American meanings, and some of them – like *ax* 'ask' – reflect Humphrey's ignorance about the history of British English, for *ax* sailed the Atlantic with its English-speakers and disembarked on American shores. Other items are indeed Americanisms, however, like *boot/to boot* 'in addition' – Humphreys

²⁸ Beal, 'Enregisterment', 148–54. ²⁹ Read, 'Boucher's linguistic pastoral'.

³⁰ Mathews, *Beginnings of American English*, 56–63.

glosses it ‘something given into the bargain’ – as in ‘Humphreys’ glossary includes many examples of Americanized pronunciation and *bona fide* Americanisms, to boot.’

The *Virginia Literary Museum* (1829–30) published Robley Dunglison’s glossary of 190 American words, in three instalments.³¹ Dunglison records many purely American words, such as *backwoodsman* – ‘A new term arising out of the circumstances of the country’, he writes – *blizzard*, *chance* to mean ‘supply, quantity’, and *succotash* ‘mixed maize and beans’.³² He implausibly attempts to derive *tote* ‘carry’ from Latin *tollere* ‘take up, take away’. Other etymologies had been proposed: ‘If we mistake not’, he wrote, ‘Mr Webster considers it a word introduced by the negroes. This is improbable’, but in fact Webster was right. The early glossaries record important facts about colonial American English, but also introduce myths that encode various American political and language ideologies. Attitudes towards the competition between British and American English were mixed into the twentieth century.³³

Webster’s *American Dictionary of the English Language* (1828) was another declaration of independence, yet from a strangely conservative position, for Webster felt he could ‘affirm with confidence that the genuine English idiom is well preserved by the unmixed English’ of the United States. ‘Franklin and Washington, whose language is their hereditary mother tongue, unsophisticated by modern grammar, present as pure models of genuine English, as Addison or Swift’, two of Samuel Johnson’s favourites, as Webster knew well.³⁴ While British English grew decadent, American English preserved an innocence appropriate to a North American Garden of Eden. According to this myth, all American English was what Allen Walker Read would later call ‘Americanism by survival’.³⁵ Of course, many Americanisms were innovations that responded to New World conditions.

Webster had already proposed spelling reform of American English, and his dictionary mildly reiterates that scheme. But the differences between his dictionary and Samuel Johnson’s *Dictionary of the English Language*, for instance, far exceed spelling, however much *color* and *colour* distinguish nations. Whereas Johnson defines *law* in five senses – including ‘A decree, edict, statute, or custom, publicly established as a rule of justice’ and ‘An

³¹ Dunglison, ‘Dunglison’s glossary’ (1927), is an edition; see also Mathews, *Beginnings of American English*, 99–112.

³² Dunglison labels *blizzard* ‘Kentucky’, but see Read, ‘The word *blizzard*’.

³³ Mencken, *American Language*, 3–89.

³⁴ Webster, *American Dictionary* (1828), 2nd page of unpaginated preface.

³⁵ Read, ‘Approaches’, 153.

established and constant mode or process; a fixed correspondence of cause and effect' – Webster resorts to twenty-six, with a certain amount of repetition, it is true, but with an eye to precision that suggests a scientific or technical approach to defining in contrast to the literary approach taken by Johnson, Charles Richardson, and *OED* (see Chapter 23). When he writes, 'Municipal or civil laws are established by the decrees, edicts or ordinances of absolute princes, as emperors and kings, or by the formal acts of the legislatures of free states. Law therefore is sometimes equivalent to decree, edict, or ordinance', he speaks from a post-Revolutionary American perspective in what is arguably a distinctively American voice.

Webster's dictionary is a typological paradox. On one hand, many – including Noah Webster – thought it was a national dictionary of a standard variety, and this seems a reasonable classification, since Noah Webster's dictionary eventually becomes the Merriam-Webster dictionary, which no one takes as 'regional' in any useful sense. On the other hand, in 1828, it announced what others had already noted, that American English was a national variety in contrast to British English: that it was English belonging to a region of the world other than England. To borrow a metaphor from James Murray's preface to *OED*, Webster's dictionary is not describing either the centre or the periphery of English vocabulary, but both. Though Webster does not emphasize it, Dunglison had already begun to enregister dialects within American English: *catch up*, *coudierpt* 'thrown into fits', *dedodgement* 'exit', *honeyfuggle* 'quiz, cozen', and *hornswoggle* 'embarrass irretrievably' supposedly all come from Kentucky, for instance.³⁶ What is regional, in other words, is a matter of scale.

Once Webster had established an American variety of English, lexicographers could return to the question of which words were Americanisms. The nineteenth century saw publication of several works on the subject, including John Russell Bartlett's *Dictionary of Americanisms* (1848, c. 3,700 entries), Maximilian Schele de Vere's *Americanisms* (1872, c. 4,000 entries), and John S. Farmer's *Americanisms Old and New* (1889, c. 5,000 entries). H. L. Mencken asserts that 'these were mainly the work of philological amateurs' and prefers Richard H. Thornton's two-volume *American Glossary* (1912), with its more carefully selected and analysed 3,700 terms.³⁷ Yet Farmer's reputation is currently strong and well supported by his thorough taxonomy of Americanisms, which among other things introduced the notion of 'Americanisms by survival'.³⁸

³⁶ Dunglison, 'Dunglison's glossary'; Mathews, *Beginnings of American English*, 99–112.

³⁷ Mencken, *American Language*, 36.

³⁸ For the taxonomy, see Mencken, *American Language*, 100.

Recall that Craigie's dictionary schemes included a historical dictionary of American English. In 1925, he assumed a professorship at the University of Chicago and embarked on the project, for which he served as chief editor while also co-editing the one-volume supplement to *OED*, with C. T. Onions, and *DOST*, with the help, primarily, of George Watson, his principal assistant on *OED* – Craigie's was a life almost wholly absorbed in lexicography. He and his transplanted Oxford staff were joined from the outset by a young scholar from Alabama, Mitford M. Mathews, somewhat later by Woodford A. Heflin and Allen Walker Read, and by James R. Hulbert, a well-established professor at Chicago, who lent an American gravitas to an American project that could not be led, some suggested, by a British lexicographer alone.³⁹

The *Dictionary of American English* (*DAE*) was published in parts beginning in 1936 and then in four volumes, biennially between 1938 and 1944. The initial printing sold quickly; the dictionary was especially well marketed. The contents, however, leave much to be desired. Having proposed a timeline to completion, Craigie intended to meet it. As a result, *DAE* is rather unambitious. It follows *OED*'s method, without significant innovation. It accounts for American English only to 1900 and excludes most slang and regionalisms. And, as with *DOST*, Craigie limited the texts under review for citations. The published text comprises 2,527 pages, enters some 35,000 head-words, and includes more or less 150,000 quotations.⁴⁰ The treatment of American lexis is thus correspondingly thin.

Well after *DAE* had been completed, Mathews accused Craigie of lacking interest in American history and culture. As a matter of principle, 'the student of words has to be at the same time a devoted student of history'.⁴¹ However, Craigie, according to Mathews, was not much interested in Americana per se. Thus, Mathews wondered that 'he ever had the temerity to undertake to compile a dictionary of the kind he did'.⁴² Craigie believed historical lexicographical method sufficient to describe a vocabulary, without regard for the language which was subject to the method or for variety within a language, and, undoubtedly, he was right up to a point, but not wholly. To discern what makes American English its regional self – not just a branch of Alfred's oak – requires attention to American history and culture. Mathews and Hulbert supplied Craigie with American perspective; too often, he resisted them and went his own way.

³⁹ Adams, 'Special Relationship'. ⁴⁰ Aitken, 'Period dictionaries', 101, 107.

⁴¹ Mathews, 'Of matters lexicographical' (1959), 203.

⁴² Mathews, 'Of matters lexicographical' (1958), 53.

The United States asserted itself in its own interest firmly within a decade of *DAE*'s completion. Mathews remained at the University of Chicago Press to produce *A Dictionary of Americanisms* (*DA*, 1951), in two volumes, mostly culled from *DAE* but not exclusively, and in many matters re-edited. Mathews learned lexicography under Craigie's supervision, but throughout the project Craigie behaved imperially and frustrated Mathews even as he taught him, and in turn Mathews restored Americanisms he thought Craigie had wilfully and foolishly excluded from *DAE* in *DA*.⁴³ *DAE* marked Americanisms – entries and subentry senses – with '+'. In *DA*, Mathews proved that they constituted, independently, a significant regional lexicon.

While they codified American English historically and distinguished it from other varieties in the process, *DAE* and *DA* did not account comprehensively or effectively for variation within the United States. *DAE* includes specifically American words and senses of words that capture American experience and culture, such as *electoral* 'Of or pertaining to the election of the president and vice-president by means of electors', as well as combinations like *Electoral College*, *electoral ticket*, and *electoral vote*. It also enters some American regionalisms, especially senses of standard, nationwide words. Almost everyone agrees that *dogfish* refers to any variety of small shark, but around Lakes Erie and Ontario it refers to bowfin, we discover in *DAE*, which is confusing because it also refers to burbot near Lake Erie. These localities are identified, not by systematic collection of evidence, but within single quotations under the senses, accidentally. Clearly, the United States required a dialect dictionary on the order of *EDD*.

The American Dialect Society (ADS) formed in 1889 with the express purpose of compiling such a dictionary, in the same year that Wright announced he would compile *EDD*.⁴⁴ ADS struggled to organize the project for more than half a century – while of course, doing other worthwhile things – until finally, in 1947, a triumvirate of its leading members – F. G. Cassidy; Allen Walker Read, who had been an assistant editor at *DAE* alongside Mathews; and James B. McMillan, who had twice been Mathews' student, first in high school in Alabama and then at the University of Chicago, and assisted significantly in the preparation of *DA* – proposed the ways and means of what would be published as the *Dictionary of American Regional English* (*DARE*, 1985–2013), under Cassidy's editorial leadership.⁴⁵

⁴³ For the imperial behaviour, see Adams, 'Credit where it's due'; for the frustration, see Adams, 'The apprentice'; for the exclusion and restoration, see Adams, 'Reading between the lines'.

⁴⁴ Markus, 'Introduction', 13; Adams, 'Words of America', 16.

⁴⁵ Cassidy, 'On collecting American dialect'; Adams, 'Before *DARE*'.

From the earliest days, Cassidy had a revolutionary dictionary in mind. *DARE* is a bold synthesis of linguistic atlas and historical dictionary. Like *OED* and *DAE*, it draws on historical documents, correspondence, and literature for quotations illustrating the use of regionally limited terms. It also includes results of a questionnaire conducted systematically across 1,002 communities between 1965 and 1970. The questionnaire, administered by fieldworkers, inquired, 'What do you call a noisy neighborhood celebration after a wedding?' and 'What games do children play around here, in which they form a ring, and either sing or recite a rhyme?'⁴⁶ There were 1,847 questions in all. Entries interweave the traditional quotations and the questionnaire data, the latter a massive 2.3 million bits of lexical information.⁴⁷

Cassidy decided to display questionnaire data relevant to certain entries on maps. He focused not on area, as in a typical geographical map, but on speakers, inventing instead a 'populational' map of the United States, in which big states with few speakers become small and small states with lots of speakers oddly, unfamiliarly big. In many instances, the maps are contrastive: dots represent the speakers surveyed, and show where speakers and how many of them call a grandfather *big daddy* (very southern, with an outlier in Pennsylvania); *pa* (infrequent and widely dispersed); *pa-paw* (dense in the middle of the country and south but present elsewhere); *granddaddy* (used all over but again densely in the south); *papa* (lightly sprinkled across the map); and *pop-pop* (focused in the Mid-Atlantic states, especially Pennsylvania and New Jersey). Presumably, northerners call their grandfathers *grandfather* or *grampa*, which do not receive maps because they are pan-regional items.⁴⁸

Besides these innovations of information and design, *DARE* is an unusually transparent dictionary. Five volumes contain front matter and the entries, but the fifth also includes lists of staff and volunteers, 'contributors of words and wisdom', and financial contributors, long lists that demonstrate how complex and collaborative making a major regional dictionary can be. The sixth volume extracts the contrastive maps (more than 1,600 of them, illustrating not only variation in usage, but also in race, age, sex, education, and community type); includes an index to regional, usage, and etymological labels; and presents the questionnaire with the responses to it, organized from most to least frequent, question by question. The first volume includes biographical essentials for the respondents, so that one can find out about, for instance, the race, age, and sex of the respondent who supplied a given

⁴⁶ *DARE*, I.lxxiv–lxxv. ⁴⁷ Adams, 'Words of America', 18–19. ⁴⁸ *DARE*, VI.184.

answer.⁴⁹ Cassidy had always envisioned these features, but he began to pass leadership of the project to his associate, Joan Houston Hall, with the second volume, and she led the project to completion, with help from many but especially Luanne von Schneidmesser, the senior editor responsible for production.

DARE so dominates American regional lexicography that few have attempted other regional dictionaries, though just as *DARE* approaches American lexis at a finer scale than *DAE* and with denser evidence at that scale, regions, subregions, and localities can refine scale and density yet further. Most prominently, Michael B. Montgomery built on the collections of Joseph S. Hall to produce a *Dictionary of Smoky Mountain English* (2004), and an enlarged version, titled *Dictionary of Smoky Mountain and Southern Appalachian English*, edited with Jennifer H. Heinmiller, is pending.⁵⁰ Some sociolinguists discuss the enregistering role of folk dictionaries; American yard sales are well stocked with them, but they are overlooked by scholars.⁵¹

As in the United Kingdom, place-name dictionaries help to define the national variety, as well as regionalisms and localisms within it. Dictionaries of broad scope, such as George R. Stewart's *American Place-Names* (1970) and William Bright's monumental *Native American Placenames of the United States* (2004), are accompanied by excellent dictionaries of names at the state level, such as Edward Callary's *Place Names of Illinois* (2008) and *Place Names of Wisconsin* (2016). F. G. Cassidy, just as he was taking the reins of *DARE*, set a standard at a more specific scale with *Dane County Place Names* (1947; ²2009). Since then, dictionaries have explained county place-names from sea to shining sea. Unlike the English, however, Americans have not approached naming of their nation systematically.

Charles J. Lovell, an American, was nonetheless devoted to Canadian English and was collecting Canadian data suitable for lexicography while he worked under Mathews as research and editorial assistant on *DA*. Craigie had resisted the very notion of Canadianisms, according to Mathews, but Mathews believed in them and proposed a dictionary of them.⁵² Lovell was eventually named chief editor on such a project, but he died in 1960, and Walter S. Avis led the editorial team that saw it to publication by W. J. Gage,

⁴⁹ For more on volumes V and VI specifically and *DARE* generally, see Adams, 'Lexical ride'.

⁵⁰ For the latter, see Montgomery and Heinmiller's article, 'Dictionary of Smoky Mountain and Southern Appalachian English'.

⁵¹ Notable among the sociolinguistic discussions is Johnstone, *Speaking Pittsburghese*.

⁵² Mathews, 'Of matters lexicographical' (1956), 202–3.

at the time Canada's leading dictionary publisher, as *A Dictionary of Canadianisms on Historical Principles* (DCHP, 1967).

DCHP includes many folk words for flora and fauna, terminology from various industries – mining and lumbering, for instance – and other words peculiar to the Canadian natural and cultural landscapes. Some words belonged to the nation, such as RCMP 'Royal Canadian Mounted Police', *returned man* 'soldier back in Canada from overseas', and *Maritimes* 'provinces of Nova Scotia, New Brunswick, and Prince Edward Island'. Others belonged to one province but not others, such as Nova Scotia's *road district* 'district whose roads are under local control'. Some terms are historical, such as *camboose* – from Canadian French *cambuse* – 'stove used by explorers; open fire in a lumber camp'. While *poke* 'sleeping bag' is undoubtedly a Canadianism, *poke* 'bag' is not, having been brought by Ulster Scots settlers to the United States as well as by Scots settlers to Canada, except that DCHP refines the definition in a specifically Canadian direction: 'a bag or small sack especially one used in carrying gold dust or nuggets', a remnant of Western Canadian mining culture.

The complexities of *poke* are not uncommon, and North American regional lexicographers have suggested that earlier dictionaries oversimplify questions of regionality – words may legitimately be North Americanisms, as well as Americanisms or Canadianisms.⁵³ Even before Stefan Dollinger, Laurel Brinton, and Margery Fee had mounted a digital DCHP-1 (2013), Dollinger, as chief editor, and Fee, as associate editor, were leading a focused, partial revision, which would be published online as DCHP-2 (2017). DCHP-1's entry for *poke* remains unrevised, but many entries of current interest have been thoroughly revised in DCHP-2. For instance, DCHP-1 enters *parkade* – a word for what most Americans call a *parking garage* or *parking ramp* – as a Canadianism, very recent at the time DCHP-1 was published. DCHP-1 illustrates the item with two quotations from 1958. DCHP-2 includes ten more, including one that antedates the record by a year. When it intervenes, then, DCHP-2 provides a richer and more accurate account of Canadianisms than DCHP-1, and it exploits the digital medium by including full-colour photographs (of, for instance, parkades) and eloquent frequency graphs.

DCHP-2 presents users with a much more discriminating typology of Canadianisms than its predecessor: Type 1 Canadianisms, for instance, are those that originated in what is now Canada; Type 5 Canadianisms are words or senses used much more frequently in Canada than elsewhere, especially

⁵³ See Dollinger and von Schneidmessa, 'Canadianism, Americanism, North Americanism?'

the United States. Further investigation of individual words will doubtless change some of their places in the typology, and in fact, *parkade*, presented as an example of both Type 1 and Type 5 on the release of *DCHP-2*, appears to belong only to the latter category.⁵⁴ This is not meant to criticize *DCHP-2* unduly: Avis had already posed the problem of untangling Americanisms and Canadianisms in his introduction to *DCHP-1*, and he pointed out that *DA*, while supposedly about Americanisms, includes an unexpected amount of Canadian material. Words such as *parkade* underscore the conceptual problems underlying regionality, with which we began.

While *DCHP-1* and *DCHP-2* distinguish Canadian English as a variety of English worldwide, some provincial Canadian lexis also receives excellent lexicographical treatment. Principal among such dictionaries is the *Dictionary of Newfoundland English* (*DNE*, 1982), edited by G. M. Story, W. J. Kirwin, and J. D. A. Widdowson, which comprises 625 double-column pages of entries and starts with unusually full front matter, including an extensive bibliography and list of collections and collectors contributing citational and other evidence for the entries – for, like *EDD* and *DARE*, *DNE* depends in part on fieldwork among living speakers. *DNE* well illustrates the ideological motives underlying many a dictionary of regionalisms, namely, to emphasize the region in question, its culture and history, as reflecting and building regional identity within communities and individuals. Newfoundland had a long history as a British colony and dominion before it became Canadian in 1949. Almost immediately thereafter, work on *DNE* began.⁵⁵ A second edition was published in 1990, and since then many of the project's underlying materials have been digitized, a precondition for a third edition.⁵⁶

Roughly a quarter of *DNE*'s size, T. K. Pratt's *Dictionary of Prince Edward Island English* (1988) presents the lexicon of a smaller population: roughly 150,000 people live in Prince Edward Island, while more than 200,000 live in the capital of the Province of Newfoundland and Labrador. It includes 873 main entries and treats about 1,000 words altogether. Many of these are not Prince Edward Islandisms – other dictionaries, including *OED*, *EDD*, *DAE*, *DA*, and *DARE* in various combinations, also include them – but the notes indicating cross-regional affiliations, placed at the ends of the relevant entries, are informative and very useful to those interested in regional English. Pratt, collaborating with Scott Burke, followed his lexical dictionary with another well-executed volume, *Prince Edward Island Sayings* (1998), presenting

⁵⁴ Considine, 'Parkade'. ⁵⁵ Webb, 'Cullers of words'. ⁵⁶ Power, 'Aaron's rod to zurr'.

material often integrated into the entries of other North American regional dictionaries, such as *DAE*, *DA*, and *DARE*.

The effects of language contact contribute greatly to regionality throughout the English-speaking world, but in some places, such as the Caribbean islands which once were or still are British colonies and territories, contact produces creoles much more difficult to treat in dictionaries than any dialect which is close to a standard variety in form and function. Intrepid lexicographers attempted description of these varieties as early as 1905 and throughout the twentieth century, with dense but miscellaneous treatment of them in the 1970s and 1980s, focusing, for instance, on the Virgin Islands and Trinidad and Tobago.⁵⁷ The dictionaries they produced were not historical, nor were they prepared as systematically as the highest lexicographical standards demand.

At that standard, dictionaries must depend on evidence from fieldwork, and so R. B. Le Page organized the Linguistic Survey of the West Indies in 1953, the very year in which Cassidy and Audrey R. Duckert published their plan for the Wisconsin Dialect Survey, precursor to the *DARE* questionnaire.⁵⁸ Cassidy and Le Page joined forces to produce a *Dictionary of Jamaican English* (*DJE*, 1967, ²1980), which, like *DARE*, is a historical dictionary – written evidence of Jamaican English appears as early as 1655 – infused with field material, including that captured by the questionnaire, but also collected by Cassidy and others, including words extracted from '[t]he spontaneous conversation of children in school playgrounds or in their homes'.⁵⁹ Some of the etymologies lead back to African-language substrates of the creole, something we might take for granted now but which was pioneering then. Work on specific Caribbean English creoles continues from *DJE*'s foundation, for instance, in John Holm and Alison Watt Shilling's *Dictionary of Bahamian English* (1982) – a significant example of sociolinguistically motivated lexicography – and Lise Winer's *Dictionary of the English/Creole of Trinidad and Tobago* (2009), like *DJE* a historical dictionary.

If *DJE*'s value lies partly in its focus on the English of one Caribbean island, Richard Allsopp's *Dictionary of Caribbean English Usage* (1996) extends ambitiously across the region, accounting for commonalities across the islands – and extending to Anglophone coastal countries such as Guyana and Belize – but also for regionalism within the region. The *Dictionary of Caribbean English Usage* contains about 20,000 words, whereas Winer's

⁵⁷ Allsopp, 'Dictionaries of Caribbean English', 354.

⁵⁸ Cassidy with Duckert, *Method for Collecting Dialect*; see also the call in Mathews, 'Of matters lexicographical' (1956), 203.

⁵⁹ Cassidy and Le Page, *Dictionary of Jamaican English*, viii.

dictionary includes just over 12,000, and *DJE* some 15,000, and from those figures one guesses that its coverage is relatively thin because it is so broad, and the other dictionaries comparatively deep explorations of more restricted lexis. However, appearances can be deceiving for, as Allsopp argues in criticizing *DJE*, Cassidy and Le Page seem to privilege Jamaican English historically, though that variety is best viewed, according to Allsopp, within 'the massive sociological samenesses and similarities of the Caribbean scene', lines of regional interrelationship rather than difference that his own dictionary draws.⁶⁰ Though not a historical dictionary, the *Dictionary of Caribbean English Usage* is nonetheless a quotations dictionary, and it draws compellingly on a Caribbean literary tradition largely ignored by others.

Each hamlet in England, Scotland, Ireland, and Wales has its distinctive speech (consider place-names again), and by no means have lexicographers accounted for all of Great Britain's regionally salient vocabulary. Yet even a brief survey like this shows how the geographical expanse of North America multiplies the array of flora, fauna, and topographical features for which Americans must have words, words that not only pin down the dragonfly, but in forms like *snake doctor* and *skeeter hawk* pin it to regional identities spreading across that expanse and through time. New political institutions from province to province and state to state, new social relations in town and country – the United States has a lot of regional words and needs a lot of dictionaries to store up its lexical treasure.

Regional Lexicography in Other Colonial and Post-Colonial Situations

The United States, Canada, the former and current Caribbean colonies and territories – one can assess the status of English in all these places according to Edgar Schneider's Dynamic Model of Colonial/Post-Colonial Englishes.⁶¹ National varieties move from applying extra-colonial norms through nativization of English to achieving endonormative stabilization and further to differentiation of the variety into dialects. Dictionaries codify the variety from within and identify dialects against a standard variety and one another. Anglophone cultures well advanced in the model have produced dictionaries as fine as those produced in Great Britain and North America.

⁶⁰ Allsopp, 'Critical commentary on the *Dictionary of Jamaican English*'.

⁶¹ Schneider, *Postcolonial English*, 21–70, but also throughout; for criticism of indiscriminate application of this model, see Denis and D'Arcy, 'Settler colonial Englishes'.

In this model, time matters, and the Anglophone cultures most likely to have codified their varieties and noticed variation within them are among the old colonies, those with long histories of English that also have been independent for a while, though the first example of such a variety and its dictionary, the Anglo-Indian *Hobson-Jobson* (1886), compiled by Henry Yule and A. C. Burnell, was decidedly pre-post-colonial. Immediately popular, it appeared in a second edition by William Crooke (1903). Like most colonial or post-colonial dictionaries, it registers effects of language contact: *gindy* 'drinking vessel; washing bowl' is borrowed from Dravidian languages, it tells us, just as American *raccoon* derives from Algonquian. But Anglo-Indian speakers invented other words to characterize their experience in South Asia: *Corporal Forbes*, according to *Hobson-Jobson*, is 'A soldier's grimly jesting name for *Cholera Morbus*'.

Hobson-Jobson attracted the attention of James Murray, chief editor of *OED*, who corresponded with Yule about Anglo-Indian vocabulary and extracted information otherwise unknown in Oxford from the proofs of his dictionary.⁶² Murray did not wish to ignore significant English lexis from around the world, and Anglo-Indian was suddenly a prominent variety, the subject not only of *Hobson-Jobson* but of G. C. Whitworth's almost simultaneous *Anglo-Indian Dictionary* (1885).⁶³ Murray's preference established *Hobson-Jobson* as the better dictionary. Interest in it has risen so high recently that it prompted a new, accessible edition of the dictionary. Even the name *Hobson-Jobson* carries ideological weight.⁶⁴ The dictionary is a crucial text in colonial and post-colonial studies, as well as an essential record of historical regional vocabulary.

By the time *Hobson-Jobson* was published, Britain had been in contact with South Asia, through the agency of the East India Company, since the early seventeenth century, and had been in control of much of the Indian sub-continent since the middle of the eighteenth, plenty of time for an identifiable variety of English to rise and demand codification. In the twentieth century, other post-colonial varieties came of age, and dictionaries followed. *The Australian National Dictionary* (*AND1*, 1988) is just such a dictionary, presenting roughly 10,000 main entries and 57,000 citations. *AND1* naturally treats words for native flora and fauna, such as *amulla* 'fruited shrub of New South Wales', unfamiliar to outsiders and perhaps to Australians who are not from or

⁶² Nagle, 'Visible and invisible influence'.

⁶³ For Murray's attitude, see Ogilvie, *Words of the World*, 53–103.

⁶⁴ Nagle, "'There is much, very much, in the name of a book'"; Lambert, 'Much tortured expression'.

resident in New South Wales. But it also includes compendious entries for well-known words such as *boomerang*, with mistaken reports of the form as early as 1790, and ingenious compounds like *boomerang bill* 'legislation that rebounds on legislators' and *boomerang cheque* 'bounced cheque'.

W. S. Ransom, the chief editor of *AND*₁, believed that distinctively Australian lexis was a thing of the past by the time the dictionary appeared, and that Australian English was destined to be the 'handmaid' of American English.⁶⁵ Subsequently, Bruce Moore and his colleagues have demurred from those assertions and produced a second edition of *AND*, with an additional 6,000 headwords and more than twice as many quotations. *AND*₂ improves on its predecessor by taking account of several smaller dictionaries published after *AND*₁, including R. M. W. Dixon, W. S. Ramson, and Mandy Thomas' *Australian Aboriginal Words in English* (1990), listing some 400 words borrowed from 60-plus Aboriginal languages. For a long time, Australians insisted on the homogeneity of Australian English, admitting to three social accents – the stereotyped Broad Accent, a General Accent, and a Cultured Accent barely distinguishable from England's Received Pronunciation – but no regional variation. Schneider's Dynamic Model, however, predicts the eventual differentiation of Australian English, and, indeed, regional variety is now documented in works such as Maureen Brooks and Joan Ritchie's *Words from the West* (1994), *Voices of Queensland* (2001), edited by Julia Robinson, and Dorothy Jauncey's *Bardi Grubs and Frog Cakes: South Australian Words* (2004). The future of Australian English may lie less in its relationship to American English than in internal variation.

AND is a historical dictionary in the *OED* tradition and so is *A Dictionary of South African English on Historical Principles* (1996). Like *AND*, the *Dictionary of South African English* contains many terms for indigenous flora and fauna and also registers complex borrowings from African languages, as well as Afrikaans and South Asian languages. Usually, lexicographers produce regional dictionaries, on whatever scale, to celebrate the relevant varieties of English – the vocabularies therein constitute identity, a region's identity or a lexicographer's, or both. 'South Africans are notorious for their inferiority complex about all things South African', Penny Silva, chief editor of the dictionary, writes in its front matter, 'and this is true too of their English.' Sometimes, then, a dictionary attempts to raise the prestige of a national or local variety beyond the willing expectations of its speakers, an ideological

⁶⁵ Ramson, *Lexical Images*, 247, xv.

burden different from that shouldered by *CSD* or *DA*. Each regional dictionary will speak, in some way, to its specific cultural situation.

Some cultural situations are unlikely, but produce English lexis nonetheless. Believe it or not, *The Antarctic Dictionary* has been compiled (by Bernadette Hince, 2000), accounting for English in the culture of a place without many permanent residents, most of whom, indeed, live not on the Antarctic continent, but in Tristan da Cunha, the Falklands, and the sub-Antarctic islands. The dictionary rests on a foundation of some 20,000 citations – 15,000 or so are printed in the dictionary – and follows historical principles.⁶⁶ Some of the words included, especially those for flora and fauna, have long histories – *blue petrel* first appears in Cook's *A Voyage towards the South Pole*, in the entry for 23 December 1772. By the nineteenth century, explorers wrote of *ice blink* 'white or yellow reflection above the horizon from distant ice, proving it's there'. Then, twentieth-century scientists and other semi-residents coined terms like *scattered pack*, a synonym – dating from 1964 – of *open pack* 'broken sea ice which is easily navigable', first attested in 1901. Reading the *Antarctic Dictionary*, though less arduous than a trek across the seventh continent – *seventh continent* is entered – is an armchair adventure.

As English ages and differentiates in the post-colonial nations, develops and shifts in the old Anglophone countries, and establishes itself newly in yet other places, lexicographers will write yet more dictionaries to capture the breadth and depth and cultural significance of regional English, from Hallamshire to the ends of the earth.

⁶⁶ Hince, *Antarctic Dictionary*, viii.

The Romance Languages from c. 1700

PASCALE RENDERS

(TRANSLATED BY JOHN CONSIDINE)

This chapter presents the history of the lexicography of the Romance languages, from the eighteenth century to the present day. In practice, emphasis is placed on French, Italian, and Spanish, together with Portuguese and Romanian. Behind these languages of the first rank, there are other languages, which will as far as possible be mentioned in this discussion.

Historians of the Romance languages distinguish the Gallo-Romance linguistic domain (comprising French, Francoprovençal, Occitan, and Gascon), the Italo-Romance domain (Italian, Friulian, Ladin, and Romansh), and the Ibero-Romance languages (Spanish, Portuguese, Galician, Asturian, Aragonese, and Catalan), to which I add the Romanian domain (divided *grosso modo* into two main varieties, Daco-Romanian and Aromanian) and Sardinian. If one likewise takes account of the geolinguistic varieties present within these various linguistic domains, the number of Romance language varieties capable of being made the objects of lexicographical description is considerable.¹

Moreover, the period treated here extends over three centuries, which have been very fertile in terms both of the quantity and of the quality of their lexicographical production. This production has followed the historical and cultural movements which have arisen on the European continent: the sixteenth and seventeenth centuries were marked by the cultural influence of the Italian Renaissance and by Spanish political domination; the history of the lexicography of the eighteenth and nineteenth centuries was, for example, strongly dominated by French culture; and, in the twentieth century, the latter has in its turn withdrawn, to give place to English as prestige language, concurrently with the emergence of American cultural influence.

¹ For an overview of the Romance languages, see Ledgeway and Maiden, *Oxford Guide to the Romance Languages*.

The sections which follow will give a summary of the lexicographical history of each of the five principal Romance languages, while also taking notice, as far as possible, of the description of the regional variants present in their respective linguistic domains. Certain languages considered, wrongly or rightly, as 'minor' are treated thereafter. Before reviewing these languages individually, I shall give a brief general treatment of some features which are common to the whole body of the Romance languages.

General Tendencies and Pan-Romance Lexicography

The history of the European lexicography of the eighteenth and nineteenth centuries is closely related to the topic of the formation of the national languages. Language academies were created successively in Spain, in Portugal, and in Romania on the models of the Accademia della Crusca and the Académie française (for which see Chapter 14), their aim being to unify the national languages and to fix their orthographical, grammatical, and lexical norms. Each of these institutions had the goal of producing a dictionary, and these dictionaries shared a prescriptive and selective approach: they excluded all but a limited part of the vocabulary, that portion which was judged to be representative of exemplary 'good usage'.

At the same time, the Age of Enlightenment saw the development of scientific knowledge, accompanying the appearance of great encyclopedic works such as the *Encyclopédie* of Diderot and d'Alembert in France. As these two movements came together, a lexicographical fault line was opened: the desire to codify the language produced works with a restrictive entry list, at just the same time as the vocabulary of the Romance languages was being enriched by numerous technical terms. The solution would be a new genre, the 'encyclopedic dictionary', which contained linguistic information and encyclopedic commentary at the same time. This genre would be very successful from the nineteenth century onwards, particularly in France and Spain.

Likewise characteristic of the nineteenth century in Europe was the beginning of modern philology. The discovery of Sanskrit and the Indo-European languages instigated a renewal of lexical studies. For the first time, the theory of 'phonetic laws' was expressed; these were supposed to explain the development of sounds from the fragmentation of Latin to the contemporary Romance languages. Romance philology first came into being as an academic discipline in Germany (the first professorial chair of Romance

languages and literatures was established there in 1822, and the first chair of 'Romance philology' in 1867).² The German model was then taken up in Paris and subsequently throughout France. In the first half of the nineteenth century, Friedrich Diez, inspired in particular by the German studies of Jacob and Wilhelm Grimm (see Chapter 22), published a grammar (*Grammatik der romanischen Sprachen*), which showed how Popular Latin had evolved to give birth to the various Romance languages. This grammar was followed by a pan-Romance etymological dictionary, the *Lexicon etymologicum linguarum Romanorum* (1853).

The desire for knowledge of the diverse language varieties of Romance-speaking Europe (often in a political context of their eradication in favour of a unified national language) led to the birth of linguistic geography and of dialectology. Linguistic atlases were published, first in France (the *Atlas linguistique de la France* of Jules Gilliéron appeared between 1902 and 1910), and then in Italy and Spain. This dialectological knowledge permitted both the making of more rigorous etymologies, and the appearance of the *Romanisches etymologisches Wörterbuch* of Wilhelm Meyer-Lübke, a Swiss Romanist who made use of the rich documentation provided by the dialectological studies and added such 'minor' languages as Dalmatian (now extinct) and Francoprovençal to those which had been treated by Diez. The material contained in his work is now being revisited within the scope of a wide-ranging lexicographical project, the *Dictionnaire étymologique roman*, directed by Éva Buchi and Wolfgang Schweickard, the distinctive features of which are the application of the principles of comparative grammar and reconstruction used for the other Indo-European languages to the Romance languages, and hence the proposal of reconstructed etyma which are not necessarily attested in classical Latin.

Philological rigour was also brought to bear on texts, leading, as we shall see, to the birth of large historical and philological dictionaries in the twentieth century. The twentieth century has likewise been marked by the development of linguistic theories, about which I shall say little, but which have had a notable influence on the conception of learners' dictionaries.³ Finally, the last decade of the century was marked above all by the emergence of large electronic corpora: the digital revolution continues to transform the methods of the compilation of dictionaries and the modalities of their reception and consultation.

² See Swiggers, 'Linguistique romane', 43–52.

³ See Rey, 'Lexicographie française depuis Littré', 1832–6.

French

In 1700, the French language was in a privileged position among the Romance languages. Indeed, it possessed three large general monolingual dictionaries – those of Richelet, the Académie française, and Furetière – which would appear in further editions throughout the eighteenth century.⁴ These three dictionaries set up lexicographical models for the following centuries, by establishing a distinction between, on the one hand, dictionaries of the language (represented by the dictionaries of Richelet and the Académie française, the former being descriptive and the latter normative) and, on the other hand, encyclopedic dictionaries, of which that of Furetière is the prototype. Thereafter, abundant dictionary production and a receptive public made French the European language which was most completely registered in dictionaries in the twentieth century, with more than 15,000 editions of general or specialized monolingual wordlists.⁵

As was noted in Chapter 14, the first edition of the *Dictionnaire de l'Académie française* followed derivational principles in its presentation of lexemes. This arrangement was inconvenient for the reader, and it was subsequently abandoned (it would reappear in certain twentieth-century dictionaries), but it would influence a morphological, rather than phonetic, tendency in French orthography. Editions after the first conformed to alphabetical order – and, from 1835 onwards, provided an official standard for the orthography of the French language. From the fourth edition (1762) onwards, the academy dictionary was considered to contain the essential, central part of the vocabulary of French, and would be used as a reference outside France, for instance in the compilation of bilingual dictionaries. In France, however, it would be the butt of many witticisms, particularly on account of the slow progress of the undertaking, and other monolingual dictionaries would soon be produced in response to it, designed to make up for its numerous deficiencies.

A first criticism directed at the dictionary of the academy was the absence of technical terms, which had been systematically excluded from its list of headwords. The dictionary of Furetière had appeared in response to this deficiency; in its wake, the genre of the 'dictionnaire universel', which combines linguistic and encyclopedic information, would develop extensively in France. In this lexicographical tradition, a special place must be given to the *Dictionnaire universel françois et latin* undertaken by the Jesuits of

⁴ Bray, 'Lexicographie française', 1801. ⁵ Quemada, 'Französisch: Lexikographie', 870.

Trévoux and more familiarly known as 'Le Trévoux'. The first edition of this work, which appeared in 1704, plagiarized an edition of the dictionary of Furetière, while modifying certain articles in the interests of anti-Protestant propaganda. The editions which followed were improved by the incorporation of new material, particularly from 1751 onwards, this being the date of the publication of the *Encyclopédie* of Diderot and d'Alembert. From the nineteenth century onwards, the genre of the encyclopedic dictionary was represented in France by the Larousse series, which would have a lasting success, continuing to the present day. Between 1865 and 1876, Pierre Larousse published a *Grand Dictionnaire universel du XIXe siècle* in seventeen volumes, in which the encyclopedic element is predominant, but citations of authors and invented lexical examples are also integrated. An abridged version intended for educational use appeared in 1905, namely the one-volume *Petit Larousse illustré*, which has been hugely successful ever since its publication. Since then, the Larousse series has diversified to address the needs of different readerships, in a wide range of abridged, updated, and specialized editions.

As well as the absence of technical terms, a second criticism directed at the *Dictionnaire de l'Académie française* was the absence of citations of authors. The academicians had, indeed, decided to make their unique expertise the basis for selecting and defining exemplary good usage. This decision makes sense in light of the fact that the golden age of French literature was precisely contemporaneous with the creation of the Académie française; nevertheless, this absence of citation hinders any diachronic study of the vocabulary of French. From 1787 (the *Dictionnaire critique* of the abbé Féraud), dictionaries which included textual citations, following the example of the dictionary of Richelet, appeared again. An important stage was passed with the publication in 1863 of the *Dictionnaire de la langue française* of Littré, which marked the beginning of historical and philological French lexicography, anticipating fine subsequent work. Émile Littré based his work on a corpus of literary texts, presenting numerous citations for every entry. However, the sequence of senses was logico-semantic rather than historical, and the provision of diachronic information remained limited. The work stands out on account both of the importance of the materials which it assembled, and of the conjunction which it effected between metalinguistic elements and literary documentation. For the linguistic period which it describes, the seventeenth and eighteenth centuries, Littré's work is still a standard dictionary. Its outdated etymologies would be improved at the beginning of the twentieth century in the *Dictionnaire général* of Hatzfeld, Darmesteter, and Thomas, which would benefit from the progress that had been made in historical linguistics

since Littré, and would set out a real history of morpho-semantic development, a landmark achievement of that kind in philological lexicography. Shortly after Littré's dictionary, the *Dictionnaire de l'ancienne langue française* of Godefroy would appear; despite its deficiencies, it is still a precious collection of citations for the medieval period (the ninth to the fifteenth centuries).⁶

After the publication of the dictionaries of Littré and Godefroy, and of the *Dictionnaire général*, no major new language-oriented dictionary of French would appear for half a century. The lexicography of the first half of the twentieth century was dominated by encyclopedic and pedagogical dictionaries. Meanwhile, a philological renewal was in preparation, and progress was being made in etymology. A remarkable example of this progress in the Gallo-Romance linguistic domain is the *Französisches etymologisches Wörterbuch (FEW)* of Walther von Wartburg (who was also the author, with O. Bloch, of a *Dictionnaire étymologique de la langue française*). This monumental work brings together the vocabulary of French, Francoprovençal, Occitan, Gascon, and all their diatopic varieties, from the first texts (of the ninth century) up to modern times. The lexis is organized from a genetic perspective, each article tracing the history of a lexical family from the etymon with which it originates. The twenty-five volumes of this work, published from 1922 to 2002, will be of considerable importance for the historical linguistics and lexicography of the Romance languages: the majority of present-day dictionaries that give a scholarly account of the vocabulary of French, Gallo-Romance, or indeed Romance more generally refer more or less systematically to the *Französisches etymologisches Wörterbuch*.⁷

Following this philological renewal, new language-oriented dictionaries of French came into being after 1950. Alongside the Larousse series of encyclopedic dictionaries appeared the Robert series. In 1951, Paul Robert published his *Dictionnaire alphabétique et analogique*, which was originally intended as an update of the dictionary of Littré, and included citations from authors of the sixteenth to the twentieth centuries. Like the Larousse dictionaries, the Robert dictionaries then appeared in forms for different readerships, including a single-volume dictionary for a broad reading public (the *Petit Robert*), a format destined to be highly successful. However, the monumental work of the second half of the twentieth century is incontestably the *Trésor de la langue française*, an institutional project which was launched in 1957. This philological dictionary documents the period from 1789 to the twentieth century – in other words, it offers a synchronic treatment of a 170-year period – complementing

⁶ See Trotter, 'Gallo-Romance II', 664. ⁷ See Buchi and Renders, 'Gallo-Romance I'.

the chronological scope of the dictionary of Littré. From its conception, it offered a major methodological innovation: it was the first language-oriented French dictionary to make use of a huge corpus of digitized texts (today, the *Frantext* database contains more than 5,000 texts, ranging from the tenth century to the twenty-first, with more than 290 million tokens). Recourse to this corpus, of great quantitative and qualitative importance, gives its linguistic description a high degree of objectivity. Another characteristic of the entries is a clear separation between the synchronic section and the section titled 'Etymology and history', which is extremely rich, but was edited independently from the first section. The digitization of the sixteen print volumes and their automated conversion to a structured electronic resource capable of being searched with precision was likewise a pioneering enterprise, comparable to the digitization of the *Oxford English Dictionary* (see Chapter 23). The online version, *Trésor de la langue française informatisé*, is now accessible, by means of a shared portal, together with other digitized dictionaries, among them some editions of the *Dictionnaire de l'Académie française*.⁸

So, historical lexicography has found its way in France in the form of period dictionaries, each covering a stage in the history of the language. After the dictionary of Littré and the *Trésor de la langue française*, which respectively describe the language in its classical form and the modern period, twentieth-century lexicographers would privilege the medieval period.⁹ The dictionary of Godefroy was joined by the eleven-volume *Altfranzösisches Wörterbuch* of Adolf Tobler and Erhard Lommatzsch (more familiarly known as 'Tobler-Lommatzsch'), which offers a more rigorous philological approach; this was followed by the *Dictionnaire étymologique de l'ancien français* of Kurt Baldinger, the *Anglo-Norman Dictionary* of William Rothwell, and the *Dictionnaire du moyen français*. These standard works, now available online, complement and often correct the *FEW* for the period in question.¹⁰ The story of the conversion of these dictionaries to electronic form moreover illustrates the diverse procedures used in the present period of digital transition: whereas the sixteen print volumes of the *Trésor de la langue française* have been completely digitized, the *Dictionnaire du moyen français* was conceived from the beginning as a dynamic electronic resource. As for the *Dictionnaire étymologique de l'ancien français*, it changed its method in the course of its compilation,

⁸ See Trotter, 'Gallo-Romance II', 666.

⁹ The vocabulary of the Renaissance is treated in Huguet, *Dictionnaire de la langue française*; see Trotter, 'Gallo-Romance II', 666.

¹⁰ Cf. Buchi and Renders, 'Gallo-Romance I', 655–6.

moving from printed fascicles to an online resource (which now provides raw materials at the same time as fully edited articles).

The description of geographical variation in the French language is in its own right a major innovation in the French lexicography of the twentieth and twenty-first centuries. Three dictionaries of regionalisms have raised the genre to a high level of scholarly description: the *Dictionnaire historique du français québécois* of Claude Poirier, the *Dictionnaire des régionalismes de France* of Pierre Rézeau, and the *Dictionnaire suisse romand* of André Thibault. Together with the *Dictionnaire des belgicisms* of Michel Francard, these works provide reliably excellent coverage of the linguistic territory. Their microstructure is exemplary: they are without any doubt the best representatives of what differential lexicography at a scholarly level can be today. Discussion of the place which regionalisms should have in reference dictionaries is particularly lively in Quebec, which has for a long time aimed at freedom from European lexicography. After a century of normative dictionaries, in which regionalisms were presented as faults, a first descriptive work, the *Glossaire du parler français au Canada*, appeared in 1930. This foundational work had a great influence on Québécois lexicography, despite the subsequent arrival of a new normative dictionary, the *Dictionnaire général de la langue française au Canada* of Louis-Alexandre Bélisle, based on a Canadian edition of the dictionary of Littré.¹¹ The idea that the lexical variation of French is capable of being described then made its way to Quebec: Québécois lexicography has recently offered non-differential general dictionaries (such as the *Dictionnaire Usito*), which give an exhaustive account of the standard French of North America, and even mark words and senses which are used exclusively in France as ‘francismes’.

Italian

The Italian language had experienced a golden age of both monolingual and bilingual lexicography in the sixteenth and seventeenth centuries (see Chapter 14). The eighteenth century seems comparatively impoverished. Apart from a fourth edition of the *Vocabolario* of the Accademia della Crusca (1729–38), the main lexicographical activity was a matter of translations of foreign works or of new editions of the works of previous centuries. For example, the *Vocabolario* of Lorenzo Franciosini would go through fifteen editions after the first (1620) and would provide a basis for the bilingual

¹¹ See Cormier and Francoeur, ‘Un siècle de lexicographie au Québec’.

Italian–Spanish dictionaries of the eighteenth and nineteenth centuries. As contacts between France and Italy became still more intense in the eighteenth century, a new contribution was made to bilingual Italian–French lexicography: the *Nuovo dizionario italo-francese* of Francesco d’Alberti di Villanuova, a lexicographer who likewise distinguished himself at the end of the century by bringing out a monolingual work of an encyclopedic nature, the *Dizionario universale critico enciclopedico della lingua italiana* (1797–1805).

Italian lexicography reached a new level in the second half of the century, with the appearance of a monumental work which definitively replaced the dictionary of the Accademia della Crusca. This was the *Dizionario della lingua italiana* of Niccolò Tommaseo and Bernardo Bellini, published from 1861 to 1879, and now searchable online along with the editions of the della Crusca dictionary.¹² Tommaseo, writer and philologist, was the most celebrated Italian lexicographer of the nineteenth century; his strong personality is visible throughout the entries of his dictionary, which is highly original compared to its predecessors. Particularly distinctive are its literary documentation, integrating citations from eighteenth- and nineteenth-century authors, and its technique: excerption of technical works by a team of collaborators, and integration of materials originating in earlier encyclopedic dictionaries. The entries are characterized above all by the direct interventions of Tommaseo, who – keen observer of the language as he was – enriched it with stylistic reflections, and with personal, and often polemical, observations. The semantic organization is highly modern, and the syntactic analysis is highly precise. The etymologies are outdated, but this deficiency is common to every work of the period. Although certain neologisms, notably in the political field, are accepted, the criteria for the selection of headwords nevertheless remain puristic.

Before 1950, highly normative monolingual dictionaries were indeed typical of Italian lexicographical production. As we saw in Chapter 14, the Accademia della Crusca had set up the Tuscan, and more precisely Florentine, language of fourteenth-century authors as the model of good usage. Even medium-sized dictionaries, which were based on living usage and not on literature (like the abridged versions of the *Novo dizionario* of Petrocchi), remained normative: the spoken language was progressively reintroduced, but Tuscan remained the standard, and specialized

¹² Online editions are available through the Accademia della Crusca webpage ‘Dizionari’.

terminology was absent.¹³ The dictionaries of neologisms which appeared from the beginning of the nineteenth century likewise took part in this tendency towards the purification of the language, indicating borrowings from foreign languages (*barbarismi*), especially French (*gallicismi*), which were introduced into Italian on a massive scale.¹⁴ Technical vocabulary was generally reserved for encyclopedic dictionaries, which appeared in increasing numbers throughout the nineteenth and twentieth centuries, in reaction to the purism of the *Vocabolario* of the Accademia della Crusca.¹⁵ One of the first synchronic usage dictionaries to be highly receptive to technical and scientific terminology would be the *Vocabolario della lingua italiana* of Nicola Zingarelli, in the first half of the twentieth century.

Another characteristic of Italian lexicography is the abundance of dialect dictionaries, which assumed a particular importance in the nineteenth century. Italy is, indeed, highly linguistically diverse. All Italians, with the exception of Tuscans, use dialect for everyday purposes and reserve standard Italian for written usage or for special circumstances. Dialect dictionaries respond to the practical necessities of translation from a dialect to supra-regional Italian, as did the non-alphabetical ('methodical') dictionaries which proliferated in the nineteenth century.¹⁶ It was only towards the end of the twentieth century that scholarly, descriptive regional dictionaries would appear in Italy, as they had in France, for instance excellent descriptions of the Italo-Romance varieties Sardinian, Friulian, and Ladin (the compiler of the last, Johannes Kramer, has also made significant contributions to papyrology, cited in Chapter 5).¹⁷

As in France, it was indeed from the second half of the twentieth century that Italian lexicography underwent a renewal. The important changes comprised new work on monolingual dictionaries and the appearance of historical and etymological dictionaries, as well as the arrival of scholarly regional dictionaries.

The standard monolingual dictionary for the Italian language, replacing that of Tommaseo and Bellini, is now the *Grande dizionario della lingua italiana* of Salvatore Battaglia (21 vols., 1961–2002). This is a true historical dictionary, a genre which had been lacking from Italian lexicography: the fifth

¹³ For the Petrocchi dictionaries, see Pfister, 'Italienische Lexikographie', 1858.

¹⁴ For *barbarismi* and *gallicismi*, see, for instance, Stocchi, *Nuovo elenco*, vi.

¹⁵ See Duro, 'Lexicographie italienne du XXe siècle', 1869–70.

¹⁶ See Zolli, 'Italienisch', 789–90, 792.

¹⁷ Zamboni et al., *Dizionario etimologico storico friulano*; Kramer, *Etymologisches Wörterbuch des Dolomitenladinischen*; Wagner, *Dizionario etimologico sardo*; Pittau, *Nuovo vocabolario della lingua sarda*; see Schweickard, 'Italian', 676–7.

edition of the *Vocabolario* of the Accademia della Crusca was definitively broken off in 1923 at the letter O, and the dictionary of Tommaseo and Bellini, although useful, did not fully meet the expected criteria of the genre, if only because of the logico-semantic ordering of senses. Besides its historical structure, providing first attestations and a chronological series of textual witnesses, Battaglia's dictionary included citations from recent and contemporary authors, so as to register the most up-to-date usage, and this was likewise an innovation in Italian lexicography. Unfortunately, the work does also have some deficiencies, and does not completely replace its predecessors, the consultation of which remains useful to linguistic specialists. As for the genre of synchronic usage dictionaries, it is now well represented, for instance by the *Grande dizionario italiano dell'uso* of Tullio De Mauro, which appeared in eight volumes at the beginning of the twenty-first century.¹⁸

Meanwhile, the Accademia della Crusca resumed its activities in 1964, after a hiatus of several years, and, abandoning the project of a sixth edition of the *Vocabolario*, decided to provide a counterpart to it: a *Tesoro della lingua italiana delle origini*, founded on a textbase of nineteen million occurrences, which constitutes a historical dictionary for the varieties of medieval Italian up to the end of the fourteenth century. Etymological lexicography has likewise sprung into life since 1950, with the notable result, besides some very good etymological dictionaries, of the launch of a *Lessico etimologico italiano* comparable to the *Französisches etymologisches Wörterbuch* of Wartburg for France, which takes into account the whole array of Italo-Romance varieties.¹⁹

Spanish

In the sixteenth and seventeenth centuries, Spain dominated the European world, making Spanish the diplomatic language of all Europe. By 'Spanish' must be understood Castilian, which had progressively extended its reach throughout Spain and to the overseas empire, as the language of royal power. The eighteenth century, by contrast, was the time of the decline of Spain as a political and economic power. The golden age of Spanish literature likewise came to its conclusion, declining at the end of the seventeenth century into the production of baroque works of inferior quality. This political and literary context encouraged a sense of linguistic decadence and corruption, which would provoke a puristic reaction against the lexical enrichment of the seventeenth century. The Real Academia Española was founded in 1713,

¹⁸ See Schweickard, 'Italian', 676. ¹⁹ See Schweickard, 'Italian', 676.

with the mission, summarized in its motto 'limpiar, fijar y dar esplendor', of re-establishing the prestige of Castilian and of codifying it.

The six volumes of the *Diccionario de la lengua castellana* of the Real Academia, which would subsequently be known as the *Diccionario de autoridades*, appeared rapidly (1726–39: only fourteen years). The Spanish academy benefited from Italian and French models (the *Vocabolario degli accademici della Crusca* and the dictionaries of Richelet, Furetière, and Trévoux) as well as from Spanish sources of high quality (namely the dictionaries of Nebrija, Covarrubias, and Oudin). The name of the *Diccionario de autoridades* is explained by the fact that, as in the *Vocabolario* of the Accademia della Crusca, the criterion for the inclusion of a word was its employment by authors (the *autoridades*, 'authorities') who were regarded as representatives of good usage. The puristic character of the dictionary is less marked than that of the *Dictionnaire de l'Académie française*, consisting primarily of the rejection of baroque coinages, of neologisms, and of Gallicisms. On the other hand, the c. 37,000 entries included words from popular language, archaisms, technical terms, and regionalisms. Spanish-language authors from the New World were among the 'authorities', and hence the dictionary contained – involuntarily – some hundreds of Americanisms. The initial character of the *Diccionario de autoridades* would unfortunately undergo damaging alteration. In 1780, the Real Academia decided to abridge the work and to suppress the citations, while increasing the number of entries to 46,000. This version of the dictionary and those that followed it would be called simply *Diccionario*, to distinguish them from the *Diccionario de autoridades*. They were on the whole more strongly puristic.²⁰ So, for instance, the Americanisms were suppressed. The editions which followed in the eighteenth and nineteenth centuries would maintain the selection criteria of the *Diccionario* of 1780. Only in the twentieth century would the work really become receptive to Americanisms and neologisms, and only with the twenty-second edition, of 2001, would there be a change in lexicographical policy and a thoroughgoing revision of method, particularly with the advent of digital corpora. The *Diccionario* is now accessible online, with other productions of the academy.

The dictionary of the Real Academia fixed the orthography of Spanish as well as its lexical and grammatical norms. By contrast with the *Dictionnaire de l'Académie française*, it enjoyed high prestige in the Spanish-speaking world, and still occupies a central position in Spanish lexicography today. It provided a model for nearly all the dictionaries which followed it, up to the second half

²⁰ See Haensch, 'Spanische Lexikographie', 1743–4.

of the twentieth century. This explains why lexicography evolved less freely in Spain than in France for a long time: as we shall see, it was only from the second half of the twentieth century that descriptive lexicography would make its appearance in Spain.

French influence was nevertheless very important in the eighteenth and nineteenth centuries, as shown by the dictionaries of Gallicisms which sought to purify the Spanish vocabulary of its numerous borrowings from French. This influence – more than the pan-European context of the Enlightenment, which would play a less important role in a Spain dominated by Catholic orthodoxy – explains why encyclopedically oriented monolingual dictionaries of Spanish started to be produced.²¹ As early as 1786, the *Diccionario castellano con las voces de ciencias y artes* of the Jesuit Esteban Terreros y Pando was published, offering some 60,000 entries, and complementing the academy dictionary. In 1846–7, the *Diccionario nacional, o gran diccionario clásico de la lengua española* of Ramón Joaquín Domínguez appeared, its title being borrowed from a French dictionary, the *Dictionnaire national* of Louis-Nicolas Bescherelle. The first dictionary which announced its encyclopedic character clearly in the title was the *Diccionario enciclopédico de la lengua española* of Eduardo Chao (1853). These last two dictionaries were remarkable for a second reason: they were the first Spanish dictionaries with titles identifying the language they treated as Spanish rather than Castilian (by comparison, the *Diccionario de la lengua castellana* of the Spanish academy only became the *Diccionario de la lengua española* in 1925, in an edition which took better account of regionalisms).

The best-known lexicographer of the nineteenth century was Vicente Salvá, who was also a grammarian and publisher. His *Nuevo diccionario* was published in Paris in the same year as the first volume of Domínguez's dictionary, but its contents were radically different. Salvá could have been the Spanish Littré, if he had distanced himself further from the academy model and cited the dates and sources of his citations. His work was distinguished by great objectivity (it reprised the whole of the academy dictionary, signalling all of its own alterations scrupulously) and by a highly modern conception of what a dictionary should be. The *Nuevo diccionario* sought to be an exhaustive inventory of all the varieties of the language: it included several different registers of the language, and numerous archaisms and regionalisms.²² The description of Americanisms was, notably, founded on the testimony of speakers of New World varieties of Spanish with whom

²¹ Haensch, 'Spanische Lexikographie', 1745–6. ²² Ezquerro, 'Spanische Lexikographie', 646.

Salvá corresponded. Salvá was thus the first Spanish lexicographer to give a conscious and deliberate account of regionalisms.

With a few exceptions (notably the beginnings in 1945 of the VOX series of dictionaries), the first half of the twentieth century saw little by way of lexicographical originality.²³ A renewal of Spanish lexicography took place after 1960, with two major innovations. The first was an increase in attention to the syntagmatic and paradigmatic elements which characterize the usage dictionary (*diccionario de uso*); the groundbreaking work in this respect was the *Diccionario de uso del español* of Maria Moliner (1966–7). The second was the appearance, starting in the 1990s, of dictionaries that were no longer founded on the wordlist of the academy dictionary but on a newly constituted corpus. Among these, the high point of monolingual Spanish lexicography was the *Diccionario del español actual* of Manuel Seco (1999). This was the first synchronic, descriptive dictionary of peninsular Spanish (it was based on the usage of the period 1955–93). It was characterized by very rich documentation, from both literary and journalistic sources; by rigorous methodology; by a consistent and reader-friendly microstructure; and also by excellent definitions.

I should note that the *Diccionario del español actual* is confined to the Spanish language as attested in Spain, and rightly so. The Spanish vocabulary of the nineteen Spanish-speaking countries of the Americas – with a total Spanish-speaking population of about 360 million, eight times as many as the 45 million speakers of Spanish who live in the Iberian peninsula – has not yet been described precisely enough for its accurate integration into general dictionaries to be possible. However, in the twentieth century the Real Academia did attempt to reintroduce to its dictionary the Americanisms that it had put aside in the nineteenth. Between 1871 and 1973, academies were created in every Spanish-speaking country, with the mandate to supply lexical materials to the Real Academia in Madrid. The original goal of these collaborations was to maintain the unity of the language rather than to describe diatopic differences precisely. Moreover, the gathering of material varied greatly from country to country, preventing a unified survey of the history and distribution of lexemes. For these reasons, the Spanish of the New World is not well represented in the editions of the academy dictionary.

Fortunately, the lexicography of these varieties is now developing strongly. Some of the milestones are as follows.²⁴ The first lists and glossaries

²³ For the VOX dictionaries, see Ezquerro, 'Spanisch: Lexikographie', 646–7, and Haensch, 'Spanische Lexikographie', 1748.

²⁴ See also Lara, *Historia mínima de la lengua española*, and Haensch, 'Spanische Lexikographie', 1751–5.

of Spanish lexemes from the New World appeared in the seventeenth century. In 1789, Antonio de Alcedo published a 'Vocabulario de las voces provinciales de América' as an appendix to a geographical work. Dictionaries of 'provincialisms' that registered the speech of one country then appeared: the first, Estéban Pichardo's *Diccionario provincial de voces cubanas*, was published in 1836. These dictionaries showed normative and prescriptive criteria: their goal was not only to describe vocabulary as a curiosity of a region (just as, in that age, geographical, botanical, and other curiosities were described) but also to gather 'barbarous' words which were not part of standard Spanish. The works of Rufino José Cuervo, a linguist from Bogotá, were an exception in this respect, but unfortunately he did not produce a general or regional dictionary.²⁵ The twentieth century saw the appearance of dictionaries of Americanisms which show the same puristic approach, but bring together the vocabulary of all the Spanish-speaking countries of the New World. One of the first of this succession of dictionaries is the *Diccionario de americanismos* of Augusto Malaret, published in 1925.

After 1970, the influence of the study of linguistic geography, and the awareness of regional variation within Spanish, led progressively towards a modernization of the lexicography of New World Spanish, which took different directions, depending on whether or not the approach remained contrastive. Notable examples of a contrastive approach were the series of dictionaries conceived as part of the 'Augsburg project' (directed by Günther Haensch and Reinhold Werner) of a *Nuevo diccionario de americanismos*, several of which have already appeared.²⁶ An approach which is original because it is not contrastive is represented in Mexico by the online *Diccionario del español de México* (a project directed by Luis Fernando Lara), based on a corpus of written and oral texts, which seeks to describe the Mexican variety of Spanish exhaustively and in its own right, without comparing it with usage beyond Mexico. Thanks to these descriptive projects, our knowledge of the New World varieties of Spanish is likely to improve considerably in the coming years.

Consideration of diatopic variation within Spanish can also be seen in etymological dictionaries, where there has likewise been progress in the twentieth and twenty-first centuries. The *Diccionario crítico etimológico castellano e hispánico* of Joan Coromines (published under the Spanish form

²⁵ See Schütz, 'Cuervos Wörterbuch'.

²⁶ Haensch and Werner (eds.), *Nuevo diccionario de argentinismos*; Haensch and Werner (eds.), *Nuevo diccionario de colombianismos*; Haensch and Werner (eds.), *Nuevo diccionario de uruguayismos*.

of his name, Juan Corominas) is an excellent example of this development: the etymological commentary takes the geographical diffusion of the words which it describes into account, from an Ibero-Romance perspective. Historical lexicography is likewise developing strongly, notable being the *Nuevo diccionario histórico* of the Real Academia, which replaces two previous unfinished dictionaries and is based on a digital corpus.²⁷

Portuguese

Portuguese can be divided into two major varieties: European and Brazilian. The lexicography of Portuguese was exclusively European until the nineteenth century. After medieval glossaries, and bilingual Renaissance dictionaries inspired by the dictionary of Nebrija, the first modern dictionaries appeared in the eighteenth century.²⁸

The first wide-ranging monolingual dictionary of Portuguese was published in Lisbon in 1789 by the Brazilian Antônio de Morais Silva, under the title *Diccionario da lingua portugueza*. This work was presented as an abridged version of an earlier bilingual dictionary, the *Vocabulario portuguez e latino* of Rafael Bluteau (1712–21). Bluteau's dictionary ran to ten volumes and combined linguistic and encyclopedic approaches. Morais Silva suppressed the Latin element, the citations, and the etymologies; he also made the definitions more concise; and thus he produced, in two volumes, a true general dictionary of the language. Further editions followed, with numerous additions. The dictionary of Morais would continue to be a standard work until the end of the nineteenth century, together with two other dictionaries of high quality which would likewise appear in several further editions in the course of the nineteenth and twentieth centuries: the *Diccionario contemporaneo* of Francisco Júlio de Caldas Aulete, distinguished by the precision of its semantic organization, and the *Novo diccionario* of Cândido de Figueiredo, the c. 110,000 entries of which bring together literary vocabulary, terms from popular language, regionalisms, and Brazilianisms, together with scientific and technical terms. The *Diccionario da lingua portugueza* of Morais became the exemplar of normative lexicography with a historical quality, as academy dictionaries had become in other countries. The tenth edition (1949–59), of no fewer than twelve volumes, was directed by the lexicographer and

²⁷ See Reinhardt, 'Iberoromance I', 637, and, for a comprehensive overview of recent Spanish lexicography, Haensch and Omeñaca, *Diccionarios del español*.

²⁸ For the lexicography of Portuguese, see Verdelho, 'Portugiesisch: Lexikographie'; Woll, 'Portugiesische Lexikographie'; and Kiesler, 'Iberoromance II'.

etymologist Pedro Machado and diverged a lot from the original. It is still much cited today: it is particularly remarkable for the richness of its vocabulary (306,949 entries) and for its systematically provided citations, taken alike from modern Portuguese and Brazilian authors.

Among the institutional lexicography of Portugal is a singularity: the Academia Real das Ciências de Lisboa, founded in 1780, produced only one solitary volume of its *Diccionario da lingua portugueza*, in 1793, for the letter A. This volume would be republished by the same academy nearly two centuries later, in 1976. While it showed similarities to its Spanish and Italian counterparts – a normative character, and citations from authors from the fourteenth century to the end of the seventeenth – the work of the Portuguese academy had numerous distinctive qualities, notably that it took better account of the different registers of the language.²⁹ Only in 2001 would the academy produce its *Dicionário da língua portuguesa contemporânea*, which is a synchronic dictionary.³⁰ The absence for two centuries of any academy dictionary was compensated for by the successive editions of the dictionary of Moraes, and by other general dictionaries, which were, in the nineteenth century, often inspired by French works.³¹ The academy had a greater impact in the nineteenth century with a specialized work, its *Vocabulário ortográfico*, which would, after the first orthographic reform of 1911, have a strong influence on the regularization of the orthography of Portuguese.³²

It was only at the end of the nineteenth century that the vocabulary of Brazilian Portuguese began to be an object of lexicographical study in its own right. There were two approaches, both contrastive. First, in the first quarter of the nineteenth century, dictionaries of Brazilianisms started to appear, showing an interest in popular language use and in the vocabulary of everyday life. This sort of dictionary would be produced until the beginning of the twentieth century and the dominance of a new approach: the inclusion of Brazilianisms, labelled as such, in general dictionaries of the Portuguese language. Among these, the third edition of the above-mentioned dictionary of Aulete stands out; published in the middle of the twentieth century, it was innovative in that it cited the Brazilian literature of the 1930s and 1940s. This dictionary was highly successful in Brazil.

Another successful Brazilian dictionary, edited for the first time in 1975, is the *Novo dicionário da língua portuguesa*, written by the Brazilian Aurélio

²⁹ Woll, 'Portugiesische Lexikographie', 1727. ³⁰ See Kiesler, 'Iberoromance II', 649.

³¹ See Verdelho, 'Portugiesisch: Lexikographie', 679–85.

³² Verdelho, 'Portugiesisch: Lexikographie', 686.

Buarque de Holanda and better known as 'Aurélio'. It is composed of 115,243 articles, with many citations coming from literary, journalistic, and oral corpora.

Brazil was the country of origin of the synchronic dictionary which is now regarded as the most complete language-oriented dictionary of Portuguese.³³ This is the *Grande dicionário Houaiss da língua portuguesa*, published by the Antônio Houaiss institute of lexicography ('2001). This is a synchronic dictionary, but it is also recognized as an indispensable historical and etymological reference.³⁴ There is still no comprehensive historical dictionary of Portuguese, but the language of the thirteenth, fourteenth, and fifteenth centuries is well documented in the *Índice do vocabulário do português medieval* of Antônio Geraldo da Cunha, complemented by his *Vocabulário histórico-crônológico do português medieval*, which appeared in digital form in 2006. Overall, the specialized lexicography of Portuguese is still modest in comparison with what has been produced for the other major Romance languages.

Romanian

The earliest written records of the Romanian language are from the sixteenth century. In that period, Church Slavic, the language of the clergy, was the dominant written language where Romanian was spoken. Western European humanism made its way into Romania in the seventeenth century, and the first glossaries appeared from 1650 onwards. In them, Romanian was first presented with Latin, Italian, or Slavic language varieties; it was subsequently presented with German, Hungarian, Russian, Greek, or Turkish. So, the history of Romanian lexicography is at first bilingual or plurilingual.³⁵

This history is bound up with that of the emancipation of the written language, which took place with the progressive borrowing from foreign languages of the words which Romanian lacked. The beginning of a cultural renaissance took place in Transylvania at the end of the eighteenth century: the Latin origin of the Romanian language was shown, and the necessity of enriching its vocabulary with neologisms taken from other Romance languages was insisted upon. In 1704, the historian Dimitrie Cantemir had already provided, as an appendix to his novel *Istoria ieroglifica*, a glossary of

³³ Benarroch, 'L'apport des dictionnaires', 623.

³⁴ It is, for instance, regularly cited in Buchi and Schweickard, *Dictionnaire étymologique roman*.

³⁵ See Miron, 'Rümanische Lexikographie'; Winkelmann, 'Rumänisch: Lexikographie'; Ernst, 'Romanian'; Rosetti, *Brève histoire de la langue roumaine*; Seche, 'Schită'.

the words of foreign origin which he had used in it. The enrichment of the language, together with the need to unify and codify it, became the major preoccupation of nineteenth-century literati, and was reflected in lexicographical works undertaken by what came to be called the Transylvanian School. In 1825, the best-known lexicographical work of the period, the Romanian–Latin–Hungarian–German *Lesicon românescu-lătinescu-ungurescu-nemțescu*, was published at Buda (now part of Budapest); hence it is also known as *Lexiconul Budan*. This monumental work, which would be used by many scholars up to the present day, includes a great number of new words.

The lexicographical works of the period are sometimes appended to grammatical or historical treatises, and are often marked by didacticism, in line with the Enlightenment ideas in circulation at the time. In the first half of the nineteenth century, Romanian became the official language of education. New bilingual dictionaries appeared, often with French or German as the second language. Between 1895 and 1925, the *Rumänisch-deutsches Wörterbuch* of Heimann Hariton Tiktin appeared. Unlike other bilingual dictionaries, words were not translated or glossed, but explained following the model of monolingual dictionaries. It was a descriptively and historically oriented work, with 40,000 entries, supplied with reliable bibliographical, grammatical, and etymological information, making it one of the classic dictionaries of the Romanian language. It is still used today, in its third edition.

The story of the monolingual dictionaries of Romanian begins in the nineteenth century, as the literary language was enriched with technical and scientific terms under the influence of the French encyclopedists. After numerous unpublished or unfinished monolingual dictionary projects, dictionaries of foreign terms and of technical terms appeared after the so-called Romanian discovery of Europe in 1821. These dictionaries, in small or medium format, covered words which had entered the daily life of the Romanian people, together with terms of foreign origin for plants, animals, and institutions. Larousse-style encyclopedic dictionaries circulated from the end of the nineteenth century with the *Dicționarul universal* of Lăzar Șăineanu and its successive editions (among which may be placed that of Ioan Oprea, which is now the standard reference). As for learners' dictionaries, they would appear in the twentieth century.

In 1866, a Romanian academy was officially founded, on western European models, with the task of producing a dictionary of the language. This academy dictionary, the *Dicționarul limbii române*, would

eventually be produced in two major stages. The first of these saw the publication, between 1913 and 1949, of five volumes covering roughly the first half of the alphabet. The remaining letters were covered in fourteen volumes edited as *Dicționarul limbii române: serie nouă* between 1965 and 2010. There are a number of differences between the two moieties, especially with reference to their alphabetical ordering, modified as and when orthographic reforms took place. The ensemble runs to about 175,000 entries.

The *Dicționarul limbii române* has provided the base for subsequent synchronic dictionaries. In the second half of the twentieth century, while the editing of the academy dictionary was in progress, various projects were undertaken at Bucharest, Iași, and Cluj, leading to the appearance of dictionaries in small and medium format. Apart from the *Micul dicționar academic*, which is an attempt to broaden the readership of the academy dictionary by presenting its contents in a more compact format, the *Dicționarul explicativ* has been particularly successful, fulfilling some of the same functions for Romanian that the *Petit Robert* has done for French.

The story of the beginnings of the Romanian language explains why etymological work has had a particularly privileged place in Romanian specialized lexicography. An interest in the etymology of Romanian has been evident since the first lexicographical undertakings of the eighteenth century. The first etymological dictionary of Romanian was Alexandru Cihac's *Dictionnaire d'étymologie daco-romane* (1870–9). It classifies about 17,645 Romanian words according to their language of origin (Latin, Slavic, Turkish, modern Greek, Hungarian, or Albanian), giving the impression – a mistaken one – of a great quantity of Slavic elements. Several other works followed, including a *Dicționarul etimologic* by Ion Aurel Candrea and Ovid Densușianu in 1907–14; a *Diccionario etimológico rumano* edited by Alejandro Cioranescu in the second half of the twentieth century; and a *Dicționarul etimologic* in course of publication by the Romanian academy. But Romanian still lacks a large etymological dictionary on the scale of the *FEW* or the *Lessico etimologico italiano*. On account of the richness of its documentation, the aforementioned dictionary of Tiktin continues, together with the academy dictionaries, to be a resource for the historical and philological lexicography of Romanian.

Other Romance Languages

A great number of languages regarded as minor have not been described in the preceding sections of this chapter. I have not, for example, mentioned the

languages of the south of the Gallo-Romance linguistic domain (the '*langues d'oc*'), namely Occitan, Francoprovençal, and Gascon. In fact, the lexicography of Occitan stands out in comparison with that of the other Romance languages, on account of a sharp dichotomy between the two stages of the language, Old Occitan and modern Occitan. Old Occitan, like Old Gascon, has been the object of lexicographical descriptions at a high scholarly level, beginning in the nineteenth century with the *Lexique roman* of François Raynouard (followed at the end of the century by an eight-volume supplement by Emil Levy), and continuing in the twentieth century with, notably, an online *Dictionnaire de l'occitan médiéval*.³⁶ Modern Occitan, by comparison, seems to have been forgotten by scholarly Romance lexicography. Efforts were made to codify it from the middle of the nineteenth century, but did not result in a consensus.³⁷ A central part – but also a strongly criticized one – in this story was played by the *Trésor dou Félibrige* of Frédéric Mistral, which sought to bring all the varieties of the *langues d'oc* together in the embrace of a revitalized Rhodanian dialect.³⁸ As for Francoprovençal, its Helvetic varieties are being described in the *Glossaire des patois de la Suisse romande*, which is an excellent descriptive, synchronic dictionary.

In the Ibero-Romance linguistic domain, something should be said about Catalan, the medieval culture of which was very rich. The distinctive feature of its linguistic history is a permanent struggle with Castilian, the language of power in Spain. As with the other Romance languages, the lexicography of Catalan was at first bilingual, the second languages being Latin and, from the end of the eighteenth century, Castilian. The history of the relations between Catalan and Castilian (marked by long periods in which the former was excluded from literary and cultural life by the power of the Spanish state), together with the permanent situation of bilingualism of the speakers of Catalan, goes a long way to explain why Catalan lexicography still lacked monolingual dictionaries at the end of the nineteenth century. In 1907, the Institut d'Estudis Catalans was founded, and in 1931–2 the *Diccionari general de la llengua catalana* of Pompeu Fabra appeared. For a long time, it played the role of normative standard reference, alongside certain encyclopedic dictionaries.³⁹ In the twentieth century, major descriptive lexicographical

³⁶ See Buchi and Renders, 'Gallo-Romance I', 658–9.

³⁷ See Kramer, 'Lexikographie des Provenzalischen, Rätoromanischen, Sardischen und Dalmatischen', 1891–5; Trotter, 'Gallo-Romance II', 669; Schlieben-Lange, 'Okzitanisch', 122–4.

³⁸ See Chambon, 'Brèves remarques'.

³⁹ See Haensch, 'Katalanische Lexikographie', 1775, and Rico and Solà, 'Katalanisch: Lexikographie'.

works were produced. Two historically oriented dictionaries stand out among them: the *Diccionari etimològic i complementari de la llengua catalana* of Joan Coromines (1980) and, above all, the famous *Diccionari català-valencià-balear* in ten volumes, which has been online since 2002.⁴⁰

I cannot give an account here of all the lexicographical records of all the language varieties of the Romance linguistic domain, although many of them are of high quality. I could, for example, have said something about the Romance creoles, the lexicographical history of which is interesting.⁴¹ I could likewise have mentioned the Walloon language varieties of Romance-speaking Belgium, which have been the object of excellent descriptions.⁴² The reader interested in Galician, in Sardinian, or again in Friulian and Romansh in the Italo-Romance linguistic domain will find information about their lexicographical histories elsewhere.⁴³

Conclusion

The survey which I have presented here is very brief, and has only cited the major products of the lexicographical history of the principal Romance languages. I have, for example, chosen to leave bilingual lexicography on the margins, although since its medieval origins it has been an important part of European dictionary production. I have likewise said little about specialized language dictionaries, such as – to name only two varieties – dictionaries of synonyms or of proverbs. Despite these omissions, this survey allows us to fill in the outlines which I sketched at the beginning of this chapter, and to take up certain differences between the traditions of the different countries.

We have seen that the principal Romance languages, namely French, Spanish, Italian, Portuguese, and Romanian, have all been the objects of institutional lexicography overseen by an academy, with the exception of Portuguese, where this normative lexicography was represented by the dictionary of Moraes rather than by the incomplete academy dictionary. This prescriptive approach contributed everywhere to fixing the lexical norm, in a pan-European context of the establishment of national languages. It was,

⁴⁰ See Kiesler, 'Iberoromance II', 650–1.

⁴¹ See Bollée, 'Lexicographie créole', and Bollée, *Dictionnaire étymologique des créoles français de l'Océan Indien*.

⁴² Remacle, *Atlas linguistique de la Wallonie*.

⁴³ See Kramer, 'Lexikographie des Provenzalischen, Rätomanischen, Sardischen, und Dalmatischen'; Dettori, 'Sardisch'; Marcato, 'Friaulisch'; Kramer, 'Ladinisch'; Pensado, 'Galician lexicography'; see also references above to the major languages of the relevant linguistic domains.

however, effected in different countries at very different moments in their linguistic and literary histories. Whereas in France the prestige of a language at its classical height was being settled, the business of the Spanish academy was, on the contrary, the purification of a language perceived as corrupted and decadent. In Romania, the situation was quite different again: there, lexicographers were participating in the codification of a language which was actively developing. But everywhere, the undertaking was very selective. Following the dictionaries produced by these academies, the lexicographical production in most European countries remained normative for a long time.

During these three centuries, and particularly in the nineteenth, France was in the vanguard of dictionary production. The leading position of French is especially evident in the field of historical lexicography: the periods of its linguistic history are well described today, whereas, by contrast, Portuguese and Romanian still lack exhaustive historical descriptions of high quality.⁴⁴ A pan-Romance project such as Buchi and Schweickard's *Dictionnaire étymologique roman* therefore plays an important role in compensating for certain delays. The study of the dictionaries of different languages also shows that the boundary between synchronic and diachronic lexicography is seldom well defined: a historically oriented work may be used as a synchronic dictionary, while a synchronic work is sometimes the only source of historical or etymological information.

A notable difference between the principal Romance lexicographical traditions is the way in which they manage transatlantic diatopical variation. French, Spanish, and Portuguese are, like English, colonial languages; each of their traditions has its own experience of the problematic (still current) of the description of regional varieties in relation to a centre which was established as a point of reference in the eighteenth and nineteenth centuries. Like English (see Chapter 24), these languages have recently seen the appearance of dictionaries conceived and realized in the Americas to give descriptions – always with methodological differences from each other – of these languages of European origin which are more centred in the New World. Portuguese, for example, has privileged a differential lexicography which integrates Brazilianisms into the fabric of general dictionaries, whereas Spanish and French at first privileged dictionaries of regionalisms, before venturing very recently along the path of a non-differential lexicography, making a comprehensive description of the French of Quebec or the Spanish of Mexico.

⁴⁴ See Groß, 'Les dictionnaires étymologiques et historiques', 91, for an overview of Romance lexicography from a historical perspective.

PART IV



THE MODERN WORLD:
MISSIONARY AND
SUBSEQUENT TRADITIONS

Missionary Traditions in South America

OTTO ZWARTJES

The story of missionary lexicography in South America begins in the sixteenth century, with the Spanish–Quechua and Quechua–Spanish *Lexicon, o Vocabulario de la lengua general del Peru* of the Dominican friar Domingo de Santo Tomás, published in 1560; its first section contains about 6,000 entries, and its second about 4,000.¹ Quechua had been the language of the Inca empire, and continued in widespread use after the fall of that empire, so that it was an important subject for dictionaries. The work of Santo Tomás was followed in 1586 by an anonymous *Arte y vocabulario en la lengua general del Peru*, presenting adaptations of his two wordlists (about 5,000 entries each) followed by a grammar. Further editions of the anonymous work appeared, for instance one of 1604; this was followed in its turn by the Jesuit Diego González Holguín's *Vocabulario de la lengua general de todo el Peru llamada lengua qquichua o del inca* in 1608 (about 13,000 entries in both sections), and by his confrere Diego de Torres Rubio's *Arte de la lengua quichua* of 1619, a grammar which includes bidirectional wordlists.²

Aymara, the language of a people of the Andes who had been subjects of the Inca, was documented a little later. Quechua and Aymara wordlists appeared in an anonymous *Doctrina Christiana y catecismo para instruccion de los Indios* of 1583–4, but the first dictionary of Aymara was the *Vocabulario de la lengua aymara* of the Jesuit Ludovico Bertonio, which was published in 1612 (9,000 entries), and was followed in 1616 by wordlists in the *Arte de la lengua aymara* of Torres Rubio.³

¹ For entry counts of the Quechua dictionaries, see Calvo Pérez, 'El concepto de diccionario bilingüe', 266–7, and Segovia Gordillo, 'El vocabulario hispano–quechua', 21; for these dictionaries, see also Niederehe, 'Grammatiken und Wörterbücher'.

² The wordlists are Torres Rubio, *Arte de la lengua quichua*, sigs. M2r–R4r (Spanish–Quechua), S1r–X3v (Quechua–Spanish).

³ Entry count for Bertonio's dictionary by E. Ridruejo in Valdivia, *Arte y gramatica general* (2007), II.150; the wordlists of 1616 are Torres Rubio, *Arte de la lengua aymara*, fos. 46r–64v (Spanish–Aymara), 81r–97r (Aymara–Spanish).

In the Andean region, the most important works on the indigenous languages were produced in the period from 1560 to 1650. According to Julio Calvo Pérez, the period which follows was one of ‘general decay’ until the nineteenth century, when scholarly interest in indigenous languages was resuscitated.⁴ As occurred in New Spain, where grammarians and lexicographers did not always continue describing and teaching the prestigious variety of the Central Valley, in the Andes works were produced after 1650 describing regional varieties of Quechua, or even different languages, such as Lule and Xebero.⁵ So, for instance, Juan de Figueredo’s new edition of the Aymara *Arte* of Torres Rubio, published in 1700, adds a wordlist of the Chinchaysuyo variety of Quechua, and Juan de Velasco’s eighteenth-century Spanish–Quechua *Vocabulario de la lengua índica* describes the variety of Quechua spoken in Ecuador, in 2,998 single-word entries, with a few phrases.⁶ Similar wordlists were compiled elsewhere in the Andes, such as the trilingual Spanish–Quechua–Xebero *Vocabulario de la lengua xebera*.⁷

In what is now Colombia, I can mention works which were produced during the seventeenth and eighteenth centuries, such as the seventeenth-century *Gramática, vocabulario, catecismo, i confesionario* of the Chibcha language, containing 3,700 Spanish–Chibcha entries, and Alonso de Neira and Juan Rivero’s Achagua grammar and dictionary, begun no later than 1703 and extant in a manuscript copied in 1762.⁸ Many other wordlists could be mentioned. From what is now Chile, for instance, we have Luis de Valdivia’s and Andrés de Febrès’ of Mapudungun, printed in 1606 and 1765 respectively, the former containing 2,800 entries.⁹ Valdivia also made wordlists, of about 800 entries, of the now-extinct languages Allentiac and Millcayac.¹⁰ An important collection of manuscript wordlists of languages such as Andaqui, Carib, Lokono, Otomaco, Siona, and Warao was gathered

⁴ Calvo Pérez, ‘El vocabulario de Velasco’, 33.

⁵ Calvo Pérez, ‘El vocabulario de Velasco’, 34.

⁶ Velasco, *Vocabulario* (1964), is an edition; see also *Lengua de Maynas* (2014), 73.

⁷ *Vocabulario de la lengua xebera* (2016) is an edition; see also *Eighteenth Century Xebero* (2016).

⁸ *Gramática, vocabulario, catecismo, i confesionario de la lengua chibcha* (1871) is an edition of the Chibcha dictionary; see also *Diccionario y gramática chibcha* (1987), 30; other Chibcha wordlists are Madrid, Biblioteca del Palacio Real, MSS 2922–4, of which transcripts are published online as part of *Colección Mutis*. An edition of de Neira and Rivero’s ‘Arte y Vocabulario de la lengua achagua’ is *Lenguas de América* (1928), 1–174.

⁹ Valdivia, *Arte y gramática general*, sigs. G1r–L8r; Febrès, *Arte de la lengua general del Reyno de Chile*, 195–414 (Spanish–Mapudungun), 415–682 (Mapudungun–Spanish).

¹⁰ They are part of his *Doctrina christiana, catecismo y confesionario en las dos lenguas mas generales que corren en la Prouincia de Cuyo*, printed in 1607 and until recently known only in fragments such as Valdivia, *Doctrina Christiana y catecismo en la lengua allentiac*, and Valdivia, *Vocabulario breve en lengua millcayac*; see Cancino Cabello, ‘Los tratados millcayac y allentiac (1607) de Luis de Valdivia’.

by Celestino Mutis in the late eighteenth century, and is now in the Biblioteca del Palacio Real in Madrid.¹¹ There are at present no systematic comparative studies of these minor lexicographical works.

Another major achievement by a Spanish-speaking missionary was the work on Guaraní undertaken by Antonio Ruiz de Montoya.¹² His Guaraní–Spanish *Tesoro de la lengua Guaraní*, which presents about 5,500 elaborate entries in just over 800 pages, was published in 1639, followed in 1640 by his *Arte y vocabulario de la lengua Guaraní*, which is a grammar with a monodirectional Spanish–Guaraní dictionary. Probably the publisher wanted to publish the *Arte* in one volume and the dictionary in two more, one for the lemmata A–F (376 pages), and a second for the letters F–Z (234 pages), since after page 376, the numbering of pages starts again. Montoya's work is notable for its detailed morphological analysis and for its rich provision of examples and grammatical information. The Spanish–Guaraní dictionary was published in a revised and augmented edition by another Jesuit, who remained anonymous, in 1722; the grammar and the Guaraní–Spanish dictionary followed in 1724, the editorial work now being ascribed to Pablo Restivo, and were published in a Latin translation as Restivo's own work in 1892.

Almost no colonial dictionary of a Brazilian indigenous language other than Tupi – the 'língua geral' ('general language'), as it was called – has survived. It has been documented that Luis Vincencio Mamiani, the author of a Kiriri grammar and catechism, composed a dictionary titled *Vocabulario kiriri* in 1696, but it has never been found.¹³ As for Tupi itself, the source for all other dictionaries was without any doubt the anonymous *Vocabulário na língua brasileira*, of 9,300 Portuguese–Tupi entries, written in 1621 or 1622, probably in the São Paulo region. It must have existed in numerous copies in all the mission stations.¹⁴ Some of the headwords are Portuguese phrases, translated into Tupi with some grammatical notes. This information is often accompanied by some diaevaluative remarks, such as that one Tupi form is

¹¹ Transcripts of the Mutis manuscripts are published online in *Colección Mutis*, and some of them have been edited in *Lenguas de América* (1928), namely Madrid, Biblioteca del Palacio Real, MS 2912 (Andaqui, *Lenguas de América*, 175–95), MS 2913 (Lokono, *Lenguas de América*, 197–212), MS 2914 (Carib, *Lenguas de América*, 213–305), MS 2915 (Siona, *Lenguas de América*, 307–79), MS 2920 (Warao, *Lenguas de América*, 441–52). See Hernández, *Lexicografía hispano-amerindia 1550–1800*, 163–4.

¹² Dietrich, 'Lexikographie des Tupi-Guaraní', 2673–4.

¹³ Fonseca, *Historiografia lingüística portuguesa e missionária*, 338; for the grammar, see Zwartjes, *Portuguese Missionary Grammars*, 175–201.

¹⁴ *Vocabulário na língua brasileira* (1952–3) is an edition.

preferable to another.¹⁵ Some entries are surprisingly extensive, such as *cobra* ‘snake’, for which the generic Tupi equivalent *boya* is given, followed by information on a number of varieties, such as snakes which kill, coloured snakes, and so on.¹⁶ There are ten anonymous Tupi dictionaries, none of which were printed during the colonial period; some of these were composed by German Jesuits, such as Anton Meisterburg, Anselm Eckart, and Rochus Hundertpfundt. Unlike lexicographers working on other South American languages, their compilers often indicated the category or part of speech of the lemma, basing their information mainly on the Tupi grammar of Anchieta.¹⁷ Their titles and word counts differ: for instance, *Caderno da Lingua* contains 3,881 entries, and *Prosodia de Lingua* contains 5,480 entries.¹⁸ A manuscript wordlist of 1756 discovered at Trier by Jean-Claude Muller in 2012 is bidirectional: the first section contains 6,095 Portuguese–Tupi entries and its counterpart contains about 2,500. It is notable that the second part is arranged according to the final syllables.¹⁹ A Portuguese–Tupi dictionary printed in 1795 as *Diccionario Portuguez, e Brasiliano*, and possibly authored by a Franciscan known as Frei Onofre, was originally also bidirectional, and the Tupi–Portuguese part was printed in the twentieth century.²⁰

The anonymous MS 69 of the library of the University of Coimbra includes a grammar and a monodirectional Portuguese–Tupi wordlist.²¹ The wordlist is not as elaborate as that of the *Vocabulário na língua Brasileira*: the word *cobra*, for instance, is simply translated ‘boya’ with no further information.²² Usually, no grammatical information about the Tupi equivalents is given, but in the Portuguese columns there are sporadic grammatical notes: *huy*, for instance, is identified as an interjection.²³ There are other wordlists in the manuscript; for instance, the grammar has an index containing 286 entries, in each of which the corresponding part of speech is marked, followed by a page number leading the user to the section where the word is explained. Compared with Ruiz de Montoya, who worked in the same period on Guaraní, a related language, the lexicographers of Tupi offered less rich and sophisticated lexical analysis. Nevertheless, the Brazilian corpus is

¹⁵ *Vocabulário na língua brasileira* (1952–3), 21, ‘he mais proprio’.

¹⁶ *Vocabulário na língua brasileira* (1952–3), 76.

¹⁷ For Anchieta, see Zwartjes, *Portuguese Missionary Grammars*, 148–63.

¹⁸ Dietrich, ‘O conceito de “Língua Geral”’, 597–8.

¹⁹ Dietrich, ‘O conceito de “Língua Geral”’, 598–9.

²⁰ The edition is *Diccionario portuguez-brasiliano e brasiliano-portuguez* (1934).

²¹ For the grammar, see Zwartjes, *Portuguese Missionary Grammars*, 168–75.

²² ‘Diccionario da Lingua Geral do Brazil’, 277.

²³ ‘Diccionario da Lingua Geral do Brazil’, 303.

important for its content, especially the documentation of terms related to flora, fauna, and other cultural-specific matters of the time.²⁴

Fewer than a handful of grammars and dictionaries describing Cariban languages were ever printed and few manuscripts survived. (The *Dictionnaire caraibe-francois* of Raymond Breton, published in 1665, actually documents an Arawakan language which he had learned in the course of missionary activity in Dominica.) Pierre Pelleprat's grammar of Carib, published in 1655, contains a brief vocabulary arranged in thematic order, though not a very systematic one.²⁵ The canonical order in such works is first spiritual matters, followed by human beings, mammals, birds, insects, and finally objects. Here, by contrast, the order seems quite random: words for heavenly bodies, wind, and sunlight; the elements; metals; kinship terms; parts of the body; and so on, with a penultimate section of words which do not fall under any of the previous headings. Like many other similar works in other regions, the work has a final section devoted to manners of speaking, titled 'Quelques verbes et façons de parler assez ordinaires'. Pelleprat's work, together with other early Carib vocabularies, works of natural history, and travel accounts, was a source for the *Dictionnaire galibi* of the agronomist Simon Philibert de la Salle de l'Etang, published in 1763, with the intention that it should be used in the ill-fated French colony which was about to be established in Guiana.²⁶ Salle de l'Etang divided the French–Carib part of his dictionary into sections: general vocabulary excluding verbs; names of animals, birds, fish, and reptiles; names of plants. The Carib–French part is in a single alphabetical sequence.

Cumanagoto, a Cariban language spoken in what is now Venezuela, was documented in Francisco de Tauste's *Arte y vocabulario de la lengua de los indios chaymas, cumanagotos, cores, parias, y otros diversos* (1680), which offers a Spanish–Cumanagoto vocabulary of about 1,500 entries, and in a vocabulary included in Matías Ruiz Blanco's *Conversion de Piritu*, which differs from any other source analysed in this chapter in its layout: it presents an alphabetized Spanish wordlist with Cumanagoto equivalents, but the Spanish headwords are printed to the right of the Cumanagoto words.²⁷

The dictionaries to which this chapter will attend most closely are González Holguín's dictionary of Quechua, Ludovico Bertonio's dictionary of Aymara, and Antonio Ruiz de Montoya's two dictionaries of Guaraní, all composed in the first half of the seventeenth century.

²⁴ Dietrich, 'Lexikographie des Tupi-Guaraní', 2671.

²⁵ Pelleprat, *Introduction a la langue des Galibis*, 15–30.

²⁶ The sources are discussed at Salle de l'Etang, *Dictionnaire galibi*, iii–ix.

²⁷ Ruiz Blanco, *Conversion de Piritu*, 47–250 (second sequence of pagination).

Intended Readership

The target readers of these missionary language tools were missionaries who had come from Europe and who had not yet mastered the relevant local language, or missionaries who had already acquired some basic knowledge of a local language and wanted to improve their skills.

The authors often developed special material for beginners, such as the *Vocabulario* of Torres Rubio, or for the more advanced learners. Generally, the works are not written as language tools for learning Spanish, although we sometimes find such objectives in the prologues. Santo Tomás identifies his target groups in his prologue: his Spanish–Quechua section is mainly written for those who know Spanish and not Quechua, whereas his Quechua–Spanish section is written both for Spanish-speakers to understand Quechua and for Quechua-speakers to learn Spanish.²⁸ In Bertonio's prologue to his Aymara dictionary we find a similar observation about the indigenous people as learners of Spanish: 'If the grammar of the Spanish language were easy to explain, I would put some rules for the Indians who want to benefit from this vocabulary.'²⁹ Such objectives might be too idealistic, since there is no evidence that these works were in fact also used by the indigenous population in order to learn Spanish, although some individuals, the so-called *ladinos*, might have been interested in these works.

Place of Publication

Although the wordlists discussed in this chapter were meant for use where the languages which they documented were spoken, quite a few of them were – in contrast with wordlists of the languages of the Philippines and New Spain – not printed locally but in Europe. Likewise, unlike what happened in Japan and the Indian subcontinent, where the Portuguese Jesuits established their own printing facilities, there was no local press in Brazil during the first centuries of colonization. So, for instance, Santo Tomás' dictionary of Quechua was printed in Spain at Valladolid in 1560, Bertonio's first two Aymara grammars were printed in Rome in 1603, and Pelleprat and Breton's seventeenth-century Carib and Arawakan wordlists were printed, respectively, in Paris and Auxerre. A number of other works were printed in

²⁸ Santo Tomás, *Lexicon*, sig. 5v, 'por que el que sabe la de España, y no la dellos [i.e., de los Indios] . . . porque el q[ue] la sabe, y no la de española'.

²⁹ Bertonio, *Vocabulario de la lengua aymara*, sig. B2r, 'Si la gramatica dela le[n]gua española fuera facil de dar a ente[n]der, pusiera algunas reglas para los indios que quieren aprouecharse de este vocabulario.'

Madrid: Ruiz de Montoya's dictionaries of Guarani in 1639 and 1640, Tauste's work on Cumanagoto in 1680, Ruiz Blanco's *Conversión de Piritú* in 1690, the anonymous work on Moxo in 1699, and Machoni's grammar and dictionary of Lule in 1732. Works on Andean languages might be published in Juli, as were Bertonio's Aymara dictionary and grammar of 1612, or, more often, in Lima (then called Ciudad de los Reyes), as were the anonymous *Arte y vocabulario en la lengua general del Peru* of 1586 and González Holguín's *Vocabulario* of 1608, and the Mapudungun wordlists published by Valdivia in 1607 and by Andrés Febrès in 1764. The most important publishers in Lima were Francisco del Canto and Antonio Ricardo. The latter had lived in New Spain, where he printed Córdova's Spanish–Zapotec dictionary in Mexico City in 1578.

Sources

Most of the missionary lexicographers of the Spanish tradition followed the Spanish–Latin dictionary of Antonio de Nebrija (c. 1495; see Chapter 14), either directly, in one of its many editions, or indirectly, by using dictionaries of New World languages which were themselves derived from it. Most authors do not mention Nebrija's name explicitly, but Santo Tomás, who probably used an edition of Nebrija of 1516, states that his vocabulary goes by the same order as that of Antonio de Nebrija, in other words, alphabetically.³⁰ The *Arte y vocabulario en la lengua general del Peru* of 1586 was naturally modelled on the work of Santo Tomás, but as Byron Ellsworth Hamann has shown in his book, *The Translations of Nebrija*, the Spanish–Quechua entries include some terms which were added to Nebrija's dictionary in an edition of 1578.³¹ González Holguín's *Vocabulario* is, according to Hamann, an expansion of the 1604 edition of the *Arte y vocabulario en la lengua general del Peru*, to which 'González Holguín also added entirely new categories, such as five delicious entries for types of potatoes.'³² Bertonio's *Vocabulario de la lengua aymara* was in its turn modelled on González Holguín's dictionary of Quechua, and the dictionary of Aymara compiled by Diego de Torres Rubio was mainly based on Bertonio's. A shortcoming of Bertonio's dictionary is that he takes his Aymara examples from liturgical texts which were not the work of native speakers of Aymara, so that

³⁰ Santo Tomás, *Lexicon*, sig. ✕5v, 'este vocabulario va por el mismo orden que el de Antonio de Nebrissa por el alphabeto'; for the edition he used, see Hamann, *Translations of Nebrija*, 55.

³¹ Hamann, *Translations of Nebrija*, 71. ³² Hamann, *Translations of Nebrija*, 79.

there are serious distortions and ecclesiastical impositions of meaning ignoring Aymara categories. For example, he gives *wawachatha* as the word meaning 'to give birth', a term used only for animals, never for humans, thus violating one of the fundamental linguistic postulates of the language . . . The actual form of the word, as cited by Bertonio, is grammatically impossible.³³

Western models for dictionaries which translate from Spanish into an indigenous language are easier to trace than models for the sections or separate dictionaries which translate from the indigenous language to Spanish. For instance, Luis de Valdivia decided to publish a Mapudungun–Spanish vocabulary without first compiling and reversing a Spanish–Mapudungun counterpart, and no European models for it can be traced; probably it is a creation of his own. He informs his readers that it contained only the most essential words which missionaries would need, and that he had the plan of writing a much more copious Spanish–Mapudungun volume.³⁴ Likewise, in compiling his Mapudungun–Spanish wordlist, Andrés Febrès obviously started with lemmata in Mapudungun. He observes that he explains them in considerable detail, often providing several synonyms or near-synonyms in Spanish, which proves that he did not simply make his dictionary by reversing material from the Nebrija tradition.³⁵

Another work which does not seem to have been based on any European model is Francisco de Tauste's *Arte y vocabulario de la lengua de los indios chaymas, cumanagotos, cores, parias, y otros diversos*. Like the anonymous *Arte y vocabulario en la lengua general del Peru*, Tauste's work places the vocabulary before the grammar (in both cases, one might expect the *arte* 'grammar' to come first in the book as it does in the title), preceded only by brief remarks on linguistic variation, letters which are not used to write the Cumanagoto language, and speaking manners and styles. As Juan Pedro Sánchez Méndez has pointed out, the two-thirds of Tauste's book that consists of words and phrases seem more like the author's notebook made for his own use than a formal alphabetized dictionary such as the great works produced in the Andes.³⁶ Probably the printing of the work had to be done in a hurry, leaving no time to revise the book and thus to impose a clearer structure on it.

³³ Hardman-de-Bautista, 'Aymara lexicography', 2685.

³⁴ Valdivia, *Arte y gramática general*, sig. Gir, 'todo lo necessario y suficiente que ha menester vn ministro Euangelico de esta lengua; El Vocabulario que comiença por Español sera muy copioso y se queda haziendo'.

³⁵ Febrès, *Arte de la lengua general del Reyno de Chile*, 421–2, 'en el Castellano he puesto la significacion con bastante explicación de lo que significa la palabra Indica, y muchas vezes he añadido otros verbos sinónimos, ò casi sinónimos en Español'.

³⁶ Sánchez Méndez, 'Arte y vocabulario', 133.

During the Enlightenment, some missionary authors mention specific sources by name. For instance, Pablo Restivo mentions a large number of sources in his new edition of Ruiz de Montoya's Guarani dictionary in 1724, and in the text abbreviations are used when examples are taken from these authors, such as 'Nic.' and 'Band.' (Nicolás Yapuguay and Simon Bandini, translators of spiritual works into Guarani), 'Mend.' (Mendoza, otherwise unidentified), and others. Febrès informs his readers that his dictionary is indebted to the work of Luis de Valdivia, and of Diego Amaya and Gaspar López, the compilers of unpublished wordlists. He also mentions grammars of other languages, such as Moxo (probably the anonymous *Arte y Vocabulario de la lengua morocosi* of 1699 or Pedro Marbán's of 1702) and the Hebrew grammar of Martín del Castillo.³⁷

Macrostructure

Dictionaries might be published with or without grammars. For instance, as we have seen, the *Arte y vocabulario* of Ruiz de Montoya includes both grammar and Spanish–Guarani vocabulary, whereas his Guarani–Spanish dictionary was published separately under the title *Tesoro*. The anonymous *Arte y vocabulario en la lengua general del Perú* of 1586 includes grammar and bidirectional wordlists.

Regarding directionality, it is remarkable that González Holguín's first section (and therefore that of the anonymous *Arte y vocabulario* of 1586) translates from Quechua to Spanish, unlike the majority of missionary dictionaries in Mesoamerica and the Philippines, which generally start with a section which translates from Spanish into the indigenous language, replacing the Latin equivalents of the Spanish–Latin dictionary of Nebrija. Without referring to González Holguín's work, Bertonio explains that he decided to place the Spanish–Aymara section of his work before the Aymara–Spanish section because it seems more natural for a novice to start with the language he already knows.³⁸

The two sections of a bidirectional dictionary might be reversed versions of each other, as is generally true, for instance, in the *Lexicon* of Santo Tomás. Julio Calvo Pérez labels such a type of bidirectional dictionaries 'bi-univocal', or 'reciprocal univocal' (or, since the reversal is not always perfect, 'quasi-bi-

³⁷ Febrès, *Arte de la lengua general del Reyno de Chile*, 416 (Valdivia, Castillo, and the Moxo grammar); 422–3 (Amaya, Valdivia, and López).

³⁸ Bertonio, *Vocabulario de la lengua aymara*, sig. B2r.

univocal').³⁹ For instance, in his Spanish–Quechua section, Santo Tomás translates the lemma ‘Aun no, *aduerbio*’ (‘not yet, adv.’) as *manarac*, and in his Quechua–Spanish section, we find a lemma *manarac*, translated as ‘aun no, *aduerbio*’. When two Quechua equivalents are given, as in ‘Ante, o delante’ (‘before’), translated as *ñauquimpi*, o *xutillapi*, we find both equivalents in the Quechua–Spanish section: *ñauquimpi*, o *xutillapi*, translated as ‘ante, o delante’ (such double equivalents in Quechua correspond to diatopic variation: in most cases the northern variety first, followed by the southern variety).⁴⁰ Likewise, Torres Rubio identifies his vocabularies as reciprocal: after his Spanish–Aymara vocabulary, its Aymara–Spanish counterpart is ‘al contrario del pasado’ (‘contrary to the previous’), and after his Spanish–Quechua vocabulary, its Quechua–Spanish counterpart is ‘al trocado del pasado’, with the same sense.⁴¹

The consequence of this macrostructure is that there are no major discrepancies between the paired vocabularies. By contrast, the *Tesoro* and *Vocabulario* of Ruiz de Montoya are not reversed versions of each other. Indeed, they are complementary, and the paratexts of the *Vocabulario* explain how to use them together: Spanish lemmata are given with Guaraní equivalents in the *Vocabulario*, but in order to learn how to use the Guaraní words in context, their entries in the *Tesoro* must be consulted. So, ‘if I search for *hombre* (“man”) in the *Vocabulario*, I find *abá*, and then I search for *abá* in the *Tesoro*, where I will find all the ways of saying things about men’.⁴² In fact, Ruiz de Montoya devotes three entire columns to the entry *abá* in the *Tesoro*, including many examples and expressions.

Likewise, Andrés de Febrès explains that, although his Spanish–Mapudungun ‘Vocabulario’ is modelled on his Mapudungun–Spanish ‘Calepino’, not all the entries of the ‘Calepino’ are included in the ‘Vocabulario’, and vice versa. Nor do they handle material in the same way. So, beginners can turn at first to the ‘Vocabulario’, where they will find the Mapudungun equivalent of a Spanish headword. If they want to get more information about the Mapudungun word, they can go to the ‘Calepino’, where they will be able to go further into its ‘proper

³⁹ Calvo Pérez, ‘El concepto de diccionario bilingüe’.

⁴⁰ Calvo Pérez, *El Arte y Vocabulario*, 24.

⁴¹ Torres Rubio, *Arte de la lengua aymara*, fo. 81r, and Torres Rubio, *Arte de la lengua quichua* (1619), sig. 81r.

⁴² Ruiz de Montoya, *Arte y vocabulario*, 101, ‘En este Vocabulario se ponen los vocablos simplemente. Para saber sus vsos, y modo de frasses, se ha de ocurrir a la segunda parte, v. g. busco aqui Hombre, hallo que es Abá, buscaré en la segunda parte, y alli hallaré lo que se dize del hombre.’

meaning'.⁴³ There was a reason for Febrès to call the more discursive of his wordlists a 'Calepino'. As Thomas Smith-Stark has pointed out in the context of lexicography in New Spain, Calepino's dictionary (especially in the eighteenth-century editions known to Febrès) is more anecdotal and words are explained and clarified, whereas Nebrija's entry style can be characterized as scientific and functional.⁴⁴ It is to the point that the final section of Febrès' *Arte* is a very brief treatment of phraseology, in which the author devotes only two pages to some expressions and examples, some in the form of questions and answers, before closing with a reference to the 'Calepino'.⁴⁵ In his mind, the 'Calepino' was in a way a continuation of the *Arte*, not a dictionary in the style of Nebrija. In fact it is not only phraseological but encyclopedic, with descriptions, definitions, and examples, some of them revealing the culture of the indigenous people, such as the entry *ñamculn*, 'To nail the head of a person on a pole, as they do when they chant victory over their enemies'.⁴⁶

In some Andean sources, the missionary linguists developed different tools for the teaching of the indigenous languages which fall outside the genres of *Arte* and *Vocabulario* in the strict sense. For instance, the anonymous *Doctrina Christiana y catecismo* of 1583–4 has a section of 'Annotations, or notes, on the translation of the Christian doctrine and catechism in the Quechua and Aymara languages', in which there are Quechua and Aymara wordlists.⁴⁷ Likewise, the anonymous *Tercero catecismo y exposicion de la Doctrina Christiana* of 1585 includes a verbal index.⁴⁸ Material like this can be called 'paralexical', supplementing dictionaries as more narrowly defined.

Bertonio also decided to break the traditional boundaries between the genres of grammar, dictionary, and texts. As early as 1596, in the prologue to his full-scale Aymara grammar, *Arte y grammatica muy copiosa de la lengua aymara*, he informed his readers that he planned to compose a more comprehensive dictionary, including all the particles which are essential for

⁴³ Febrès, *Arte de la lengua general del Reyno de Chile*, 295, 'los principiantes, despues de hallar una palabra Indica en este Vocabulario, la vean en el Calepino, para enterarse mas de su proprio significado'.

⁴⁴ Smith-Stark, 'Lexicography in New Spain', 13–15.

⁴⁵ Febrès, *Arte de la lengua general del Reyno de Chile*, 97–8.

⁴⁶ Febrès, *Arte de la lengua general del Reyno de Chile*, s.v. *ñamculn*, 'enclavar una Cabeza de gente en un palo, como hacen quando cantan victoria de sus enemigos'.

⁴⁷ *Doctrina Christiana y catecismo* (1584), fos. 74r [misnumbered 83]–84r ('Annotaciones, o scolios, sobre la traduccion de la Doctrina christiana, y Catecismo en las lenguas Quichua, y Aymara'), including Quechua wordlist at fos. 76v–77v and Aymara wordlist at fos. 79v–84r.

⁴⁸ *Tercero catecismo y exposicion de la Doctrina Christiana* (1585), sigs. ¶14r–¶17v.

Aymara, not only individually but also in combination. The large grammar and a smaller *Arte breve* were both printed in 1603; he observed in the *Arte breve* that it was not sufficient to learn the ‘rudiments’ of the Aymara language, namely the parts of speech with their ‘accidents’, based on Latin morphology, but that it was much more important to learn about the particles, which were mainly described in the large grammar. His monumental *Vocabulario de la lengua aymara* was published in 1612, as was an *Arte de la lengua aymara, con una silva de phrases*. In the epistle to the reader of the latter work, Bertonio explains that it contains two books. The first deals with the eight parts of speech and their syntax. The second book, the alphabetical ‘silva de phrases’, is written with the objective to learn Aymara words easily, together with their ‘particles’, and other matters which make the *Arte* more perfect. He adds that it is in fact a supplement (or complementary section) to his dictionary, which enables the learner to find examples where the words are used in context. This paralexicographical section is one of several pioneering features of this coherent and complete programme of language instruction.

Markedness and Glottonyms

In South America as elsewhere, missionary lexicographers did not have a unified approach to markedness, whether diatopical or diastratical.

As for the marking of diatopical variation, we see two tendencies. Many lexicographers ignored diatopical variety, or recognized the existence of regional differences, but decided not to mark them as such. An example is the first extant printed Quechua dictionary, that of Santo Tomás, which was collected from informants from both the Central and Peripheral branches of the Quechua language family.⁴⁹ According to the author, the data are the most commonly used forms throughout the whole region, a supra-dialectal variety or *koine* which is intelligible in most regions.⁵⁰ On the other hand, there are sources which include – systematically or not – information related to regional varieties. For instance, the anonymous *Arte y vocabulario* of 1586 occasionally marks a word with the abbreviation ‘(Chin.)’, which means that it is used particularly in the variety of Chinchaysuyo, the northern variety. Aymara sources include more details related to local varieties than most others. In the Aymara wordlist in the anonymous *Doctrina Christiana y catecismo para instruccion de los Indios* of 1583–4, we find an extensive list

⁴⁹ Mannheim, ‘Lexicography of Colonial Quechua’, 2681.

⁵⁰ Santo Tomás, *Lexicon*, sig. 6v, ‘la mas comun’.

with abbreviations referring to diatopical varieties: A. stands for 'Aymaraes del Cuzco', C. for '[C]ara[n]gas, y Charcas', L. for 'Lupacas', P. for 'Pacajes', Po. for 'Potosí', Q. for 'Quillacas', and 'alibi' indicates 'diuersas naciones'.⁵¹ Bertonio, by contrast, decided to describe only the variety spoken by Aymara people of the province of Lupaca.⁵²

The question of diatopical markedness is related to that of the naming of South American languages more generally. In Brazil, the 'general language' often does not have a specific glottonym, as the title of Anchieta's grammar indicates: it is simply 'a língua mais usada na costa do Brazil' ('the most commonly used language of the coast of Brazil'). Luis de Valdivia also decides also not to give a specific glottonym for the language he is describing, although he sums up several regional varieties in his grammar: it is simply 'lengua de Chile' (the terms *Mapuche* or *Mapudungun* are never used, although the term *Chillid[un]gu* is included in his vocabulary, translated as 'lengua de Chile'). In his grammar, Valdivia points at some differences between the varieties of the tribes from the south (Beliches), and those from the north, in particular the variety spoken in the diocese of Santiago, but these local differences are not marked in his dictionary. We also see the opposite, where rather than giving one glottonym or ethnonym in the title of a dictionary or just referring to the 'general' language, a lexicographer indicates a list of native peoples, as occurs in the title of Francisco de Tauste's dictionary of the language of 'the Chaymas, Cumanagotos, Cores, Parias, and others from the Province of Cumaná or New Andalusia'. In his grammar, no specific information is given for the linguistic varieties of each group, and in the dictionary we just find one variety. By contrast, Ruiz Blanco, who worked in the same region, marked certain forms as 'Pal[enque]' or 'Cariue' in the margins of his wordlist, though not in the entries themselves, these being two of the native peoples in the region of which he describes the language.⁵³ Antonio Machoni de Cerdeña's *Arte y vocabulario de la lengua lule, y tonocote* has two glottonyms in the title: as Adelaar observes, 'the question whether Lule and Tonocoté were indeed identical, and if not, whether Machoni's grammar deals with the former or the latter, may remain open', since these languages have long been extinct.⁵⁴ In other cases two different glottonyms used in different works refer to the

⁵¹ *Doctrina Christiana y catecismo para instruccion de los Indios* (1584), fo. 79v.

⁵² Bertonio, *Vocabulario de la lengua aymara*, sig. A2r.

⁵³ Ruiz Blanco, *Conversion de Piritu*, 27, includes 'Palenques' and 'Cariues' in a list of 'diuersas naciones de Indios'.

⁵⁴ Adelaar, *Languages of the Andes*, 385–6.

same language: indeed, the anonymous *Arte y vocabulario de la lengua morocosi* and Pedro Marbán's *Arte de la lengua moxa, con su vocabulario* are almost identical.

Diastratical markedness does not occur so frequently in the dictionaries discussed in this chapter, but in the Brazilian dictionaries we find the word *vulgo* (abbreviated to *vulg*, *vul*, or even *v.*) in at least three anonymous dictionaries of the 'Lingua geral amazônica', all written between 1744 and 1756 and derived from the anonymous *Vocabulário na língua brasilica* of 1621–2. Generally, the authors first mention the traditional word or expression as given in the source dictionary. Then, by marking a word or expression as *vulgo*, the authors refer to the usage of their time (130 years later) and region (we do not know if the 1621–2 dictionary was written in Amazonia or in the São Paulo region). So *vulgo* absolutely does not mean 'vulgar speech' and therefore is not pejorative.⁵⁵ It means 'Today and here people say ...', indicating diachronic rather than diaevaluative markedness. By contrast, when Pablo Restivo updated Antonio Ruiz de Montoya's dictionary of Guaraní of 1639–40 and adapted it to the needs of his time, the mid eighteenth century, many of the original lemmata were no longer in use, but Restivo did not decide to mark such archaic forms systematically.

Symbols are seldom used as markers. Bertonio's *Vocabulario de la lengua aymara* is an exception: he explains that the symbol † is used when there are 'other matters to tell related to the same word', the symbol ¶ is used 'when a grammatical rule is given, or when the general meaning of the verb or noun is given', and the symbol)(is used for antonyms.⁵⁶ Bertonio also gives the prepositions which have to be combined with the verbs, and notes that, when the term 'accusative' is added, this means that the verb does not govern any prepositional phrase, a modernity for the time under study.⁵⁷

⁵⁵ Wolf Dietrich, personal communication. According to Dietrich, the abbreviation does not stand for *vulgar* or *vulgarmente*, but for the Latin *vulgo*.

⁵⁶ Bertonio, *Vocabulario de la lengua aymara*, sig. Bir, 'La † se pone cuando se van diciendo otras cosas pertenecientes al mismo vocablo. La ¶ se pone donde se da alguna regla general perteneciente a la gramática, o a la significación general del no[m]bre, o verbo', repeated with some variation at 3 (second sequence of pagination), with the addition 'Esta señal)(... significa que el vocablo q[ue] se le sigue, es co[n]trario al q[ue] le precede.'

⁵⁷ Bertonio, *Vocabulario de la lengua aymara*, sig. Bir, 'Cua[n]do en el vocabulario se hallare, Ro, Na, Tha, Mpi, entender que el verbo pide su caso co[n] aquellas preposiciones y esto se hizo breuitatis causa. y quando dize Acusatiuo, entender que ha de ser sin preposicio[n].'

Citation Forms

Most lexicographers inform their users of which citation forms are used in the dictionaries as headwords. In his Quechua dictionary, Domingo de Santo Tomás takes the first-person and second-person singular of the verb as the citation form, even when he provides a Spanish infinitive as the gloss: the entry *canini*, *gui* is translated with the infinitive ‘morder’ (‘to bite’), although *canini* is the first-person singular, and *-gui* the ending for the second-person.⁵⁸ In the Aymara wordlist in the anonymous *Doctrina Christiana y catecismo* of 1583–4, a prefatory note states explicitly that the first-person form of the verb, ending in *-tha*, is used as the citation form.⁵⁹ Modern dictionaries, starting in the nineteenth century, have given the infinitive as the citation form.

When infinitives are non-existent, the first-person singular becomes a problematic citation form when it is formed with pronominal prefixes. When alphabetized by the first letter of citation forms in the indigenous language, translating to Spanish, all the verbs that shared a given pronominal prefix would be placed together, making individual lemmata hard to find, especially for the user who was not very familiar with the inflectional rules of the language. Several authors invented creative solutions for this problem, as Diego Basalenque did in his dictionary of Matlatzinca (see Chapter 27). In the anonymous Chiquitano–Spanish wordlist of the eighteenth century in the *Arte y vocabulario de la lengua chiquita*, the roots of the verbs are arranged alphabetically, disregarding the first-person singular prefixes *i-*, *ña-*, and *z-* which always accompany them, with a note explaining this arrangement, and a supplementary alphabetical table of roots following the main wordlist.⁶⁰ A similar lemmatization procedure occurs in Pedro Marbán’s dictionary of Moxo, a language in which ‘verbs begin with the possessive markers’, according to Marbán’s description. If the author had followed the lexicographical tradition of the Andes, where it is usual for the citation form of verbs to be the first-person singular, his dictionary would have consisted almost exclusively of lemmata beginning with the prefixes *ne-*, *ni-*, and *nu-*. For this reason, he decided to begin the ‘second syllable of the verb’ (i.e. the first of the root) with a capital letter, and to begin the ‘possessive marker’ with a lower-case letter.⁶¹ In the *Dictionnaire caraïbe-françois*, Raymond Breton uses

⁵⁸ Calvo Pérez, *El Arte y Vocabulario*, 23.

⁵⁹ *Doctrina Christiana y catecismo para instruccion de los Indios* (1584), fo. 79v.

⁶⁰ *Arte y vocabulario de la lengua chiquita* (1880), 71 (note on alphabetization), 129–36 (‘Tabla alfabética de las raíces [o temas primarios] de la lengua chiquita’).

⁶¹ Marbán, *Arte de la lengua moxa*, 362, ‘en esta Lengua todos los verbos empieçen por notas de posesion . . . se ha cogido por medio, que empiese el Vocabulario, por la segunda

the imperative as a citation form since it does not have a prefix, but then follows the entry for the imperative form of a verb with entries for inflected forms, which disturb the alphabetical sequence: for instance, *erébae*, ‘take it!’, is followed immediately by *neerériem*, ‘I take’.

Ruiz de Montoya’s dictionaries are constructed differently. He explains in the prologue of his Guarani–Spanish *Tesoro* that the user seeking a word has to find ‘lo fixo’, in other words the root, although he does not use this term here. For example, he explains, *ñémboé* is a compound form which can be segmented into three ‘particles’: *ñé* is ‘reciprocal’, *mô* is an ‘active particle’, and *e* means ‘destreza’ (‘skill’), so that all of them together mean ‘adestrarse’ or ‘aprender’ (‘to train oneself’ or ‘to learn’); adding the first-person singular pronominal prefix *A* gives *Añémboé*, which Ruiz de Montoya translates ‘Yo aprendo’ (‘I learn’). So, when the user wants to search for the meaning of *añémboé*, he has first to remove the ‘particles’ in order to find the root (as in the tradition of Hebrew-language instruction, where the ‘servile letters’ have to be removed, in order to determine the ‘radical letters’ containing the semanteme: for a similar example from Arabic, see Chapter 19). Conversely, the user of Ruiz de Montoya’s Spanish–Guarani *Vocabulario* who searches for the translation of the verb *enseñar* (‘to teach’) will find the equivalent ‘Amboé. E.11’. Turning to the *Tesoro* and searching under *E*, he will find twelve numbered lemmata *E*. The eleventh, for which the entry is two columns long, begins ‘Destreza, enseñanza’ (‘skill, teaching’), and among the forms which follow is the word *añémboé*, translated ‘aprendo’ (‘I learn’). In other entries in the *Vocabulario*, Ruiz de Montoya notes after each Guarani verb which of the eight forms of the supine it takes (he provides the rules for determining the form of the supine which is taken by any given verb in his *Arte*).⁶²

The Brazilian dictionaries describing Tupi usually did not separate the personal prefixes from the root of the verb and, since the infinitive is non-existent, they arranged these forms in alphabetical order, starting with the personal prefix of the first-person singular. No explanations are given about the difference between the two classes of prefixes *a-* (the actor has control over the action) and *xe-* (the actor does not have control).

Ruiz Blanco’s Cumanagoto vocabulary in his *Conversion de Piritu* of 1690 is not only, as we have seen, idiosyncratically laid out, but also has an unusual

sylaba del verbo, poniendo el posesivo con letra pequeña, y la segunda sylaba con letra majuscula’.

⁶² Ruiz de Montoya, *Arte y vocabulario*, 102 (prefatory note to *Vocabulario*), 25–8 (fuller treatment in *Arte*); see also Ruiz de Montoya, *Tesoro*, fo. 2r.

lemmatization. The Spanish column which gives the wordlist its alphabetical order starts with the preposition *a* 'to', but this is presented in prepositional phrases rather than by itself, so that the first two entries are '*Paragua pona* | A la mar' and '*Huerena pona* | A los llanos' ('to the sea' and 'to the plains' respectively). A series of such phrases is followed by a note on the Cumanagoto 'particles' which correspond to Spanish prepositions, despite being placed after the noun. A series of longer phrases then follows, starting with '*Paragua ponaitepai huaze* | Quiero ir à la mar' ('I want to go to the sea'). This is in keeping with the title of the vocabulary: 'Treasury of nouns and verbs of this language, with certain phrases and particular manners of speaking'.⁶³ The information given in the first ten pages of the vocabulary is quite unsystematic but detailed, but thereafter Ruiz Blanco decides to limit the lemmata to two citation forms, one for the present infinitive, from which the other 'tenses' and the negative are formed.⁶⁴

Selection of Headwords, Inflection, and Derivation

Lexicographers in all traditions have to make difficult decisions about which headwords to select as lemmata, and this is especially true in agglutinative and or polysynthetic languages, where the potential for suffixation seems to be endless. For a lexicographer, it would not make much sense to include all the derivational patterns of the Turkish verb *ölmek* 'to die', such as *öldürmek* 'to kill', *öldürtmek* 'have someone killed', and *öldürttürmek* 'have someone kill somebody', except when the meaning of a certain form is unpredictable, as in *öldürmek*, which also means 'tenderize' as a culinary term.⁶⁵ Including every derivative form for each headword or root of many New World languages would make the dictionary 'endless', as was observed by Juan de Córdova in his grammar of Zapotec (see Chapter 27).

So, in South America, missionary lexicographers had to make decisions concerning the selection of lemmata, and the degree to which it would be necessary or relevant to include derivative forms. In some cases we have to reconstruct implicit selection criteria, but most authors explicitly inform their users how they solved this problem. González Holguín explains to his readers

⁶³ Ruiz Blanco, *Conversion de Piritu*, 47 (second sequence of pagination), 'Tesoro de nombres, y verbos de esta lengua, con algunas frases, y modos de hablar particulares'.

⁶⁴ Ruiz Blanco, *Conversion de Piritu*, 56 (second sequence of pagination), 'Solamente los presentes infinitivos, y indicatiuos, de los quales se forman los màs tiempos, y modos negatiuos'.

⁶⁵ Examples from Zgusta, 'Influence of scripts and morphological language types', 300.

in the prologue of the Spanish–Quechua section of his dictionary that a dictionary has to be ‘complete’ and ‘genuine and natural’; the first condition means that for each word, all the ‘relatives’ are included, so that the user will be able to find it among all its offspring and parents which share the same basic meaning.⁶⁶ But he also refers to the *Arte*, his Quechua grammar of the previous year, in particular the third book where the particles are described and explained, together with other phrases and manners of speech, writing that his dictionary is larger than its predecessor because it draws new material from the *Arte*, which must be used with it, for a dictionary can be complete only if the user is able to use the particles in an appropriate way (explained in the *Arte*), combined with the corresponding lemmata of the dictionary.⁶⁷ In fact, González Holguín’s selection criteria are rather free, as has been observed by Mannheim: ‘The *Vocabulario* has entries for fully inflected and derived Quechua words and phrases along with the usual stem listings . . . The phrase entries are sometimes loan translations, in both directions, from Spanish to Quechua, and from Quechua to Spanish.’⁶⁸

Bertonio also wanted to limit the number of entries in his dictionary, particularly in the case of synonyms. He explained in the preliminaries of his dictionary that when one Spanish word has several equivalents, derivatives are given only for one of them. So, the Spanish word *acequia* ‘(irrigation) ditch’ is given three Aymara equivalents, *larca*, *pincha*, and *irpa*. The derivations of *larca* follow: ‘Hacer acequia, o haverla’ (‘make a ditch’) is *larcachatha*, and so on. Bertonio does not, however, give the corresponding derivatives of *pincha* and *irpa*, which would be *pinchachatha* and *irpachatha*.⁶⁹

Valdivia’s treatment of Mapudungun is similar: we cannot be sure whether this is a result of communication between lexicographers of different South American languages, as certainly happened in New Spain, or of the independent development of similar solutions to a shared challenge. He explains that his *Vocabulario* is relatively brief, and gives two reasons why: not all the

⁶⁶ González Holguín, *Vocabulario de la lengua general de todo el Peru*, 1 (second series of pagination), ‘Todos dessean Vocabulario con dos condiciones, vna es q[ue] sea cu[m]plido, y otra que sea verdadero y proprio. La primera condicio[n] que es dela copia se cumple con q[ue] en cada vocablo se le añadan todos sus agnatos, que son los descendie[n]tes de aquel vocablo, o los q[ue] nacen del, para que el que busca vn vocablo le halle entre todos sus hijos y parientes que participan de su misma significacion.’

⁶⁷ González Holguín, *Vocabulario de la lengua general de todo el Peru*, 2 (first series of pagination), ‘este aumento lo he sacado de nuestra Arte en especial del tercero libro, que es de la copia y otras phrasis que enseña toda la arte, y assi es necessario tener el arte para entender este aumento’.

⁶⁸ Mannheim, ‘Lexicography of Colonial Quechua’, 2682.

⁶⁹ Bertonio, *Vocabulario de la lengua aymara*, sig. A8v.

‘composed’ verbs (in other words compounded or affixed forms) are included, and not all the nouns which are derived from these words are included, since everyone who is familiar with the rules given in chapter 18 of the *Arte* can apply the appropriate particles to make the derivative forms.⁷⁰ As Emilio Ridruejo observes, most particles are included in the dictionary.⁷¹ In the next century, Andrés Febrès’ explanation of why his Spanish–Mapudungun wordlist is relatively short is comparable: he usually gives only a basic Spanish lemma from which many other forms – which are given in his Mapudungun–Spanish section – can be derived. So for example, although there is a Spanish word *peligrar* ‘to be in danger’, which is related formally and semantically to phrases such as *estar peligroso* and *ponerse en peligro*, these are not given head-word status. To find them and their Mapudungun equivalents, the user has to start searching under the noun *peligro* ‘danger’, from which the verb *peligrar* is derived. There he will find the Mapudungun equivalent, which is *cuñium* – and when he looks this form up in the Mapudungun–Spanish section, he will find all the associated Mapudungun forms with their Spanish equivalents. As for derived forms which are in neither of the wordlists, they can be constructed with the aid of the *Arte*.⁷²

Pedro Marbán seems to be acquainted with the approaches of his colleagues describing Andean languages. In the prologue to his *Arte de la lengua moxa*, he explains that in the Spanish–Moxo section the user will not find a great number of entries, but in the Moxo–Spanish section many more words will appear. Like other lexicographers, he decided not to include too many words constructed from ‘interposed particles’ in his dictionary, referring to his *Arte*, where the learner can find them. He includes a limited number of particles, only the most frequently used, since he did not want his work to grow too large.⁷³

Particles

This brings us back to the ways, already touched upon in the discussion of macrostructure above, in which the demands of South American missionary

⁷⁰ Valdivia, *Arte y gramática general*, sig. Gir (referring to chapter 19 of the *Arte*, but the relevant material is in chapter 18).

⁷¹ Ridruejo, in Valdivia, *Arte y gramática general* (2007), II.153.

⁷² Febrès, *Arte de la lengua general del Reyno de Chile*, 295–6, citing *ibid.*, chapter 6, i.e. pages 60–80 (he directs his readers in particular to paragraphs ‘161, y 162, y otros’).

⁷³ Marbán, *Arte de la lengua moxa*, 364, ‘No se ponen en este vocabulario los verbos, con todas aquellas particulas de interposicion . . . se ponen algunas, las mas frequentes, porque no crezca tanto la obra’; cf. *Arte y vocabulario de la lengua Morocosi*, sig. Azv.

lexicography put the concept of the dictionary itself under pressure. Was a dictionary a collection of *dicciones*, 'words'? If so, lexicographers had to think afresh about what constituted a 'word', and about how to handle bound morphemes or, as they called them, 'particles'. Usually, in the European tradition, grammar describes the eight parts of speech (sometimes expanded to ten), each of which was considered as a *dictio*. Any element which fell outside the scope of this classification was described at the end of the grammar, where we sometimes find particles (as the word was used in the description of European languages), manners of speaking, and the like.

When missionary lexicographers and grammarians started to describe non-Western languages, the treatment of bound morphemes was of great importance. They could not yet benefit from an elaborate and detailed theory regarding bound and free morphemes, and often they classified the use of particles as something extra, not essential (whether they were actually obligatory or not), since they did not represent any specific part of speech in the traditional model of eight parts. However, lexicographers often found creative solutions to the problem of how to deal with particles. They were often included in the dictionaries themselves, but some authors devote special sections or even separate paralexicographical works to them, notably Bertonio, describing Aymara, and Ruiz de Montoya (and his follower Pablo Restivo), describing Guarani.

The second of the two grammatical books of Bertonio's *Arte de la lengua aymara, con una silva de phrases* of 1612 contains a section devoted to the particles which are interposed in the verbs, including a list of seventy-two alphabetically arranged particles, intended for memorization: in effect, a brief dictionary of particles, with information about the contexts in which they occur.⁷⁴ The third, phraseological, book has an independent prologue, explaining that the groups of phrases it presents are each based on a certain keyword, arranged alphabetically. The main body of phraseological material is followed by alphabetical indexes of the Aymara and Spanish words which occur in the phrases, keyed to page numbers. This means that Bertonio's work has much to offer for the user: the treatment of particles in the grammar shows how to use them to make more complex forms, or how to decompose complex forms into smaller entities, and the indexes give access to a great many phrases which complement the basic treatment of words in the dictionary. So, as well as his separately published *Vocabulario*, we may say that Bertonio's *Arte* of 1612 offers a 'particulario', a dictionary of particles, and

⁷⁴ Bertonio, *Arte de la lengua aymara*, 91–2.

a 'phrasario', or a phrase book arranged alphabetically according to the keywords in these phrases. In an afterword to the second part of the *Vocabulario*, Bertonio writes that he first decided to publish the *Arte* of 1612 as the final section of his dictionary, making a third part after the Spanish–Aymara and Aymara–Spanish sections, probably in quarto or folio, but that he subsequently decided that it was more practical to use a smaller format for both the *Vocabulario* and the *Arte*.⁷⁵ In the end, the *Arte* was published in octavo (15 cm) and the 910-page dictionary in quarto (20 cm), dividing the lexicographical and paralexical material. Bertonio's oeuvre is probably one of the most complete programmes of language instruction of this period. Apart from these different sections and separate books, which each have their own prologue, he also includes a detailed theory about didactics and its methodology, resulting in his paratext titled 'Modo de estudiar esta lengua' ('The way to study this language').⁷⁶

In the tradition of linguistic studies on Guaraní, the earliest reference to a work exclusively devoted to the particles is the *De linguae guaranae particulis, quibus nimirum omnis eius ornatus definitur* ('On the particles of the Guaraní language, by which all its elegance is largely shaped'), written by Alonso de Aragona. He composed several other works, including a 'huge vocabulary [of Guaraní], in two parts', all considered to be lost, except his grammar titled 'Breve introducción para aprender la lengua guaraní', composed between 1620 and 1627 during his mission in what is today western Uruguay.⁷⁷ In his grammar, Aragona seems to make a distinction between obligatory particles and those which are used only as 'elegant speech', or 'embellishment'.⁷⁸

Ruiz de Montoya's Guaraní–Spanish *Tesoro* states right at the beginning of the text that the particles are the foundation of the Guaraní language. Many of them do not have a meaning when used separately, but in combination with others, 'intact or divided', since many of them are syncopated in composition, they make significative words.⁷⁹ The distinction underlying this statement, between *consignificativa*, those elements which co-signify, and *significativa*, nouns and verbs, which signify, is derived from the classical

⁷⁵ Bertonio, *Vocabulario de la lengua aymara*, 398.

⁷⁶ Bertonio, *Vocabulario de la lengua aymara*, sigs. B1r–B2r.

⁷⁷ B. Melià, in Aragona, 'Breve introducción para aprender la lengua guaraní', (1979), 24; Ribadeneira, *Bibliotheca scriptorum Societatis Jesu*, 16, 'Vocabularium ingens, duabus partibus comprehensum'.

⁷⁸ Aragona, 'Breve introducción para aprender la lengua guaraní' (1979), 27, 56.

⁷⁹ Ruíz de Montoya, *Tesoro*, sig. Arr, 'El Fundamento desta lengua son particulas, que muchas dellas por si no significa[n]: pero compuestas con otras, o enteras, o partidas (porque muchas las cortan en composicion) hazen vocas significatiuas.'

tradition. The third section of Ruiz de Montoya's *Arte de la lengua guarani* as reworked by Pablo Restivo ends with the usual *Laus Deo*. On the following page, a new section is appended, titled 'Particulas de la lengua guarani', and introduced by its own separate prologue.⁸⁰ Restivo explains that many particles are described in the *Arte*, but that, in order to find them more easily, they are arranged alphabetically in this section, which complements the grammar and the dictionary. It is probable that Restivo was the author of the anonymous 'Phrases selectas', a Spanish–Guarani work of 633 double-column pages, based on Ruiz de Montoya's *Tesoro*.⁸¹ The pseudonym 'Blas Pretovio', an anagram of Restivo's name, appears on a manuscript 'Arte de la lengua guarani' and a manuscript 'Particulas de la lengua Guarani'.⁸² It is significant that, in addition to the grammar and the dictionary, Restivo should have composed other independent works in order to have special tools for searching other entities than the word, one devoted to particles and another to phrases.⁸³

Content

The criteria for including or omitting headwords are not only morphological (as discussed above), but also semantic. Like the morphological criteria, the semantic ones are closely related to the users' needs. In the vocabularies which are modelled on Nebrija, the Spanish wordlist was prefabricated, and in some cases, New World lexicographers added items of their own, or omitted other lemmata, since the wordlist had to be adapted to the New World context. Since Nebrija did not compile his dictionary for religious or proselytizing purposes, missionary lexicographers tried to fill the gap of missing religious and ecclesiastical terminology. In some cases, they invented new forms in the indigenous languages, whereas in other cases they decided to use Spanish loans.

As Hamann points out, these missionary dictionaries 'provide fascinating glimpses of lost cultural worlds. In the indigenous New World, such dictionaries are often our earliest – and our richest – records of Native American languages', but on the other hand, 'the formal simplicity and linguistic transparency of translating dictionaries are deeply deceptive. In many cases, their entries have nothing to do with Native American cultures – they are

⁸⁰ Ruiz de Montoya, *Arte de la lengua guarani* (1724), 117–253 (second sequence of pagination).

⁸¹ Chamorro, 'Un diccionario manuscrito castellano–guaraní anónimo', 2–5.

⁸² Chamorro, 'Un diccionario manuscrito castellano–guaraní anónimo', 4.

⁸³ Chamorro, 'Un diccionario manuscrito castellano–guaraní anónimo', 12.

simply categories borrowed from a Castilian–Latin dictionary originally compiled to help early modern readers understand the pagan Mediterranean’.⁸⁴ Mannheim likewise observes that, ‘The first lexicographers of Quechua might be compared to modern-day ethnographers. They often provided important and subtle insights into the nature of Quechua language and culture . . . On the other hand, they also distorted Quechua terms and cultural patterns by forcing them into Western frameworks. After all, Westerners were the lexicographers’ audience.’⁸⁵ And as Torero demonstrates, the first, Spanish–Quechua part of the *Lexicon* of Domingo de Santo Tomás reflects the classical European worldview of Nebrija’s Spanish–Latin dictionary, whereas this is not the case in the second part, where the author attempts to give a more authentic image of the indigenous world.⁸⁶

In Breton’s monodirectional French–Arawakan *Dictionnaire caraïbe-françois*, the concepts in the indigenous language are frequently translated with encyclopedia-like explanatory material. So, for instance, detailed information is given about family relations: *ninántaganũ* is explained as ‘this is what uncles call nephews who are the sons of their sisters’, followed by much more on kinship terms of address.⁸⁷ Some entries contain entire anecdotes: the one for *cheboũitoũmain-áli* ‘he has run the gauntlet’ begins ‘In their festivals, I have seen a young man station himself in the middle of the communal house, his hands on his head . . .’ and goes on to explain Arawakan practices of endurance of pain.⁸⁸ Not all the lexicographers documented local cultures in a thorough way, but Ruiz de Montoya’s *Tesoro*, for instance, contains lemmata related to shamanism, dance, songs, the country, mountains, agriculture, hunting, fishing, indumentaria, proverbs and sayings, and many other topics.⁸⁹ Breton and Ruiz de Montoya were not exceptional in this respect.

Conclusion

González Holguín, Bertonio, Ruiz de Montoya, Febrès, and Restivo did groundbreaking work in lexicography and didactics. They combined

⁸⁴ Hamann, *Translations of Nebrija*, 85.

⁸⁵ Mannheim, ‘Lexicography of Colonial Quechua’, 2677; see also Moreno Fernández, ‘Antonio de Nebrija y la lexicografía americana’, 79.

⁸⁶ Torero, ‘Entre Roma y Lima’, 287–8.

⁸⁷ Breton, *Dictionnaire caraïbe-françois*, 11, ‘c’est ainsi que les Oncles appellent leurs nepueux fils de leurs sœurs seullement’.

⁸⁸ Breton, *Dictionnaire caraïbe-françois*, 132, ‘T’ay veu dans leur festins des ieunes hommes se planter au milieu du carbet, les mains sur la teste’.

⁸⁹ Melià, in Ruiz de Montoya, *Tesoro* (2011), xxvi.

grammatical rules and language instruction with the acquisition of basic vocabulary in their Spanish–indigenous language wordlists, and provided examples and words in context for the more advanced students in the wordlists which translated to Spanish. When they decided to give comprehensive analysis and overviews of particles they were working in a field which was unknown in any grammatical or lexicographical framework from Europe. Apart from these monumental works, a great number of vocabularies were produced describing an impressive number of languages, many of them now extinct or highly endangered, for instance Xebero, Chibcha, and Millcayac. These may have been less groundbreaking in the field of the history of lexicography, but they are without any doubt important sources of linguistic documentation.

Missionary Traditions in Mesoamerica

OTTO ZWARTJES

Although even today missionaries are active in linguistic fieldwork, the chronological focus of this chapter is on lexicographical production during the colonial period. Its geographical focus is the culturally defined area that is traditionally called Mesoamerica. To the east and south, Mesoamerica extends to the Caribbean Sea and the southern borders of the area where Mayan languages are spoken, excluding the Caribbean islands and the southern part of central America. The frontier to the north is more problematic. The current border between the United States and Mexico is not the appropriate frontier, because it does not correspond to earlier borders of the Spanish territories; moreover, the lexicographical traditions from what is sometimes called the Greater Southwest, in other words northern Mexico, are quite different from those of the central part of Mesoamerica. Those traditions will therefore be treated in Chapter 28.

In this chapter, the main focus will be on the most important lexicographical works of Tarascan; Nahuatl; the Oto-Manguean languages, such as Zapotec, Mixtec, Otomi, and Matlatzinca; and the Mayan languages. Apart from these greater works, some experiments in linguistic documentation and teaching will be discussed as well, such as an introduction to Mixe. An overview of the dictionaries to be discussed will be followed by accounts of their sources; their organizational principles; some innovative language-instruction tools and pedagogical approaches; and some topics related to the content of the dictionaries: did missionary lexicographers in Mesoamerica mainly try to find equivalents for lemmata from Western culture, or did they also attempt to give a record of the world of the other, or did they do both?

Historical Overview

The first pioneering dictionaries in New Spain appeared in the sixteenth century, most of them being the work of Franciscans in the Central Valley,

mainly studying Nahuatl, and Dominicans working in the region of Oaxaca, describing and documenting Oto-Manguean languages such as Mixtec and Zapotec. An important lexicographical tradition was developed on the peninsula of Yucatán and Guatemala, where Mayan languages were documented.

Around 1545 an anonymous author or authors compiled a trilingual Spanish–Latin–Nahuatl dictionary by copying the 15,260 entries of an early edition of Elio Antonio de Nebrija’s Spanish–Latin dictionary (for which see Chapter 14) and then adding about 11,000 Nahuatl glosses to them. This, the very first substantial dictionary of any language of the Americas, remained in manuscript throughout the period under discussion.¹ In 1547, the Franciscan Andrés de Olmos completed his Nahuatl grammar, which includes a vocabulary of 2,062 words, and also remained in manuscript during the colonial period.² The first printed dictionary of the New World – the second printed dictionary of any language other than those of Christian Europe, after Pedro de Alcalá’s dictionary of Arabic – was completed in 1555 by the Franciscan Alonso de Molina, containing 13,866 Spanish headwords (drawn from an edition of Nebrija published in 1545) with 29,742 Nahuatl equivalents.³ This was a monodirectional Spanish–Nahuatl dictionary. In 1571, Molina produced a bidirectional version, with 17,410 Spanish entries and 37,433 Nahuatl equivalents in its Spanish–Nahuatl section, and 23,625 Nahuatl entries in its Nahuatl–Spanish section.⁴ A new bidirectional dictionary of Spanish and Nahuatl was published by Pedro de Arenas in 1611; it was reprinted frequently in the seventeenth and eighteenth centuries. Lexicographers of Nahuatl in the seventeenth and eighteenth centuries did not always continue describing and teaching the prestigious variety of the Central Valley, changing their focus to more peripheral varieties, such as those spoken in Tierra Caliente or Jalisco.⁵

¹ Téllez Nieto, *Acercamiento filológico*, includes a partial edition; Hamann, *Translations of Nebrija*, 44–7, gives an account of the dictionary, identifying its source as the pirated edition of Nebrija, dated 1516 and probably printed around 1520, and illustrates a page of the manuscript.

² Olmos, *Arte de la lengua mexicana* (1993), is an edition; entry counts for this wordlist and those of Molina are from Clayton and Campbell, ‘Alonso de Molina as lexicographer’, 336–8; see also Hamann, *Translations of Nebrija*, 47–8.

³ Molina, *Aquí comienza un vocabulario* (2001), includes a facsimile of the 1555 edition; for the source, see Hamann, *Translations of Nebrija*, 48–50.

⁴ Hamann, *Translations of Nebrija*, 59.

⁵ See Zwartjes’ entries on Juan Guerra, *Arte de la lengua mexicana* (1692); Manuel Pérez, *Arte de el idioma mexicano* (1713), and Jerónimo Tomás de Aquino Cortés y Zedeño, *Arte, vocabulario, y confessorio en el idioma mexicano* in *Corpus de textes linguistiques fondamentaux*.

Printed dictionaries of other Mesoamerican languages followed very soon. The first was Maturino Gilberti's work on Tarascan, published in 1559, and offering 6,254 Tarascan–Spanish and 13,668 Spanish–Tarascan entries, making it the first bidirectional dictionary printed in the New World.⁶ Its Spanish wordlist is a close copy of that of the 1555 edition of Molina's Nahuatl dictionary.⁷ It was followed by Juan Baptista de Lagunas' Tarascan–Spanish dictionary of 1574. A Spanish–Zapotec dictionary by Juan de Córdova was published in 1578, drawing on the 1571 edition of Molina: its 28,352 entries made it the most extensive of the Mesoamerican dictionaries of its day that had Spanish as a source language.⁸ A Spanish–Mixtec dictionary published by Francisco de Alvarado in 1593 drew both on Córdova's work and also directly on the 1571 Molina.⁹

These printed bilingual dictionaries were sometimes elaborately marked up with equivalents in a third language. As early as 1557, Andrés de Castro added 25,000 or more words of Matlatzinca to the margins of a copy of the 1555 edition of Molina's Nahuatl dictionary.¹⁰ Glosses in Otomi were added to copies of the 1555 and 1571 editions of Molina's dictionary and to a copy of Gilberti's Tarascan dictionary.¹¹ A similar annotated printed book must have been the parent of the extant manuscript of Alonso Urbano's large Spanish–Nahuatl–Otomi dictionary, dated 1605, for the Spanish–Nahuatl part of each of Urbano's entries is copied from the 1555 edition of Molina's dictionary.¹²

Other wordlists were never printed during the colonial period and are conserved in manuscript form. Thomas Smith-Stark lists nine or ten from the period before 1611.¹³ The earliest is a Kaqchikel–Spanish wordlist, giving forms in K'iche' and Tz'utujil when they differed from the Kaqchikel form, which was compiled by Domingo de Vico, who died in 1555; this is the oldest

⁶ Gilberti, *Vocabulario en lengua de Mechuacán* (1997), is an edition; entry counts are from Monzón, 'Tarascan lexicographic tradition', 173.

⁷ Hamann, *Translations of Nebrija*, 51–5.

⁸ Córdova, *Vocabulario Castellano–Zapoteco* (1942), is a facsimile; for the source, see Hamann, *Translations of Nebrija*, 62; for the entry count, see Smith-Stark, 'Lexicography in New Spain', 26.

⁹ Alvarado, *Vocabulario en lengua mixteca* (1962), is a facsimile; for the sources, see Hamann, *Translations of Nebrija*, 73–4.

¹⁰ Schuller, 'Unknown Matlatzinka manuscript vocabulary'; further information in Hamann, *Translations of Nebrija*, 51, 109, and 172 n. 25.

¹¹ Hamann, *Translations of Nebrija*, 50–1, 55, 59, 108–11.

¹² Hamann, *Translations of Nebrija*, 77–8; Urbano, *Arte breve dela lengua otomí y vocabulario trilingüe* (1990), is a facsimile edition.

¹³ Smith-Stark, 'Lexicography in New Spain', 19–21; see also Hernández, *Lexicografía hispano-amerindia*, 29–35 and *passim*.

extant wordlist of a Mayan language.¹⁴ It was followed by a Spanish–Kaqchikel wordlist of the 1570s (it is based on the 1571 edition of Molina’s Nahuatl dictionary, and a manuscript of c. 1578 is extant, so it must have been composed between those dates) by Juan Alonso.¹⁵ A bidirectional dictionary of Tzeltal and Spanish, with 7,769 Spanish entries, was compiled by Domingo de Ara around 1560.¹⁶ A Spanish–Maya dictionary of about 9,000 or 9,500 entries and a bidirectional dictionary of Maya and Spanish are both attributed to Alonso de Solana, and belong to the period around 1580.¹⁷ The manuscripts of these dictionaries are known respectively as ‘Diccionario de Motul II’ and ‘Diccionario de San Francisco’, the latter taking its name from the Franciscan convent of that name in Mérida. Another Spanish–Maya dictionary was compiled by Gaspar González de Nájera by 1582.¹⁸ The Maya–Spanish dictionary of 15,975 entries compiled by Antonio de Ciudad Real at the end of the century was the most extensive of the dictionaries with a Mesoamerican source language; a seventeenth-century source reports that its compilation took forty years.¹⁹ An anonymous compiler produced a Spanish–Tzotzil dictionary of 8,077 entries at the end of the sixteenth century or the beginning of the seventeenth, now extant only in a twentieth-century transcript.²⁰ A bidirectional dictionary of Spanish and Tarascan was compiled by an anonymous Augustinian friar at the end of the sixteenth century, offering about 22,000 Spanish–Tarascan and 27,000 Tarascan–Spanish entries.²¹ The otherwise obscure Dionisio de Zúñiga Marroquín compiled a bidirectional dictionary of Spanish and Poqomchi’ with a Latin–Poqomchi’ section around 1608.²² To these may be added the Matlatzinca–Spanish and Spanish–

¹⁴ Smith-Stark, ‘Lexicography in New Spain’, 24; Hernández, ‘Indigenismos’, 69–71; Hernández, ‘Vocabularios Hispano–Mayas’, 142–3.

¹⁵ Hernández, ‘Vocabularios Hispano–Mayas’, 141–2; Hamann, *Translations of Nebrija*, 60.

¹⁶ Ara, *Vocabulario de lengua tzeldal* (1986), is an edition (in the introduction of which a date c. 1560 is assigned); see also Hernández, ‘Vocabularios Hispano–Mayas’, 140–1, and, for the entry count, Laughlin in *Great Tzotzil Dictionary of Santo Domingo Zinacantan* (1988), 1.28.

¹⁷ Hernández, ‘En torno al vocabulario hispano–maya’ (entry count of 9,500 at 115; cf. the lower counts reported by Smith-Stark, ‘Lexicography in New Spain’, 26); see also Hamann, *Translations of Nebrija*, 60–1.

¹⁸ Hernández, ‘Vocabularios Hispano–Mayas’, 134.

¹⁹ [Ciudad Real], *Calepino maya de Motul* (1995), is an edition; for the entry count and the forty years, see Smith-Stark, ‘Lexicography in New Spain’, 26, 28.

²⁰ *Great Tzotzil Dictionary of Santo Domingo Zinacantan* (1988) presents a facsimile of the extant transcript, with a freely modernized edition; Tzotzil–English and English–Tzotzil dictionaries based on the edition; a grammar; and an introduction: entry count 1.28.

²¹ *Diccionario grande de la lengua de Michuacán* (1991) is an edition; entry counts from Monzón, ‘Tarascan lexicographic tradition’, 181.

²² Smith-Stark, ‘Lexicography in New Spain’, 55 (and see 76 for information about the manuscripts).

Matlatzinca dictionaries of Diego Basalenque, of which the manuscripts are dated 1640 and 1642 respectively, and anonymous eighteenth-century dictionaries of Kaqchikel and K'iche' with Spanish, one of which has been edited as *Vocabulario en lengua 4iche otlatecas* (the character 4 represents the same ejective consonant as the combination *k'* which is now used: see the section on organizational principles below).²³

It is remarkable that the most comprehensive dictionaries – and also the best, from our modern perspective – were composed in the period 1555–1650 (the same applies to South America, but not to the Philippines: see Chapters 26 and 29 respectively). It is surprising that some of the minor grammars were often reprinted (such as the five editions of Vázquez Gastelú's grammar of Nahuatl), whereas the major lexicographical works were generally not reprinted during the colonial period: there were, for instance, no reprints of Molina's dictionary after the 1571 edition, and Olmos' work was never printed at all.²⁴ This is probably a consequence of the changing nature of the pedagogical tools in language instruction. By the 1630s, Miguel de Guevara's 'Arte doctrinal y modo general para aprender la lengua Matlaltzinga' and Diego de Nágera Yanguas' *Doctrina y enseñanza en la lengua maçahua* were combining information about vocabulary and about grammar, and Agustín de Quintana's *Confessionario en lengua mixe, con . . . un compendio de voces mixes* would likewise mix genres in the following century. When there was a need for an update to the dictionary record – for instance, when words of the dictionaries of centuries before were no longer in use, when archaic usage had to be avoided, or when specific dialectal varieties were missing in the pioneering works – authors tended not to publish new dictionaries as independent works, but included them as sections of other publications, such as Cortés y Zedeño's *Arte, vocabulario, y confessionario* (1765).

The variety of the lexicographical works which I have surveyed above is great, and in particular the works that translate from the indigenous language to Spanish are far from uniform. Almost every lexicographer followed his own methodology which was developed for the specific features of the language under study.

When dictionaries are bilingual and bidirectional, the two sections are seldom equal in length. So, for instance, the Spanish–Nahuatl first part of

²³ Basalenque, *Arte y vocabulario de la lengua matlaltzinga vuelto en la castellana* (1975), and Basalenque, *Vocabulario de la lengua castellana vuelto en la matlaltzinga* (1975), are editions of the first two; for the latter, see the table in *Diccionario k'iche' de Berlin* (2017), 12.

²⁴ For the several reprints of Vázquez Gaselú's grammar, see O. Zwartjes' entry for the author in *Corpus de textes linguistiques fondamentaux*.

Molina's dictionary of 1571 contains 121 folios, whereas the Nahuatl–Spanish second part contains 162 folios. By contrast, the Matlatzinca–Spanish part of Diego Basalenque's work is less comprehensive than the Spanish–Matlatzinca part. Maturino Gilberti's Spanish–Tarascan section is twice as big as his Tarascan–Spanish section, and, unlike Molina, Gilberti puts the section of his dictionary which starts with Spanish in second place.²⁵

Some lexicographers composed works other than dictionaries. Alonso de Molina, for instance, also published a Nahuatl grammar and a *Doctrina christiana breve traduzida en lengua mexicana*. Juan de Córdova wrote a grammar as well as a dictionary of Zapotec. Grammar and dictionary were not always by the same person: for Mixtec, for instance, we have an author who composed the first printed grammar, namely Alonso de los Reyes, whereas another missionary, Francisco de Alvarado, composed the dictionary. But it was not unknown for a single author to compose a whole suite of grammatical and lexicographical texts, as in the case of Basalenque, whose dictionaries of Matlatzinca are accompanied by a note on the use of the Roman alphabet to write Matlatzinca, 'Cartilla Matlaltzinga'; a grammar, 'Arte de le lengua Matlaltzinga'; and a treatise on particles, to which we shall return, 'Tratado de las particulas de la lengua Matlaltzinga'.²⁶ This means that sometimes there is a tight relation between a dictionary and a grammar or other texts, which results in a frequent use of cross-references, whereas in other cases the works can be quite different from each other.

Sources

Although other sources were used, most Mesoamerican dictionaries were, as we have seen, inspired by the Spanish–Latin *Vocabulario* of Antonio de Nebrija, first published in 1495. Some authors followed Nebrija's entries strictly, while others used the *Vocabulario* more freely, adding new entries, suppressing others, adapting Nebrija's work to the languages they studied, or just following their own creativity. As was the case with the use of Nebrija's Latin grammar, there was not just one version which served as a model for missionary lexicographers. As Byron Ellsworth Hamann observes in *The Translations of Nebrija*, after the *editio princeps* thirty-four further editions of the *Vocabulario* were published in nine European cities by the early 1600s; their 'lists of entries were in constant flux', and 'the constantly changing

²⁵ Monzón, 'Tarascan lexicographic tradition', 173.

²⁶ Basalenque, *Arte y vocabulario de la lengua matlaltzinga* (1975), 7–12 ('Cartilla'), 13–117 ('Arte'), 119–42 ('Tratado').

nature of Nebrija's "Castilian–Latin dictionary" has been underappreciated, but it is of fundamental importance'. Nebrija's dictionary, Hamann continues, 'gave birth to multilingual offspring', which had offspring of their own: 'The later we move through the sixteenth century, the more complicated these genealogies become . . . in some cases, dictionaries are descended from two or even three progenitors.'²⁷ As we saw above, Alvarado's Mixtec dictionary of 1593 draws on Córdova's Zapotec dictionary of 1578, and also directly on a source of Córdova's, namely Molina's Nahuatl dictionary in its edition of 1571; Hamann points out that it also draws directly on an edition of the ultimate source of Córdova's and Molina's work, the *Vocabulario* of Nebrija.

References in dictionary prefaces to 'Antonio' as a forerunner often refer generically to any later dictionary in the Nebrija tradition; likewise, 'Calepino' is often used as a generic term for dictionaries in the long and complex Calepino tradition, and these too were sources for missionary lexicographers in Mesoamerica (for Calepino, see Chapter 14, and for the contrast between the Nebrija and Calepino traditions, see Chapter 26). Lagunas explicitly mentions the name of Calepino in his dictionary of Tarascan. As he informs his reader in the prologue, the text 'could be called a small dictionary [*diccionario*], because the author (for the benefit of the students) imitates the order and manner of Ambrogio Calepino. Though in the order of letters it is impossible to do this in this language, given the different meanings.'²⁸ As Cristina Monzón notes, it is remarkable that Lagunas does not mention Gilberti, who was apparently one of his sources, but Calepino, although the difference between Lagunas' work and that of Calepino is enormous. Monzón suggests that Lagunas decided to mention Calepino as an authority, and concludes that 'it seems plausible to postulate that Lagunas was also acquainted with the *Regio 1502* [the first edition of Calepino] that was on the shelves of the library of the Colegio de Santa Cruz de Tlatelolco'.²⁹

Spanish-speaking lexicographers working in the New World (and in the Philippines) did not always take one of Nebrija's dictionaries as their model: as we have seen, Gilberti's Spanish–Tarascan dictionary of 1559 is a translation of Molina which replaces the Nahuatl equivalences with Tarascan: 'Any debt which Gilberti owes to Nebrija was acquired via Molina.'³⁰ The shorter

²⁷ Hamann, *Translations of Nebrija*, 3.

²⁸ Lagunas, *Arte y diccionario*, 1 (second sequence of pagination), 'podra se llamar Dictionarito, porq[ue] el auctor (p[ar]a mas p[ro]uecho d[e] los estudia[n]tes) sigue e[n e]l orde[n] y modo del Ambrosio Calepino. Au[n]q[ue] e[n e]l orde[n] d[e] las letras es imposible e[n e]sta le[n]gua, por los distintos significados'; translated in Monzón, 'Tarascan lexicographic tradition', 189.

²⁹ Monzón, 'Tarascan lexicographic tradition', 191.

³⁰ Smith-Stark, 'Lexicography in New Spain', 8.

wordlists characteristic of the eighteenth century are generally not derived from Nebrija or from Molina's dictionary.³¹

The section of a bidirectional dictionary that starts with the indigenous language was not necessarily the inverted version of the section which starts with Spanish, so even if Nebrija was the source for one section of a bidirectional dictionary, the other might be an original compilation. So, for example, in Molina's bidirectional dictionary of 1571, 'the Nahuatl/Spanish side is a new compilation rather than a reversal of the Spanish/Nahuatl section'.³² The same is true of Gilberti's *Vocabulario* of 1559: the Spanish–Tarascan section follows the first edition of Molina's dictionary closely, whereas the Tarascan–Spanish section is 'a creation by Gilberti and his team since he did not [invert] the Spanish–Tarascan entries'.³³ In the lexicographical tradition of Mayan languages, Ciudad Real's dictionary of Maya and de Ara's dictionary of Tzeltal are monodirectional, starting with the indigenous language, and as far as we know, they never had counterparts starting with Spanish. In these dictionaries, no trace of Nebrija can be found.³⁴ In the historical introduction of Michael Dürr and Frauke Sachse to their edition of the *Vocabulario en lengua 4iche otlatecas*, Nebrija is not identified as a source.

In Hamann's *Translations of Nebrija*, only brief references are included to bilingual dictionaries that translate to Castilian from another language. Without any doubt, the study of Hamann is important, but seen from the perspective of a historian of lexicography, the sections which translate from the indigenous languages to Spanish are often much more interesting, since the author had to make important decisions about how to organize the lemmata, in particular for polysynthetic languages with a great number of affixes, such as Nahuatl, or languages which have a great number of prefixes, such as Matlatzinca. How to separate the headwords or roots from other elements, such as bound morphemes and pronominal prefixes, was a great challenge for missionary lexicographers; we now turn to their solutions to these problems.

Organizational Principles

The alphabetical sequences by which headwords were ordered in the Mesoamerican missionary dictionaries cannot all be treated in detail

³¹ Yáñez Rosales, 'Presencia y ausencia de Antonio de Nebrija'.

³² Karttunen, 'Nahuatl lexicography', 2658.

³³ Monzón, 'Tarascan lexicographic tradition', 173.

³⁴ Hernández, *Lexicografía hispano-amerindia*, 139–55.

here.³⁵ Missionaries followed Nebrija's model, but also developed a special alphabet for the languages under study. They gave new values to existing letters; created new combinations of letters, resulting in digraphs and trigraphs; and added diacritics whenever necessary and whenever the printing facilities were able to reproduce them. In some cases they developed new letters, such as the frequently quoted alphabet devised in the mid sixteenth century by Francisco de la Parra to represent uvular and glottalized phonemes in languages such as K'iche' and Kaqchikel: A B C CH E H Y K L M N O P Q R T V X E 4 H TZ.³⁶ Sometimes the digraphs are ordered alphabetically, which means for instance that <th> is found under the letter <t> or immediately after it, and sometimes they appear at the end, after the last letter of the Spanish alphabet, as in Basalenque's dictionary of Matlatzinca.

In pre-modern times, suprasegmentals were not always recognized or described. The tones in Oto-Manguean languages were described relatively late, and before the grammars of the Jesuits Antonio del Rincón (1595) and Horacio Carochi (1645) vowel length and the presence of the glottal stop as a phoneme were not described systematically.³⁷ This underdifferentiation often led to ambiguity, or gave a false image of polysemy when words which were not in fact homophones were written as homographs. One example is Molina's entry *auatl*, which is translated as 'enzina, roble, gusano lanudo. o espina' ('oak, woolly caterpillar, or thorn'). This entry represents three different Nahuatl nouns: *āhuatl* 'oak', *āhuātl* 'woolly caterpillar', and *ahhuatl* 'thorn' (where the first of the *h*'s represents a glottal stop).³⁸

Apart from these questions of how to reproduce the phonology of Mesoamerican languages with the Roman alphabet, the organizational problems confronted by the missionary lexicographers were morphological: the concept of the root; polysyntheticity and affixation; the particle; the treatment of derivational processes. After these have been discussed, this section will conclude with some remarks on citation forms, markedness, and the structure of the dictionary entry.

³⁵ Fuller accounts include Smith-Stark, 'Phonological description', 12–21, and Monzón, 'Tarascan lexicographic tradition', 183–5.

³⁶ Smith-Stark, 'Phonological description', 17; M. Dürr and F. Sachse in *Diccionario k'iche' de Berlin* (2017), 21.

³⁷ Karttunen, 'Nahuatl lexicography', 2659; see also Smith-Stark, 'Phonological description', 15 (Rincón and the glottal stop), 23 (Rincón, Carochi, and vowel length), 24–6 (tone).

³⁸ Molina, *Vocabulario en lengua castellana y mexicana* (1571), fo. 9r, cited in Karttunen, 'Nahuatl lexicography', 2658.

As for the concept of the root, the Tarascan lexicographical tradition can be distinguished from the others, since missionary lexicographers working in this tradition started from the beginning to distinguish roots from words. Although Molina describes the distinction between the roots and the ‘servile letters’ in his Nahuatl grammar, there is no systematic treatment of roots in his dictionaries. Gilberti’s *Vocabulario*, by contrast, has a macrostructure designed to engage with the presentation of roots. Its first section contains the Tarascan words with Spanish translations, followed by an intermediate section which is a list of ninety-two roots, followed by the Spanish–Tarascan section. An addendum contains 123 additional Spanish–Tarascan items.³⁹ The intermediate section containing roots laid the foundations for later lexicographers. The section is preceded by a short introduction, in which Gilberti informs his readers that ‘the following contains certain verbs in alphabetical order, which some people wish to call roots, because it seems that, separated from the limbs or, better, from those that do service, the root will be left without meaning, like a trunk that has no branches, but is ready to produce them’.⁴⁰ The second dictionary of Tarascan, by Lagunas, also describes roots: ‘they are like a base or a foundation to build upon, or like roots ready to produce, or first and true etymological positions, on which are constructed and built or produced the true edifice or productive branches for the composition of verbs, verbal nouns, and adverbs with the elements of the interpositions’.⁴¹

Although the concept of the root had already been mentioned by Molina and had been explored further by Gilberti, Lagunas has a very creative and original lexicographical approach. He explains in his prologue that he marks the root with the symbol of a cross, after which the lemma starts with the root in Tarascan, followed by its translation. Next, more complicated words are given, and Lagunas gives in the margin in italics the several affixes in alphabetical order, which are explained separately (he often called them

³⁹ Monzón, ‘Tarascan lexicographic tradition’, 173.

⁴⁰ Gilberti, *Vocabulario en lengua de Mechuacan* (1559), fo. 80r, ‘SIGVENSE CIERTOS VERBOS POR EL AL- | phabeto, a los quales algunos quieren llamar rayzes: porque parece que apartados los miembros, o para mejor dezir las seruiles quedara la rayz sin significar nada, como el tronco sin ramos: solamente dispuesto a producirlos’, translated in Monzón, ‘Tarascan lexicographic tradition’, 174.

⁴¹ Lagunas, *Arte y dictionario*, 38 (second sequence of pagination), ‘son como fundamento, o vasas para edificar, o como rayzes aptas a produzir, o primeras posiciones ethymologicas. i. verdaderas, sobre quien se arman y edifican, o producen el verdadero edificio, o ramos productiuos de la composicion en los verbos y nombres verbales y aduerbios, mediante los materiales de las Interposiciones’, translated in Monzón, ‘Tarascan lexicographic tradition’, 176.

‘interposiciones’, but we also find terms such as ‘particulas’ and ‘preposiciones’).⁴² So, for instance, the root *andà* appears in its place in the alphabetical sequence of roots, marked by a cross, and translated into Spanish.⁴³ The sublemma which follows (not marked by a cross) is the word *andambezcanitominis* which is analysed and translated, with the ‘interposicion’ *bez* beside it in the margin. The next sublemmata, *andacazcani* and *andahcazingani*, are similarly treated, with the ‘interposiciones’ *cazca* and *cazin* beside them in the margin, and so on, so that the user can find a given root in one alphabetical sequence, and can then see the ‘interposiciones’ with which it can be combined in a secondary alphabetical sequence. The forms given in the margin are also listed in a section of the *Arte* which is in effect a separate glossary of ‘interposiciones’.⁴⁴ Lagunas was aware that he was innovating: as he observes in his prologue, his dictionary is ‘briefer and more profitable, and [written] in a curious manner and in alphabetical order, and not so much according to the letters, but rather in the profitable phrases, following the *Arte* and “interposiciones”’.⁴⁵

The problems presented by polysyntheticity and affixation are summed up in an example given in Richard Andrews’ *Introduction to Classical Nahuatl*: Molina’s translation of Spanish *monacordio* (‘clavichord’) into Nahuatl is *petlacalmecaueutl*.⁴⁶ This ‘word’ is in fact a complete phrase, meaning literally ‘it is an upright drum with strings that has the form of a wickerwork coffer’. In Nahuatl, the absolutive ending *-tl* always indicates nominal predication, whose translation is a phrase in English. For a lexicographer, it will not make any sense to include separate lemmata for the forms in polysynthetic languages which correspond to English phrases like ‘I see the house’, ‘I see the mountain’, ‘I see the forest’, and so on. The word ‘dictionary’ is in fact a misnomer when polysynthetic languages are concerned – or at least, a dictionary of a polysynthetic language must be, as has been said in the context of Inuktitut, ‘a dictionary without words’.⁴⁷ It has to be observed that the concept of *dictionary* is in fact a collection of *dictiones*, and if we follow

⁴² Lagunas, *Arte y diccionario*, 3 (second sequence of pagination).

⁴³ Lagunas, *Arte y diccionario*, 10 (second sequence of pagination).

⁴⁴ Lagunas, *Arte y diccionario*, 144 (first sequence of pagination), ‘De las interposiciones que ya comienzan por su orden Alfabético’.

⁴⁵ Lagunas, *Arte y diccionario*, 2 (second sequence of pagination), ‘Mas breue y prouechoso que ningun otro diccionario, y en modo curioso y Alfabético, y no tanto en las letras como e[n] las prouechosas sentencias, siguiendo al Arte y interposiciones.’

⁴⁶ Molina, *Vocabulario en lengua castellana y mexicana* (1571), s.v. *monacordio* (*petlacalmecaueutl* is, conversely, a headword in the Nahuatl–Spanish section); Andrews, *Introduction to Classical Nahuatl*, 20.

⁴⁷ Cornillac, ‘De la nécessité de concevoir’, 206, ‘un dictionnaire sans mots’.

Priscian's definition of the term *dictio – pars minima orationis constructae* ('the minimal part of a constructed sentence') – there would be no place in a dictionary for smaller entities, such as 'clitics', 'particles', or 'affixes' (see below), nor for larger entities, such as the constituent or the phrase.

Agglutinative languages with (mainly) suffixation did not cause major problems for missionary lexicographers, since suffixation did not affect the alphabetical sequence. By contrast, prefixation does affect the alphabetical sequence, and this problem had to be solved, in particular in the dictionaries which translated from the indigenous language to Spanish. So, for instance, compiling the Spanish–Matlatzinca section of his dictionary was an easy task, according to Basalenque. With 'our Spanish dictionary' he went to 'the bilingual natives', and started asking them to translate.⁴⁸ It was much more complicated to make the inverted Matlatzinca–Spanish section. Matlatzinca, according to Basalenque, is different from other languages of the region (he probably compared Matlatzinca with Tarascan, which has mainly suffixes) since it uses *iniciales*, or prefixes. There is no infinitive form of the verb which could be used as the citation form of the lemma. Basalenque found a solution, and the structure of the entries of this section is unique: the entry for each root is preceded by three columns, in which he fills in the prefixes which can be combined with the root. Nouns often, according to Basalenque, take single prefixes, such as *ca*, *hue*, *huebe*, *huebu*, and *huebete*, explained in the prologue of the Matlatzinca–Spanish dictionary.⁴⁹ Verbs can take one, two, or three prefixes. So, an entry for a root which is used in the composition of verbal and nominal forms might look like this:

<i>qui</i>	<i>tu</i>	<i>tzitzí</i>	<i>yo como cosa de fruta</i> ('I eat something, as a fruit')
<i>qui</i>	<i>tu</i>	<i>tu tzitzí</i>	<i>yo doy de comer a otro</i> ('I give someone else to eat')
	<i>in</i>	<i>tzitzí</i>	<i>la comida</i> ('the food'). ⁵⁰

We now turn to the question of the particle, an element which, as Gerda Haßler observes, falls between grammar and dictionary.⁵¹ In most Mesoamerican missionary grammars of this period, we find a main section devoted to the eight parts of speech. The final section of the grammar is often devoted to 'particles' or *addiciones*, an *addictio* in this sense meaning an element which is used in combination with a *dictio* (compare *adverb* and

⁴⁸ Basalenque, *Arte y vocabulario de la lengua matlaltzinga* (1975), 145, 'nuestro vocabulario castellano fui preguntando a los naturales ladinos'.

⁴⁹ Basalenque, *Arte y vocabulario de la lengua matlaltzinga* (1975), 147.

⁵⁰ Basalenque, *Arte y vocabulario de la lengua matlaltzinga* (1975), 318.

⁵¹ Haßler, 'Las partículas entre la gramática y la lexicografía', 87.

adnoun). Juan de Córdova's *Arte* closes with a list of seventy of these *additions*, arranged alphabetically, some of them labelled as particles and others as adverbs or interjections (there is no systematic classification here, but most do not occur in the preceding sections devoted to these indeclinable parts of speech). Basalenque's treatment of particles, in a 'Tratado de las partículas', with its own half-title page and prologue, between the 'Arte' and 'Vocabulario' in his guide to Matlatzinca, is, as far as I have been able to trace, unique in this period in the tradition of New Spain. After this treatise, the dictionary does not start immediately: several chapters follow, including a supplement to the 'Arte', so that it is difficult to ascertain whether Basalenque viewed 'Tratado' and 'Arte' as completely distinct, but it is more independent from its grammatical context than any other contemporary treatment of particles.

For an example of the missionary lexicographers' engagement with the multiplicity of derivational forms in some Mesoamerican languages, we turn to the Zapotec dictionary of Juan de Córdova. As we saw above, this is the most extensive of the early Mesoamerican dictionaries with Spanish as a source language. In the first *aviso* ('monition') in his prologue, Córdova explains that his dictionary is in fact 'more copious' than others. The main reason is that the speakers of Zapotec often have an enormous amount of terms in a given semantic field, which are not known in Western languages. He gives the example of the words for different sounds, such as the sounds for snakes when they move along, of birds, the beating of the heart, the cooking of water, and all possible 'interjections' humans can produce. Although the number of entries in Nebrija's dictionary related to sounds is impressive, the Zapotec dictionary adds even more entries that are not drawn from Nebrija. The second *aviso* of the prologue explains how the dictionary deals with derivational morphology. According to Córdova, it would not make sense to give all the 'derivations and compositions' for each verb, which would make his dictionary endless: almost forty 'vocables between nouns and verbs' can be derived from one single 'principal verb'. He includes a paradigm of the verb whose first-person singular is *tol löbaya* 'I sweep', with a list of all these forty forms.⁵² This example demonstrates that Córdova was developing his own strategies and lemmatization procedures, and supported his choices by theoretical explanations, different from Antonio de Nebrija, who also includes derivation patterns quite systematically, but without any theoretical explanation in his prologue.

⁵² This is found in the dictionary and in the grammar: Córdova, *Arte en lengua zapoteca*, fo. 36r.

The citation forms of verbs presented a more simple problem.⁵³ In Nebrija, the normal citation form is the infinitive. Since the infinitive is often lacking in Mesoamerican languages, the first-person singular is usually given, as in Latin: this was done by Molina for Nahuatl, by Córdova for Zapotec, and by Alvarado for Mixtec. In Tarascan, by contrast, the lexicographers could find an equivalent for the infinitive, and, since this form was available, they selected it as the citation form. Molina also ‘provides information about transitivity and reflexivity by listing after the verb a sample set of the prefixes it may take’, so that the verb *ihtoā* ‘to say something’ is listed in three successive entries with different prefixes and the appropriate different forms of the preterite.⁵⁴

Missionary lexicographers might provide information about features such as frequency, style, origin, dialect, or sociolect, although their marking was unsystematic. From Molina onwards, they showed an innovative interest in metaphorical usage.⁵⁵ Diaevaluative information is sometimes included (for instance, when insults are translated, the lemmata can allude to appreciative, derogatory, or offensive meaning, or taboo words). The use of the expressions ‘lo mesmo’ or ‘idem’ indicates that in the indigenous language no equivalent is given, and that instead the Spanish loanword is used. In Molina, this occurs in about 200 lemmata.⁵⁶ In an anonymous dictionary of Kaqchikel and K’iche’, a cross is used when a form is not only Kaqchikel but also K’iche’. Marking of dialectal or linguistic varieties is not, however, the rule, although diatopical variation is often described in missionary grammars. So, although the grammar of Mixtec of Alonso de los Reyes describes a great number of diatopical varieties in the prologue, it is surprising that in Alvarado’s dictionary of Mixtec, which appeared in the same year, almost nothing is found related to dialectal variation, beyond a preliminary reference to Alvarado’s desire to serve people of the high and low parts of the Mixteca region.⁵⁷ Grammatical categories were also marked unsystematically, for instance to distinguish *a* as a pronoun from *a* as an interjection.⁵⁸

⁵³ Overview in Smith-Stark, ‘Lexicography in New Spain’, 60–1.

⁵⁴ Karttunen, ‘Nahuatl lexicography’, 2658, citing Molina, *Vocabulario en lengua castellana y mexicana* (1571), fo. 43r (second sequence of foliation).

⁵⁵ Hernández, ‘La marca de uso metafórico’; see also Clayton and Campbell, ‘Alonso de Molina as lexicographer’, 347–8.

⁵⁶ Karttunen, ‘Nahuatl lexicography’, 2658.

⁵⁷ Alvarado, *Vocabulario en lengua mixteca*, fo. 3r, ‘aprouechar a los naturales de los pueblos de la Misteca alta y baja’; cf. Reyes, *Arte en lengua mixteca*, sig. ¶7r, ‘las grandes diferencias y modos distintos de hablar esta lengua’.

⁵⁸ M. Dürr and F. Sachse in *Diccionario k’iche’ de Berlín* (2017), 27.

Since Nebrija was undoubtedly the most important source for most lexicographers, his entry style was also the most common one in dictionaries translating from Spanish to a Mesoamerican language. In the tradition of Mayan lexicography, however, we find totally different lemmatization patterns. One example which illustrates this is Dürr and Sachse's analysis of the internal structure in the *Vocabulario en lengua 4iche otlatecas*.⁵⁹ In the lemma, the grammatical function of the word is often indicated first. When verbs are analysed, the author often adds information related to the conjugation class, combined with information on the first-person prefixes for transitive and intransitive verbs. Then, the equivalent meaning of the lemma is given in Spanish. In most entries, whole phrases are given as examples where the headword is used in a specific context, and when verbs are discussed several derivations of the verb are added as well. In other lemmata we also see related meanings, sometimes antonyms, with sketches of semantic fields related to the headword. We often also find observations related to ethnography. It is remarkable that many K'iche' examples are not translated into Spanish.

New Tools and Pedagogical Approaches

In New Spain, during the first two centuries, the documentation, teaching, and learning of the indigenous languages concentrated on the trilogy of dictionary, grammar, and catechism. These three genres were still in use until the end of the colonial period, but we see also that authors composed new learning methods. Most grammars and dictionaries in New Spain were composed by and for missionaries, but there is an exception: Pedro de Arenas published his successful *Vocabulario manual* of 1611 for use in a wide variety of communicative situations, as the subtitle shows: it contains 'the most common and ordinary words, questions and answers, which are used in communication between Spaniards and Indians'.⁶⁰

Only a few missionaries followed Arenas' method. Miguel de Guevara's 'Arte doctrinal y modo general para aprender la lengua Matlatzinga' of 1638 is one example. His method of teaching the Matlatzinca language brings paradigms and wordlists together with idiomatic expressions ('everyone is laughing at you!'), proverbs ('the one who lives well is always as lovely as a rose'),

⁵⁹ In *Diccionario k'iche' de Berlín* (2017), 21–2.

⁶⁰ Arenas, *Vocabulario*, title page, 'palabras, preguntas, y respuestas mas com[m]unes, y ordinarias que se suelen ofrecer en el trato y comunicacion entre Españoles é Indios'.

and similes ('the priests are walking from one village to another, like birds, today here and tomorrow there').⁶¹ Another remarkable text which does not follow the strict genres of the grammar, the dictionary, and the religious text is Diego de Nágera Yanguas' *Doctrina y enseñanza en la lengua maçahua* of 1637. The grammatical section is presented as a series of *advertencias* ('remarks'), rather than as a traditional grammar. As the title *Doctrina y enseñanza* indicates, the work is a mixture of grammar, texts for confession, liturgy, and a method for learning to speak in daily situations. As Dora Pellicer has put it, introducing a neologism for a new genre, Nágera Yanguas did not compile a 'diccionario' but a 'conversacionario'.⁶² The method is important seen from a pedagogical standpoint: learners could read the examples aloud, or they could learn them by heart. I am not aware that there circulated a dictionary parallel to Nágera Yanguas' method, since no dictionary of Mazahua has been preserved. What we know is that Nágera Yanguas included lexicographical sections, arranged thematically, in his work, for instance numbers, kinship terms, names related to the house, names of places in the city, and words to do with sewing, colours, and the parts of the body. The inclusion of a list of toponyms is not a very common practice, although there are earlier examples, such as the Mixtec grammar of Alonso de los Reyes, which has several lists enumerating regions. However, Nágera Yanguas did not copy from Reyes, but made his own list.

Augustín de Quintana's description of Mixe in his *Confessionario en lengua mixe, con una construccion de las oraciones de la doctrina christiana, y un compendio de voces mixes* is another experiment from the eighteenth century. His work is not a traditional 'Arte' which has to be combined with a 'Vocabulario' and religious texts in the teaching of Mixe, but a mixture of bilingual texts, detailed annotations, wordlists, and grammatical rules. The verbal index to Quintana's bilingual *Instruccion christiana* is, in effect, a Spanish–Mixe theological wordlist.

Content

The final topic to be analysed in this chapter is the content of the dictionaries. Again, they are far from uniform. In the more encyclopedia-like dictionaries we not only see many Western concepts which are translated into other languages, but also the other way around.

⁶¹ Guevara, 'Arte doctrinal' (1862–3), 218–19, 'Todos se ríen de tí . . . el que vive bien siempre está hermoso como la rosa . . . los sacerdotes andan de pueblo en pueblo como los pájaros hoy aquí y mañana allí.'

⁶² Pellicer, 'Confesión y conversación', 31.

Concepts from the Old World might be carried over into dictionaries of Mesoamerican languages. So, the entry *mezquita* ('mosque') in Nebrija is also included in the 1571 edition of Molina's *Arte*, where we even find Nahuatl translations: *Mahomacalli* ('it is the house of Muḥammad') and *Mahomatlatlatlauhtilizcalli*, a polysynthetic construct combining the root *Mahoma* with *tlatlatlauhtiliztli* ('prayer') and *calli* ('house'). Other lexicographers might explain *mezquita* with forms which included *Mahoma*, as Urbano did; or, in a spirit of hostility to Islam, as 'house of the devil', using the Spanish word *diablo* 'devil' as a loanword, as Gilberti did; or using forms which included words for 'devil' in a Mesoamerican language, as Córdova did.⁶³ Loanwords, calques, and literal translations were likewise used to express Christian vocabulary by lexicographers of K'iche'.⁶⁴ Mesoamerican gods and goddesses were compared with Greek and Roman analogies – for instance, Huitzilpochitli was 'another Mars' – and Nebrija's entries for kinds of divination in ancient Rome became the basis for entries in Córdova's dictionary which coined Zapotec words for what were in fact divinatory techniques of the ancient Romans.⁶⁵

On the other hand, missionary lexicographers often expanded an entry from Nebrija, adapting it to the the New World context. Nebrija's entry *altar donde sacrifican* ('altar where they sacrifice') became two entries in Alvarado's Mixtec dictionary, a Christian altar being *chiyo* (originally 'foundation') and a pagan one being *tayu quacu* or *tayu dzana*. In Córdova's Zapotec dictionary, a Christian altar became *pecógoláya nitaca missa* ('throne where Mass is held').⁶⁶ Using these translations, pre-Hispanic models were preserved. Likewise, Nebrija's *relumbrar, o reluzir* ('shine') is given four different Nahuatl equivalents in Molina's dictionary, expressing Nahua categories of reflected light (*tlanextia*, *peptlaca*, *pepetzca*, *tzotzotlaca*), so that 'a whole world of visual distinctions is opened up by these definitions: the luster of silk and feathers is distinguished from the sun glittering on the water and the sparkle of precious stones'.⁶⁷

⁶³ Urbano, *Arte breve dela lengua otomí y vocabulario trilingüe* (1990), s.v. *mezquita*, has *naxæcâmbemcangû. magînquexæcambenimahoma*, which is a calque of the Nahuatl translation by Molina (*mahoma* is the final element). Gilberti, *Vocabulario en lengua de Mechuacan*, s.v. *mezquita*, has *diabloeueri quahtaqueri* ('house of the devil'; elsewhere, *quahta* glosses *casa* 'house'). Córdova, *Vocabulario Castellano-Zapoteco* (1942), s.v. *mezquita*, has *lichi pezè láo* ('house of the devil'; elsewhere, *lichi* glosses *casa* 'house' and *pezèlâotào* glosses *diablo* 'devil').

⁶⁴ *Diccionario k'iche' de Berlín* (2017), 37–40. ⁶⁵ Hamann, *Translations of Nebrija*, 86–91.

⁶⁶ Hamann, *Translations of Nebrija*, 91–6. ⁶⁷ Hamann, *Translations of Nebrija*, 98–9.

Missionary lexicographers generally included more specific information related to the indigenous cultures in their dictionaries which translated from the indigenous language to Spanish. A clear example is the *Vocabulario otlatecas* of K'iche', which contains lemmata related to food, local religion, political and social organization, daily life, agriculture, how to participate in a social event, such as buying meat, or how to have a conversation or discussion about a certain topic.⁶⁸

As regarding their content, we can conclude that missionary dictionaries oscillate between two extremes: on the one hand they tried as best as they could to translate Western concepts, imposing their own culture on the indigenous languages, but on the other hand, most frequently in the dictionaries which translate from the indigenous language to Spanish, we find meticulous and surprisingly detailed information related to the culture of the other.

⁶⁸ *Diccionario k'iche' de Berlin* (2017), 41–8.

Missionary and Subsequent Traditions in North America

WILLEM DE REUSE

In view of the recent flurry of indigenous-language dictionaries published in North America, as well as the large amount of unpublished manuscripts in missionary and museum archives, this account will necessarily be selective. The only other historical account of the subject I know of is a short article in Franz Josef Hausmann et al.'s *Wörterbücher: ein Internationales Handbuch zur Lexikographie*.¹ Other basic references are a very thorough account of published and unpublished material on Eskimo-Aleut languages by Michael Krauss; Victor Hanzeli's *Missionary Linguistics in New France*, the most detailed account of missionary linguistics and archives of French-speaking Canada; M. Dale Kinkade's account of the history of research on North-west Coast languages; and Victor Golla's *California Indian Languages*, which has detailed sections on the history of research on California languages.² *Making Dictionaries: Preserving Indigenous Languages of the Americas*, a book on indigenous-language lexicography edited by William Frawley, Kenneth Hill, and Pamela Munro, is also of historical interest.

As for geographical coverage, I will cover the United States and Canada, as well as Greenland, which is geographically, though not politically, part of North America. I will not cover the Russian Far East, even though Eskimo languages are spoken there (their lexicography is covered by Krauss in the article cited above). I also cover the portion of northern Mexico which is non-Mesoamerican culturally, thus, the isolate Seri and the Uto-Aztecan languages O'odham, Pima Bajo, Northern Tepehuan, Eudeve, Opata, Yaqui, Mayo, and Tarahumara, but not the cluster of Uto-Aztecan languages further south; for the missionary lexicography of Mesoamerican languages, see Chapter 27.

¹ Rhodes, 'Lexicography of the languages of the North American Indians'.

² Krauss, 'Eskimo-Aleut'; Hanzeli, *Missionary Linguistics in New France*; Kinkade, 'History of research in linguistics'; Golla, *California Indian Languages*.

I start with the Catholic missionaries, then the Protestant missionaries. For the sake of space I will not be able to treat early American museum lexicography, and what I would like to call the 'Berkeley school of lexicography', by which I indicate the large body of indigenous-language dictionaries produced under the impetus of Mary R. Haas at the University of California, Berkeley. The next section is a brief discussion of the modern trends in indigenous North American lexicography, but without dictionary examples. I conclude with an illustrative account of the lexicography of the Western Apache language. While no single language can be representative of all the problems and pitfalls which early wordlist collectors, missionaries, and linguists have experienced, this detailed account provides the reader with a snapshot of what a complete history of the lexicography of one North American indigenous language would look like, including an exhaustive (as far as I know) coverage of early wordlists.

Catholic Missionary Lexicography

As noted by Edward Gray, one characteristic of Catholic lexicography is that it was intended for the missionary rather than for the indigenous people.³ The idea was for missionaries to learn the language so they could better minister to their charges, and could hear confession from them.

A characteristic of Catholic lexicography in the north, which distinguishes it from Mesoamerican lexicography, is that many of the dictionaries were not printed, and remained in manuscript. For example, if a dictionary from New France needed to be printed, it had to be printed in Paris. The Jesuits of New France requested a printing press, but never received one.⁴ The Franciscans of the California missions apparently did not get their dictionaries printed either, since only two dictionaries written by Franciscans saw print, later on in the nineteenth century.⁵ These were Felipe Arroyo de la Cuesta's 'vocabulary or phrase book' of Mutsun, and Buenaventura Sitjar's vocabulary of Antoniano Salinan.

Another characteristic of Catholic lexicography in the north, which also distinguishes it from Mesoamerican lexicography, is that the best early lexicographical work was carried out by Jesuits and Oblate missionaries, and not so much by Franciscans. This contrasts again with Mesoamerica, where there seems to have been more competition among the different

³ Gray, 'Missionary linguistics', 934. ⁴ Gray, 'Missionary linguistics', 933.

⁵ Bartholomew and Schoenhals, *Bilingual Dictionaries for Indigenous Languages*, 276.

orders (Franciscans, Dominicans, Augustinians, and Jesuits, among others) to produce the best materials. This might explain why an important recent proceedings volume, focusing on missionary lexicography, has only one article on North America.⁶

The work of the Jesuits in Canada and New France has been described in Hanzeli's *Missionary Linguistics in New France*, and in a useful summary by John Hewson.⁷ A list of their remarkable lexicographical achievements on the Algonquian languages of the area, some recently published, and some still in manuscript, includes dictionaries of Abenaki by Sébastien Rale, begun in 1691, and by Joseph Aubery in the mid eighteenth century; a dictionary of Miami-Illinois by Pierre-François Pinet, probably made in 1696–1700, and two slightly later dictionaries of the same language, one attributed to Jacques Gravier and one made by Antoine-Robert LeBoullenger; and dictionaries of Montagnais by Antoine Silvy, of c. 1678–84, and by Bonaventure Fabvre, extant in a manuscript dated 1696.⁸ As for their lexicographical achievements on the Iroquoian languages of the area, they include Pierre-Philippe Potier's impressive work on Huron from the mid eighteenth century, that of Jacques Bruyas on Mohawk from the later seventeenth or early eighteenth century; and an anonymous seventeenth-century dictionary of Onondaga.⁹

Not all of the missionary dictionaries of Algonquian and Iroquoian languages were by Jesuits, however. The Recollect missionary Gabriel Sagard published an influential dictionary of Huron as a separately paginated supplement, with its own title page, to his *Grand voyage du pays des Hurons* (1632).¹⁰ The linguistically sophisticated Sulpician priest Jean-André Cuoq stands out for having written dictionaries of the Algonquin language, a variety of Ojibwa, which belongs to the Algonquian family, as well as of the Mohawk

⁶ The volume is Zwartjes et al., *Missionary Linguistics IV*; the limited representation of North America therein is remarked on in Tomalin, 'Review'.

⁷ The latter is Hewson, 'Study of the native languages of North America: the French tradition'.

⁸ Rasles, 'Dictionary of the Abnaki language' (1833), dated 1691 before the first entries; Aubery, 'Dictionnaire abnaquis-françois' (manuscript), and Aubery, *French Abenaki Dictionary* (1995) (edition); [Pinet], 'French-Miami-Illinois dictionary' (manuscript), for which see Costa, 'St-Jérôme dictionary of Miami-Illinois'; [Gravier], 'Illinois-French dictionary' (manuscript), and Masthay, *Kaskaskia Illinois-to-French Dictionary* (an edition of Gravier's dictionary); LeBoullenger, 'French and Miami-Illinois dictionary'; Silvy, *Dictionnaire montagnais-français* (1974); Fabvre, *Racines montagnaises* (1970).

⁹ Works by Potier are edited in Fraser, *Huron Manuscripts*; editions of the other dictionaries are Bruyas, *Radices verborum Iroquaeorum* (1863), and *French-Onondaga Dictionary* (1860).

¹⁰ Modern editions are in Sagard Theodat, *Histoire du Canada* (1866), vol. IV, unpaginated appendix, and Sagard, *Grand voyage* (1998).

language, which belongs to the Iroquoian family, a language family quite different from Algonquian.

The Jesuits were also very active in northern Mexico. An extensive dictionary of Opatá, an extinct language, was published by Natale Lombardo in 1702, and a dictionary of Northern Tepehuan by Benito Rinaldini in 1743.¹¹ Thomás de Guadalupe's *Compendio del arte de la lengua de los tarahumares y guazápares* of 1683, the first account of the Tarahumara language, does not appear to contain a Tarahumara–Spanish vocabulary, although some references seem to imply it does.¹² The Tarahumara dictionary by the German-speaking Jesuit Matthäus Steffel (1809) is the next lexicographical work on this important language. These four early documents are presently being studied by Mexican scholars. According to Doris Bartholomew and Louise Schoenhals, the following seventeenth- and eighteenth-century Jesuits also compiled dictionaries although, as they point out, at least some of the attributions are uncertain: Juan Fonte wrote ‘something of a dictionary’ of Northern Tepehuan, which remained unpublished at his death in 1616; Tomás Basilio is said to have written a grammar and vocabulary of Yaqui not long afterwards (there is certainly reason to believe that the Yaqui catechism issued with the grammar and vocabulary was by him, though the linguistic works have also been attributed to his confrere Juan Bautista de Velasco); Jerónimo de Figueroa wrote a ‘Vocabulario copioso de la lengua tepehuana y tarahumara’ later in the seventeenth century; Adán Gilg, from Moravia, who flourished into the first decade of the eighteenth century, is said to have written a ‘Vocabulario de las lenguas eudeve, pima, y seris’; and the famous Jesuit pioneers Eusebio Kino and Jacobo Sedelmayer are said to have written on Pima (O’odham).¹³ Certain other anonymous dictionaries of the Pima Bajo and Eudeve languages were most likely by Jesuits.¹⁴ Jesuits were also active in Baja California and in Spanish Florida, although it appears they did not produce dictionary-size materials, or if so the materials no longer exist.¹⁵

¹¹ The original edition of Lombardo's work is rare, but there is a modern edition, Lombardo, *Arte de la lengua tegüima* (2009); Rinaldini's dictionary is available in a facsimile reprint of the 1743 edition.

¹² Guadalupe, *Gramática Tarahumara* (2010), is a modern edition.

¹³ Bartholomew and Schoenhals, *Bilingual Dictionaries for Indigenous Languages*, 275–6; for Basilio and the Yaqui catechism, see *Arte de la lengua cahita* (1890), viii and 236; this *Arte* is the work also attributed to J. B. de Velasco.

¹⁴ These dictionaries are available in modern editions: *Arte y vocabulario de la lengua dohema, heve o eudeva* (1982) and *Vocabulario de la lengua névome* (1979).

¹⁵ Bartholomew and Schoenhals, *Bilingual Dictionaries for Indigenous Languages*, 272–3.

In the Tarahumara area, lexicographic work by Jesuits has continued to the present, as evidenced by the dictionaries of Tarahumara and Spanish by David Brambila (1960, 1983).¹⁶ Jesuits have also been active in South Dakota, as evidenced by Eugene Buechel's Lakota dictionary of 1970, re-edited in 2002. While this work has been influential, the re-edition is not an improvement upon the original as, among other things, it introduces mistakes in the syllabification of words. This work is now supplanted by Jan Ullrich's *New Lakota Dictionary* (2011), a superb work, exemplary both in coverage and linguistic accuracy.

Jesuits were also active in other parts of the United States, such as Montana, as shown by Joseph Giorda's Kalispel dictionary, which suffered from a lack of understanding of Kalispel phonology, perhaps not surprising considering the phonetic complexity of a Salishan language such as Kalispel. Also not phonetically perfect, but otherwise very sophisticated linguistically and ethnographically, is Jules Jetté's monumental dictionary of Koyukon, brought to completion long after his death by a native speaker linguist, Eliza Jones, and published in 2000.

The Oblate missionaries (Oblates of Mary Immaculate) have been very active, particularly in Canada. Their lexicographical work in the nineteenth century included a Cree dictionary by Albert Lacombe (1874), and dictionaries of various Athabaskan languages and of Western Canadian Inuktitut by Émile Petitot (both 1876).¹⁷ An influential dictionary of Ojibwa by Bishop Frederic Baraga, who was not himself an Oblate, appeared in a new edition of 1878 'by a missionary of the Oblates', namely Albert Lacombe. In the twentieth century, we have Georges Lemoine's Montagnais dictionary of 1901 and the same Oblate's Ojibwa dictionary of 1911. The Oblate Adrien-Gabriel Morice, quite conversant in Athabaskan linguistics, wrote a monumental dictionary of the Carrier language. The influential *Dictionnaire esquimau-français* and *Dictionnaire français-esquimau* of Eastern Canadian Inuktitut by Lucien Schneider (both 1970, the former being a new edition of a work of 1966) have been re-edited and translated into English as *Ulinaisigutit* (1985). Not all the Oblate work is of the same high quality, however: for example, the popular Inuktitut dictionary by Arthur Thibert (1954, revised 1958) suffers from a poor spelling system.

In North America, the Franciscans have produced Navajo dictionaries. The collectively produced *Ethnologic Dictionary of the Navaho Language* and

¹⁶ Discussed in Bartholomew and Schoenhals, *Bilingual Dictionaries for Indigenous Languages*, 289–90.

¹⁷ Lacombe's dictionary is discussed in Wolfart and Carroll, *Meet Cree*, 114–15.

Vocabulary of the Navaho Language, of 1910 and 1912 respectively, were important from an ethnographic point of view but suffered from a poor understanding of Navajo phonology. The dictionaries by a later Franciscan, Berard Haile, written under the influence of the celebrated linguist Edward Sapir and published in 1950 and 1951, remedied this problem, and are still useful reference works for linguists. These early Franciscan dictionaries have not been entirely superseded by the superb, monumental, and linguistically sophisticated dictionaries by Robert Young and William Morgan Sr, published in 1980 (revised 1987) and in 1992. I am not aware of any early Catholic lexicographical work among the Pueblo or Apache groups of the American south-west.

Protestant Missionary Lexicography

Early Protestant missionaries in New England tended to write fewer dictionaries, since, to the extent that they were interested in the native languages, they concentrated on translating Scripture for the indigenous peoples to read.

In a class by itself is Roger Williams' *Key into the Language of America* of 1643, a phrasebook of Narragansett, which contains a lot of interesting lexical material and is sympathetic to indigenous culture.¹⁸ The Puritan missionary Josiah Cotton wrote an early vocabulary of Massachusetts or Natick, which was eventually published by the nineteenth-century lawyer and philologist John Pickering (whose own lexicographical interests extended from the indigenous languages of North America to the English of the United States, and also to Ancient Greek). In the same vein, James Hammond Trumbull composed a *Natick Dictionary* (1903) on the basis of the writings of the well-known seventeenth-century New England missionary, and pioneer linguist, John Eliot.

There is a strong lexicographic tradition in Greenlandic which started with the Lutheran missionary Poul Hansen Egede's Greenlandic–Danish–Latin dictionary of 1750, and continues with the works of the Lutheran missionary Otto Fabricius (1804) and of the Moravian missionary Samuel Petrus Kleinschmidt (1871). Indeed, Greenlandic lexicography continues to this day, although it is now undertaken by secular linguists. Of the Moravian missionaries, by far the most perceptive linguistically was Kleinschmidt, but generally their lexicographical output was copious and of high quality. Johann Jacob Schmick's work on Mahican, edited by Carl

¹⁸ Gray, 'Missionary linguistics', 934.

Masthay in 1991, stands out. Two other productive Moravians in the eastern United States were John Gottlieb Heckewelder and David Zeisberger. A sample of the work of the Moravians might include the Lenape or Delaware dictionary from the Moravian archives at Bethlehem, Pennsylvania, edited by Daniel Garrison Brinton and Albert Seqaqkind Anthony; Heckewelder's manuscript 'Mahicanni words, taken down from the mouth of one of that nation, who had been born in Connecticut'; and Zeisberger's English–German–Onondaga–Delaware dictionary, posthumously published in 1887.

In the south-east of the United States, there were influential dictionaries of Choctaw (1915, reprinted 1973 and 1978) by Cyrus Byington, a missionary affiliated with the largely Congregationalist American Board of Commissioners for Foreign Missions, and of Muskogee or Creek (1914, reprinted 1964) by Robert Loughridge, a Presbyterian missionary, and David Hodge, a member of the Creek Nation. Both of those dictionaries were widely used by people who wanted to read the Bible and other material in the native languages. Interestingly, I do not know of an early Cherokee dictionary, even though the Cherokees achieved fame for their high level of native-language literacy in the nineteenth century. Farther west, there were Mennonite missionaries such as Rodolphe Petter, who wrote an English–Cheyenne dictionary, a work of 1,127 pages printed at his home, by his son, from 1913 to 1915; it does not appear that the promised Cheyenne–English counterpart ever saw publication. The Presbyterian missionaries in Minnesota wrote influential dictionaries of the Dakota language: a bidirectional dictionary 'by the members of the Dakota Mission' was edited by Stephen Riggs and published in 1852, with an extensively revised version of the Dakota–English portion in 1890, followed by an English–Dakota dictionary by John Williamson in 1902. To conclude the account of these early missionaries, it is interesting to note that some of the missionary works were actually published by the Smithsonian Institution of Washington, DC, as was the Dakota dictionary of 1852, or by its Bureau of American Ethnology, founded in 1879, as were the *Natick Dictionary* of Trumbull in 1913 and the *Dictionary of the Choctaw Language* of Byington in 1915.

Some of the most productive lexicographers in North America are the linguists of the Summer Institute of Linguistics (SIL), now called SIL International. This is a Protestant organization which has always concentrated on literacy rather than on explicit proselytism or church building. Most of its members' work on American indigenous languages has been in Mesoamerica and in South America, but their North American output is

not negligible. The dictionaries published by SIL tend to be short, but the spelling system is almost always a reliable one. Here are a few samples. The Inupiaq dictionary by Donald Webster and Wilfried Zibell (1970) was the first dictionary on this Eskimo variety with a consistent spelling, although the Inupiaq spelling was modified somewhat in recent years; the Carrier dictionary by Francesca Antoine and members of the Carrier Dictionary Committee with two linguists from the SIL contains helpful illustrative sentences; the *Tlingit Verb Dictionary* and the *Tlingit Noun Dictionary* by Gillian Story and Constance Naish (1973, 1976) are among the earliest linguistically sophisticated works on this complex language; the Papago (now called O'odham) dictionary by Dean Saxton and Lucille Saxton (1969) was liked for its illustrations, which unfortunately were removed in the later edition (1983).¹⁹ For northern Mexico, SIL has published a dictionary of Seri by Edward Moser and Mary Moser (1961), now supplanted by a much more impressive work also by SIL linguists, Mary Moser and Stephen Marlett (2005, ²2010); a dictionary of Tarahumara by K. Simon Hilton (1959, substantially revised 1993); and a dictionary of Mayo by Howard Collard and Elisabeth Scott Collard (1962). I will discuss SIL's Western Apache dictionary in the last section of this chapter.

For curiosity's sake, let me mention a Hebrew–Dakota dictionary, compiled by Samuel Pond in 1842, and a Hopi vocabulary written in the Deseret alphabet, a writing system at one time proposed by the members of the Church of Latter-Day Saints (or Mormons).²⁰

Modern Trends

Chronologically overlapping with the Berkeley school mentioned at the beginning of this chapter, and not entirely distinct from it, there is a new resurgence of modern dictionaries, which started around 1984–5, years which saw the publication of the work of Rosalie Bethel et al. on Mono; Steven Jacobson on Yup'ik Eskimo; Philip LeSourd on Passamaquoddy; and the pioneering work by Richard Rhodes on Ojibwa.²¹ Less ambitious works published at the same time were Roy Albert and David Shaul's *Concise Hopi*

¹⁹ The members of the Carrier Dictionary Committee who worked on the Carrier dictionary were Francesca Antoine, Catherine Bird, Agnes Isaac, Nellie Prince, and Sally Sam; their co-workers from the SIL were Richard Walker and David B. Wilkinson.

²⁰ Plaut, 'Hebrew–Dakota dictionary'; Beesley and Elzinga, *1860 English–Hopi Vocabulary*.

²¹ For Bethel et al. on Mono, see Kroskrity, 'Language renewal', 178–80; for Rhodes on Ojibwa, see Rhodes, 'Multiple assertions'.

and *English Lexicon* and P. David Seaman's *Hopi Dictionary*, both now supplanted by the *Hopi Dictionary/Hopiikwa Lavàytutuveni* (1998) produced by the Hopi Dictionary Project, the most impressive and scholarly-looking of all North American indigenous language dictionaries to date.²² This is a trend accelerating in the twenty-first century, probably fuelled by recent electronic and computational technical advances.

Indeed, whereas the previous trends of museum lexicography and the Berkeley school tended to focus on dictionaries for linguists only – for linguistic typological, historical-comparative, or areal study, and for reading texts collected mostly by other linguists – the focus has now changed to make dictionaries more accessible to heritage learners, typically non-linguists, and to work together with indigenous communities towards language revitalization or even revival. This new trend is accompanied by new issues that linguists never really had to deal with in the past, such as 'orthography wars' (called 'community debates' in politically correct parlance), and questions such as 'how do we make more dictionaries accessible and/or portable for indigenous people?' or 'how do we make dictionaries more accessible to learners, including child learners?' or 'how does one respectfully deal with communities who do not want their language written down, or simply do not want any dictionaries?' The frustrations go both ways, as discussed in more detail in Frawley, Hill, and Munro's *Making Dictionaries*.

Lexicographers often attempt to be as comprehensive as possible before publication, thereby delaying availability to endangered-language speakers. They often make the indigenous–English portion of the dictionary the basic one, so that the English–indigenous section is little more than an index to the indigenous–English portion. They use complex linguistic terminology. They tend to make even a pedagogical dictionary more etymologically sophisticated than is required, so that it ends up being used more by comparativists than by the speakers themselves. There is also a lack of acknowledgement that indigenous-language bilingual dictionaries are not so much used for checking unknown words or their spellings, but rather as phrasebooks and guides to language learning curriculum development.

Indigenous users experience the following sorts of frustration. They expect the dictionary to be in a standard spelling but, if there is no standard spelling, there will be arguments over what a correct spelling should be. Inexperience

²² See Hill, 'On publishing the *Hopi Dictionary*'. Contributors to the Hopi Dictionary Project included Kenneth C. Hill, Project Director and Editor-in-Chief; Emory Sekaquaptewa, Cultural Editor; Mary E. Black, Associate Editor; Ekkehart Malotki, Senior Contributing Editor; and Michael Lomatuway'ma, Contributing Editor.

with the procedure of devising a spelling by consensus or committee can be a problem. They will also be impatient with the fact that the production of a dictionary takes so long (people say things like, ‘what can be so hard about it? You take all the words of the language and you put them in alphabetical order!’). If there is no simple alphabetization technique, they might be frustrated by their ignorance of the complex linguistic basis on which the alphabetization rests, and will tend to consult only the English–indigenous index. In the case of a dictionary available as an electronic database, there might be insufficient familiarity with computers on the part of the elderly speakers (who are often the best), or frustration with outdated computer systems.

The History of Western Apache Lexicography as an Illustration

Western Apache (henceforth WA) refers to the mutually intelligible varieties of Apache presently spoken on five reservations in the state of Arizona. The WA language is part of the Apachean or Southern Athabaskan branch, a clearly definable subgroup of the Athabaskan (or Dene) language family (Athabaskan-Eyak-Tlingit or Na-Dene phylum). The other Apachean languages are Navajo (also called Diné), spoken in Arizona, New Mexico, and Utah; Chiricahua-Mescalero Apache, spoken in New Mexico; Jicarilla Apache, spoken in New Mexico; Lipan Apache, originally from Texas, but now dormant; and Plains Apache (formerly called Kiowa-Apache), originally from Oklahoma, also now dormant. There is very marginal mutual intelligibility between WA and Navajo, its Athabaskan neighbour to the north, and between WA and Chiricahua Apache, originally its Athabaskan neighbour to the south-east. The Apachean languages were originally part of a U-shaped dialect chain, with the westernmost sections (Navajo and WA) most influenced by Puebloan cultures, and the easternmost sections (Lipan and Plains Apache) part of the southern plains cultural area.²³

The ethnic population of all WA groups is about 20,200, according to *Ethnologue*. The number of speakers of WA can presently be estimated at 5,000–6,000. This includes semi-speakers, so it is not clear exactly how many fully ‘fluent’ speakers exist. Following Grenville Goodwin’s widely accepted ethnographic classification, five groups are distinguished within WA: Northern Tonto, Southern Tonto, Cibecue, San Carlos, and White Mountain.²⁴ Through

²³ Rice, ‘Linguistic evidence regarding the Apachean migration’, 255.

²⁴ Goodwin, *Social Organization of the Western Apache*.

my own research, focusing on phonological differences and lexical data in the varieties spoken by these groups, and informed by ethno-historical considerations, I have concluded that it is possible to identify three geographic groups of WA dialects.²⁵ This is in spite of the considerable dialect mixture stemming from clan exogamy requirements that stimulate marriages between people of different groups. These three groups, which cross-cut reservation boundaries, are listed below in order of speaker number size:

- (1) An Eastern group, corresponding to Goodwin's White Mountain group, as well as the Coyotero of the Bylas area of the San Carlos Reservation.
- (2) A Central group, roughly corresponding to Goodwin's San Carlos and Cibecue groups and probably containing most of the Southern Tonto group as well.
- (3) A Northern group, roughly corresponding to Goodwin's Northern Tonto group and maybe some Southern Tonto as well.²⁶ I will refer to this group as Dilzhe'e.

The Eastern and Central groups each have approximately 2,000 to 3,000 speakers, but for both groups fewer than 1 per cent of the children entering kindergarten speak the language. The fact that the overwhelming majority of children now enter school as English-speaking monolinguals projects low vitality in the near future for WA as a whole and is cause for serious concern. The vitality of the Northern group (the Dilzhe'e) is an even greater cause for concern. With a few exceptions, the Dilzhe'e varieties have not been spoken to children for more than fifty years, and as a result only the grandparent and great-grandparent generation are fluent speakers. The communities involved are much smaller to start with and, additionally, are engulfed in an Anglo-American majority population. So, while both the Eastern and the Central groups have several thousand speakers each, the Northern dialect group probably has fewer than twenty-five speakers. With one or two exceptions, the speakers of Dilzhe'e are all elderly.

The table lists all pre-1964 lexical material, ordered by date recorded: the year 1964 is the date of the first scientific study of WA, Marshall Durbin's 'Componential analysis of the San Carlos dialect of Western Apache', and also the time at which the spelling system tends to become the relatively fixed SIL system.

²⁵ De Reuse, 'New classification of Western Apache dialects'.

²⁶ Mierau, 'Concerning Yavapai-Apache bilingualism'; de Reuse, 'Tonto Apache and its position within Apachean'; de Reuse, 'Yuman phonetic or phonological influences on Apachean'; Rocha, 'Dilzhe'e Apache learners' dictionary'.

Pre-1964 lexical sources for Western Apache, ordered by date recorded

Writer	Date	Place	Source	Variety or group
A. W. Whipple	1851×1853	[uncertain]	Whipple, 'Vocabularies' (1855), 81–4	Pinaléño [SC]
E. Palmer	before 1867	on campaign	Palmer, 'Pinella and Aviva Apache vocabulary'	Pinaléño and Aravaipa [SC]
N. S. Higgins	1866	on campaign	Higgins, 'Ethnographic notes and vocabulary', 23–9	Sierra Blanca [WM], Pinal [SC], Coyotero
C. Smart	1866	Camp McDowell	Smart, 'Imperfect vocabulary of Tonto or Coyotero Apache'; see Hagemann, 'Surgeon Smart and the Indians'	Coyotero or Tonto [D]
W. L. Sherwood	1869×1872	on campaign	Sherwood, 'Sierra Blanca and Coyotero vocabulary'	Sierra Blanca and Coyotero [WM]
J. G. Bourke	c. 1872–1886	on campaign	Bourke, <i>Vocabulary of the Apache or 'Inde' Language</i>	Sierra Blanca [WM] and SC
O. Loew	1873	Camp Apache	Gatschet, <i>Zwölf Sprachen</i> , and Gatschet, 'Classification'	Aravaipa [SC]
J. B. White	1873–1875	San Carlos	White, 'Vocabulary of the Apache and Tonto language' and White, 'Corrected wordlists'; partially published in Gatschet, <i>Zwölf Sprachen</i>	[WM?]
G. K. Gilbert [anonymous]	1874 c. 1880	Camp Grant Fort Apache?	Gatschet, 'Classification'	Aravaipa [SC]
A. Gatschet	1883	Hampton, VA	Hough, 'Vocabulary collected from the White Mountain Apache'	WM
C. Ruby	1886	[uncertain]	Gatschet, 'Pinal Apache'	Pinal [SC]
			Ruby, 'Chiricahua Apache vocabulary'	Chiricahua [but actually WM]

J. Plocher	starting in 1893	[uncertain]	Plocher, 'English Apache dictionary'	SC, some WM
E. S. Curtis	c. 1900	[uncertain]	Curtis, <i>North American Indian</i> , I.139-44	[Coyotero?, WM?]
A. B. Reagan	c. 1901-1903	Cibecue	Reagan, 'Grammar treatise of the White Mountain Apache Indian language'	WM
B. Freire-Marreco	1913	Fort McDowell	Freire-Marreco, 'Tonto Apache vocabulary and information'	'Tonto' [D]
F. J. Uplegger	c. 1919-1960	San Carlos, Peridot, Bylas	Uplegger, 'San Carlos Apache dictionary' and 'Thesaurus of Apache language forms'	SC [and some B]
G. Goodwin	1927-1940	many locations	Goodwin, 'Word lists, phonology, conjugations'	[all varieties, but mostly B]
H. Hoijer	c. 1933	[uncertain]	Hoijer, 'San Carlos Apache linguistic notes'	SC [but probably Southern Tonto]
R. Young	c. 1939	[uncertain]	Young, 'Comparative English-Athabaskan vocabulary'	[WM]

Note: C = Cibecue; D = Dilzhe'e (formerly called Tonto); SC = San Carlos; B = San Carlos, town of Bylas; WM = White Mountain.

The earliest contacts by Europeans with speakers of WA must have been with Spanish (and later Mexican) military and missionaries, but the early Spanish-language sources, although quite informative regarding Apache names of early groups and Apache leaders' names, have left us with no clear evidence that WA names were being recorded, nor did Spanish-speakers elicit any wordlists.²⁷

The first wordlists of WA were compiled in the latter half of the nineteenth century by United States military or government agents: Higgins, Hough, Palmer, Ruby, Sherwood, Smart, Whipple, and White. Apart from Whipple's printed text, all of these are in the National Anthropological Archives at the Smithsonian Institution. The spellings of these earliest documents are unsatisfactory in the extreme, and, with the exception of White's 'Vocabulary' and 'Corrected wordlists', they are only a few pages long. Nevertheless, they are important, because they represent older stages of the language.

In this period, we also encounter the first documents by United States military and government agents with some anthropological training, and by the earliest anthropologists in the area. Two works stand out. Captain John Gregory Bourke's *Vocabulary of the Apache or 'Inde' language*, collected from information provided by Western Apache scouts during the campaigns of General George Crook against the Chiricahua in the 1870s and 1880s, is the most interesting lexical source from the nineteenth century, and Albert B. Reagan's manuscript 'Grammar treatise of the White Mountain Apache Indian language; also a vocabulary of one thousand words of the same language with their English equivalents appended thereto', written in 1903, contains intriguing lexical material on cultural and religious matters.²⁸ Other wordlists were recorded by Albert Gatschet from the Smithsonian Institution, who also published some earlier short vocabularies, and by the anthropologists Edward Sheriff Curtis and Barbara Freire-Marreco.

The anthropologist Grenville Goodwin wrote extensive notes on all aspects of WA life.²⁹ In these notes, there are lexical data classified by semantic domain and by verb paradigm, which amount to about 557 pages. Goodwin's spelling of WA words was not very accurate, as phonemic distinctions of vowel length and tone, as well as some distinctions in the consonant system, were not made.³⁰ In these lexical data, there is evidence that Harry Hoijer, to whom we shall return very shortly, coached Goodwin

²⁷ De Reuse, 'Apache names in Spanish and early Mexican documents'.

²⁸ For Bourke, see Porter, *Paper Medicine Man*.

²⁹ See Ferg, *Western Apache Material Culture*.

³⁰ See Opler, *Grenville Goodwin Among the Western Apache*, 28–9.

in transcribing in his own reliable system, and even wrote some forms into his notes himself.

Two early unpublished dictionaries were compiled by Lutheran missionaries, originally from Germany. The earliest one is the English–WA dictionary of Johannes Plocher (1893), a manuscript of 266 pages. The second, much larger, dictionary, also English–WA, was compiled by Francis Uplegger, who worked on the San Carlos Reservation starting in 1919, between approximately 1919 and 1960, and is 498 sheets long. It is not clear whether Plocher's manuscript had an influence on Uplegger's. It is particularly interesting for its attempt to list paradigms for major verbs. Uplegger's spelling system indicates all phonemic distinctions, but it does not do so consistently. However, his system is important, because it is the only one (previous to Willem de Reuse and Phillip Goode's *Practical Grammar of the San Carlos Apache Language* of 2006) to indicate the phonetic differences between low, mid, and high tones.

The noted Athabaskanist Harry Hoijer, a student of Edward Sapir, carried out the first accurate linguistic work on WA. His manuscript linguistic notes on the San Carlos dialect of WA (possibly Southern Tonto rather than San Carlos) consist of 1,912 typewritten 4" by 6" fileslips, roughly alphabetized by stem-initial consonant. Most fileslips contain one verb base, with complete inflectional paradigms; other fileslips contain nouns, postpositions, or particles, with morphological analyses. Hoijer's notes are the most linguistically informed and morphologically detailed of the existing lexical collections on WA and are a necessary point of departure for analytic lexicographic work.

In the 1930s, Robert Young, the celebrated expert in Navajo language and linguistics, collected an Apache vocabulary as part of a larger comparative Athabaskan vocabulary project. It is clear that the forms are in the White Mountain variety, but Young identifies the list simply as 'Apache' and provides no consultant data. The main interest of this vocabulary is that it is written by someone with an excellent ear for Athabaskan languages, although it reveals some phonetic biases, due to Young's profound knowledge of the Navajo language.

The first linguistically accurate published dictionaries are the direct or indirect product of linguists affiliated with the Summer Institute of Linguistics. The *Western Apache Dictionary*, identified as having been compiled by the 'Staff of the White Mountain Apache Culture Center' (1972), was largely written by SIL linguist Faith Hill. It is in English–WA format only; it is quite short (c. 4,000 entries); and it contains very little morphological information. Verbs are typically given in the third-person singular subject

imperfective mood only. It does contain useful and illustrated appendices on pronunciation, a grammatical sketch with a few sample paradigms, plant names, words borrowed from Spanish, kinship terms, names of the months, and times of the day or night. The work of Margaret Quay et al. (1987), mostly circulated in photocopy, took the same material and re-alphabetized it according to the WA words, in effect creating a WA–English index to the 1972 dictionary.

The latest published lexicographic work on WA is the elegantly produced *Western Apache–English Dictionary: A Community-Generated Bilingual Dictionary* (1998), edited by the late Dorothy Bray, formerly from Central Arizona College, and written in collaboration with the White Mountain Apache Tribe. This is a substantial work, with more than 10,500 WA entries in the WA–English section, and therefore deserves a more extensive critique.

This dictionary started as an expanded version of the *Western Apache Dictionary* of 1972, and mentions Faith Hill as a linguistic adviser. Entries added are words from transcriptions of spoken materials, business terms, place-names, and words from the WA translation of the New Testament. The number of 10,500 entries is deceptive, however, since the large majority of the added entries are inflectional forms of items already given, in another inflectional form, in the 1972 dictionary. Whereas the 1972 dictionary had ‘love, to, *bił nzhq̄q̄*’, which is understood as actually being: ‘s/he loves him/her’, and one example, ‘*shił nzhq̄q̄* I love him’, the Bray dictionary has, as seemingly random separate entries:

bił ch’izhq̄q̄ his love (for a person)
bił danjq̄q̄ he loves us
bił danlj̄q̄q̄ he loves us
bił danolshq̄q̄ni those who love you (PL)
bił danzhq̄q̄yúgo if they love him
bił nshq̄q̄ he loves material things
bił nzhq̄q̄ he loves him
bił nzhq̄q̄ni the one he loves
shił daanolshóni you (PL) whom I love
shił danohshóni you (PL) whom I love
shił danolshq̄q̄ I love you (PL)
shił nzhq̄q̄ I love him.³¹

³¹ [Hill et al.], *Western Apache Dictionary*, 46; Bray et al., *Western Apache–English Dictionary*, 45–7 (entries in *bił*), 241–2 (entries in *shił*).

These are completely derivable, either as inflections, or through the addition of clitics, or as dialect variants, from the entry given in the 1972 dictionary. As seen in the example above, the work also contains many entries that are phrases with clitics and, since these constructions are perfectly predictable, they do not belong in a dictionary. As David Samuels' review points out, 'The effect – and I know this was not intended by the editors – is to make English appear logical and organized, whereas Western Apache seems to be somewhat arcane, with lots of phrases for talking about "looking", but no actual word for the action.'³²

Furthermore, it is not possible to find a principle according to which one inflected form of a word occurs but another one is omitted. This is unfortunate, since the lexicographer needs a principled way for choosing which inflected word counts as the entry, and a principled way to relate the inflected words to each other. For example, if one looks at the English–WA section, which happens to bring these inflected forms together under one English word, one finds, under the verb *like*, WA expressions for 'do you like him', 'they like it', and 'he likes it'. Under the word *love*, one finds WA expressions for 'he loves us', 'those who love you', 'if they love him', 'he loves material things', 'he loves him', and several more (as shown above). This conceals the fact that the inflectional potential of the verbs meaning 'to love' and 'to like' is the same in WA.

In sum, despite its larger size, Bray's dictionary is a marginal improvement upon its predecessor. Native speakers of WA have repeatedly voiced their frustration to me about finding useful information in this work. Neither its size nor its design would justify calling it a scientific or comprehensive dictionary. Other Apacheanists will have to do better.

³² Samuels, 'Review', 217.

Missionary Traditions in East Asia

OTTO ZWARTJES

East Asia, which for our purposes includes the Philippines, Vietnam, Japan, and China, is not an area defined by a uniform culture. Before Christian missionaries arrived there in the sixteenth century, it was the home to the long and complex Chinese lexicographical tradition, and to those of what may be called the Chinese periphery; these are discussed from their beginnings to 1700 or 1800 in Chapters 3, 6, and 10. These older traditions continued, variously modified, after contact with the new traditions of missionary lexicography: for that of Chinese, see Chapter 15, and for those of Japanese, Korean, and Vietnamese, see Chapter 16.

A result was that missionary lexicographers in East Asia, unlike their counterparts in the Americas, Africa, and Australasia, could benefit from existing educational systems and methods, particularly of teaching Japanese and Chinese. So, many dictionaries produced by missionaries in East Asia share a distinctive set of features. Of course, some basic features of European origin were shared by all missionary dictionaries, but generally the works in Asia form a highly diverse, but coherent, group.

Whereas the first missionary dictionary from Mesoamerica was compiled around 1545, and the first from South America was published in 1560, the extant record of missionary lexicography in the Philippines, China, and Japan begins towards the end of the sixteenth century. As in the Andes (but not in Mesoamerica), the Jesuits were the most important lexicographers in much of East Asia. In the Philippines, we also occasionally see lexicographical work carried out by Jesuits, such as Juan de Noceda and Pedro de San Lúcar's dictionary of Tagalog, but the most important works from the first two centuries of the colonial period were written by Franciscans and Augustinians.¹ The following overview will present the missionary

¹ Fernández Rodríguez, 'Lexicography in the Philippines', 5.

lexicography of Chinese from the Philippines and China, then of other languages of the Philippines, then of Japanese, and, finally, of Vietnamese.

Historical Overview

The first missionary dictionaries of any variety of Chinese were very possibly compiled in Manila – where there was a Hokkien-speaking community, whose members were sometimes referred to as the Sangleys – rather than in China itself.² The Hokkien spoken there may be distinguished from the Hokkien of mainland China by calling it Early Manila Hokkien. The first dictionary of this variety of which there is a clear report is the lost *Arte y vocabulario de la lengua china* attributed to Martín de Rada, who was in the Philippines from 1565 to 1578.³ Henning Klöter has tabulated records of twelve other missionary dictionaries of Hokkien and other Mǐn varieties composed between the sixteenth century and eighteenth, of which six have survived.⁴

The oldest extant dictionary of Hokkien from the Philippines is the *Dictionarium Sino Hispanicum* of Pedro Chirino, which survives in two manuscripts, one of them dated 31 March 1604 in the compiler's hand.⁵ The dated manuscript consists of eighty-four folios, with an average of thirteen entries on each, ordered from top to bottom and from right to left; the Chinese character which is the most conspicuous part of each entry is accompanied by a transcription and a translation. The order of entries is not always clear, though the first five folios show some grouping by radicals, or semantic constituents of the character, in a Chinese tradition going back 1,500 years to *Shuōwén jiězì* (see Chapters 3 and 6), and there is some subject grouping thereafter.

A more extensive Hokkien–Spanish dictionary from a slightly later date, the title of which began *Diccionario de la lengua Chincheo*, survived in manuscript until the nineteenth century, and not only contained characters representing words, but also phrases and complex expressions.⁶ Another, the 'Bocabulario de la lengua sangleya por las letraz de el A.B.C.', which is alphabetically ordered and presents nearly all of its Hokkien material in transliteration, is still extant, and can be dated around 1617–18 on internal evidence.⁷ These are all fairly limited dictionaries, of somewhat over a thousand entries, but other early Hokkien dictionaries compiled by missionaries in the Philippines no later than the 1630s run to as many as 20,000

² Klöter, *Language of the Sangleys*, 9, 32–3. ³ Klöter, *Language of the Sangleys*, 52.

⁴ Klöter, *Language of the Sangleys*, 53–6. ⁵ Klöter, *Language of the Sangleys*, 56–62.

⁶ Klöter, *Language of the Sangleys*, 66–8. ⁷ Klöter, *Language of the Sangleys*, 68–73.

entries.⁸ After the 1630s, the Dominicans who had been making dictionaries of Early Manila Hokkien extended their missionary activities to the mainland and turned their lexicographical energies to Mandarin.

Although the missionary lexicography of Chinese may have begun with the Hokkien of Manila, the earliest extant dictionaries are of Mandarin and come from the Chinese empire. The first of these is an untitled dictionary in which Portuguese (and some Italian) headwords are followed by Chinese equivalents, first Romanized and then in Chinese characters. It is attributed to the Jesuits Michele Ruggieri and Matteo Ricci (though doubts have been raised: see the reference to the dictionary in Chapter 15) and was probably prepared between 1583 and 1588 when Ruggieri and Ricci were in Zhaoqing. It remained in manuscript until its rediscovery by Pasquale D'Elia in the twentieth century; a facsimile has been published under the title *Dicionário Português–Chinês*.

The Mandarin–Spanish *Vocabulario de letra China con la explication castellana hecho con gran propiedad y abundancia de palabras* of Francisco Díaz, completed around 1642, is extant in several manuscripts.⁹ The one in the Jagiellonian Library at Kraków presents 7,169 characters, arranged 12 to a page, in 3 rows and 4 columns, a layout resembling that of Chirino's *Dictionarium Sino Hispanicum*, but unlike that of any European dictionary of the time. The lemmata are arranged alphabetically according to the Romanization of the Chinese character, so that the first three elements in an entry are Romanization; original character; and Spanish equivalent. Most lemmata also have some examples with context, and sometimes grammatical information or references to a lost *Arte*.¹⁰ Another manuscript of the same dictionary, of 199 folios, now in the Bodleian Library at Oxford, presents 6,831 characters, of which approximately 1,123 are neither transliterated nor translated, in a somewhat different layout.¹¹ A significant proportion of the equivalents are Portuguese rather than Spanish.

Francisco Varo, who is best known for his posthumously published grammar, *Arte de la lengua mandarina*, also compiled two dictionaries of the same language, a Portuguese–Mandarin 'Vocabulario da lingua mandarina' of 1670 (with some translations into French) and a Spanish–Mandarin 'Vocabulario

⁸ Klöter, *Language of the Sangleys*, 73–4.

⁹ Zwartjes, 'El Vocabulario de letra china de Francisco Díaz', 74–81.

¹⁰ References to the *Arte* are given in Kraków, Jagiellonian Library, Ms Berol. Ms Sin 13. Hisz. chin., XVII, fos. 274, 289, 509, and 541. Entry count from the colophon of the manuscript.

¹¹ Oxford, Bodleian Library, MS Marsh 696, discussed in Zwartjes, 'El Vocabulario de letra china de Francisco Díaz', 75–6.

de la lengua mandarina con el estilo y vocablos conque se habla sin elegancia' of 1677–87.¹² He wrote in the prologue to the latter that his aim had been 'arranging in order and collecting the terms which are only used in speaking, adding many which are missing in the other glossaries, and not including the words which are used in writing'. He added that words used in writing rather than speech 'can be looked up in other works, in particular in the *cabecillas* ["small heads": see below] compiled by the venerable Father Fray Francisco Dias, priest of our religion, which by poll of all those who have seen it is the best work compiled for this purpose in this kingdom'.¹³ Although Varo referred to the work of Díaz with such respect, his own differed from it not only in content but also in macrostructure. Indeed, the macrostructures of the missionary dictionaries of Chinese are exceptionally diverse, even more so than those from the Mesoamerican tradition.

In the Philippines, at least seven dictionaries of Philippine languages were written between 1580 and 1610, but all of these early works appear to have been lost.¹⁴ Three survive from the seventeenth century, and eight, including two reprints, from the eighteenth; again, there are records of a number of lost dictionaries from both centuries.

The first extant dictionary of Tagalog, and indeed of any Philippine language, is the *Vocabulario de la lengua tagala* of Pedro de San Buenaventura, published in 1613, offering about 16,350 Spanish headwords and about 14,500 Tagalog equivalents.¹⁵ Its Tagalog–Spanish section is only an index of roots which leads the user from a root back to the Spanish–Tagalog entry: far from being a fully bidirectional work made by the systematic reversal of entries, the *Vocabulario* is essentially monodirectional.¹⁶ It was succeeded by a smaller Tagalog dictionary compiled by Domingo de los Santos and published posthumously in 1703. A third, slightly larger than San Buenaventura's, was compiled by Juan de Noceda and Pedro de San Lúcar and published by San Lúcar in 1754, after his confrere's death. Noceda and San Lúcar's dictionary begins with a Tagalog–Spanish section of 619 double-column pages, followed by an appendix of 34 pages, and, finally, an index of Spanish entries with the corresponding root in Tagalog, of 190 double-column pages.

¹² The former remains in manuscript; the latter has been edited as Varo, *Vocabulario de la lengua mandarina* (2006).

¹³ Translated by W. S. Coblin, in Varo, *Vocabulario de la lengua mandarina* (2006), I.16.

¹⁴ Sueiro Justel, *Historia de la lingüística española en Filipinas*, 171; Fernández Rodríguez, 'Lexicography in the Philippines', 4.

¹⁵ The following overview of dictionaries of Philippine languages is from Fernández Rodríguez, 'Lexicography in the Philippines', 4–8 (entry counts at 6).

¹⁶ Wolff, 'The *Vocabulario de lengua tagala*', 44.

The first monodirectional Tagalog–Spanish dictionary was composed before 1620 by the Franciscan Francisco de San Antonio, also called ‘Orejita’; it remained in manuscript until modern times.¹⁷ One of the extant manuscripts presents 6,749 entries, and the other 5,058. This work is, according to its modern editor, a derivative of San Buenaventura’s Tagalog–Spanish section but, as well as presenting Tagalog roots, it also presents compound forms.¹⁸ Even if San Antonio’s work draws on San Buenaventura’s, however, it adds enough to count as a new creation. For example, San Buenaventura has five entries for the headword form *a*, and San Antonio has ten, with a great number of examples; likewise, San Buenaventura has twenty entries for *ca*, and in San Antonio we find thirty-nine.

As for other major languages of the Philippines, the first extant dictionary of languages of the Visayan family is the *Vocabulario de la lengua bisaya, hiligueyna, y haraya* of Alonso de Métrida, published in 1637, in which, unusually, the Visayan–Spanish part was longer than the Spanish–Visayan part, with about 18,000 and 10,900 entries respectively. The first dictionary of Bicol was compiled by Marcos de Lisboa, who died in 1628, but was published only in 1754; a dictionary of Ibanag compiled by José Bugarín, who died in 1676, was published only after successive revisions in 1854; the first of Pampango, compiled by Diego de Bergaño, was published in 1732. The first of Ilocano, ‘Calepino ylocano’, largely compiled by Pedro de Vivar, who died in 1771, was transmitted in manuscript, though an adaptation was published in the nineteenth century.¹⁹ All of these dictionaries are bilingual and bidirectional to and from Spanish, with the exception of the Ilocano–Spanish ‘Calepino ylocano’.

The first Europeans arrived in Japan in 1542 or 1543, and missionary activity began there under the leadership of St Francis Xavier in 1549. The study of the Japanese language by missionaries led almost at once to the compilation of grammars and dictionaries, for instance a ‘Vocabulario da lingua Japoneza’ by Duarte de Silva, but these have all apparently been lost.²⁰ Plans for a trilingual Latin–Portuguese–Japanese dictionary, *Dictionarium Latino Lusitanicum, ac Iaponicum*, began in 1581, a final draft was produced by three European and two Japanese lexicographers from 1593 onwards, and the dictionary was

¹⁷ San Antonio, *Vocabulario tagalo* (2000), is an edition.

¹⁸ A. Postma, in San Antonio, *Vocabulario tagalo* (2000), xii–xiii.

¹⁹ Fernández Rodríguez, ‘Lexicografía de la lengua ilocana’, is a detailed study; the edition which forms part of it is cited as Vivar, ‘Calepino ylocano’ (2012). Fernández Rodríguez, ‘El *Calepino ilocano* (c. 1797) del P. Vivar’, is more accessible in print.

²⁰ Zwartjes, *Portuguese Missionary Grammars*, 93–4.

finally printed in Amakusa in 1595.²¹ It was based on a bulky polyglot edition of Calepino (see below), and the text of its 32,000 entries was judiciously abridged so as to make the book portable in the mission field.²²

One remarkable work which calls for notice here as a dictionary printed, although not compiled, by missionaries, was published by the Jesuit press in Japan. 'This new dictionary, a campaigning work for the Jesuits in their religious controversies with the bonzes, and an instructive work for learning to read Buddhist texts set forth in *kanji*, sought to be, at other levels as well, useful for preaching' and for other purposes.²³ Called *Rakuyōshū* ('Collection of fallen leaves'), or, in the spelling and formulation of its title page, *Ra cu yo xu sive dictionarium iaponicum*, it was the first movable-type dictionary to incorporate two *kana* innovations which enabled the reading of Chinese characters (*kanji*) by Japanese people. It can be characterized as a 'dictionary of characters' and not as a 'dictionary of words': it is a tool for learning the equivalences between Chinese characters and spoken words rather than for learning the meaning of words.²⁴

The first book gives the two readings available for each *kanji* character in the syllabic *kana* system: on the right the Chinese reading, called *koe* (= *koe*, in modern Japanese *ondoku*, 'the pronunciations borrowed into Japanese from Chinese'), and on the left, the Japanese reading, called *yomi* (*kundoku*, 'the pronunciations of the Japanese words represented by the Chinese characters'; for *kana*, *ondoku*, and *kundoku*, see Chapter 10).²⁵ The Japanese syllabary has a conventional order, and the order of the *kanji* characters in this section depends on the order of the syllabic transcription of their *on*-readings. They appear with *koe* to their right and *yomi* to their left. The second part is a list titled *Irohajishū*, in which the *kanji* characters and their compounds are arranged in the order of the syllabic transcription of their *kun*-readings, with *yomi* to their right and *koe* to their left. At the end of the *Irohajishū*, a list of a hundred professions of Japanese government and administration and their Chinese equivalents is included, with an appended list of the Japanese provinces. The third part, the *Shōgokuhen*, is a list of the *kanji*, arranged in 12 thematic categories, such as astronomy, geography, and

²¹ Kishimoto, 'Adaptation of the European polyglot dictionary of Calepino', 210–11.

²² Kishimoto, 'Adaptation of the European polyglot dictionary of Calepino', 214–16.

²³ Debergh, 'Débuts des contacts linguistiques', 34, 'Ce nouveau dictionnaire, œuvre militante des Jésuites dans leurs controverses religieuses avec les bonzes, et œuvre de méthode pour apprendre à lire les textes bouddhiques rédigés en *kanji*, se voulait, à d'autres niveaux aussi, utile à la prédication [etc.]'

²⁴ Bailey, 'The *Rakuyōshū*', 300; Bailey, 'The *Rakuyōshū* II', 263–4.

²⁵ Bailey, 'The *Rakuyōshū*', 291; see also Debergh, 'Débuts des contacts linguistiques', 35.

human beings, which are subdivided in a total of 105 subcategories, such as 'sun', 'moon', 'earth', 'water', and 'fire'. Every *kanji* is accompanied by its *koe* and *yomi* written in *kana*. The *Rakuyōshū* did not have a marked influence on later Japanese lexicography (see Chapter 10).

The *Vocabulario da lingoa de Iapam com a declaração em Português* ('Vocabulary of the language of Japan with an explanation in Portuguese'), printed in Nagasaki in 1603 with a long supplement of 1604, was a monodirectional Japanese–Portuguese dictionary of 800 pages with 32,798 entries; its 'objectivity, thoroughness and wealth of cross references are quite remarkable for a work of that time', a point to which I shall return.²⁶ A Japanese–Spanish version, *Vocabulario de Iapon declarado primero en portugues por los padres de la compañía de Iesus de aquel reyno, y agora en Castellano en el Colegio de Santo Thomas de Manila* ('Vocabulary of Japan, first explained in Portuguese by the Jesuit fathers of that kingdom, and now in Spanish, in the College of Saint Thomas at Manila'), was printed in Manila in 1630.²⁷

The number of missionary dictionaries of Japanese is limited, as a consequence of the relatively short period of Christian activity in Japan.²⁸ One feature they share is that, with the exception of the *Rakuyōshū*, the Chinese and Japanese characters are never used in them; Japanese always appears in the Latin alphabet. As we have seen, Chinese dictionaries such as those of Ruggieri and Ricci and of Díaz used Chinese characters. Likewise, the Tamil dictionary of Antão de Proença uses Tamil script (see Chapter 30). Perhaps the preference of the missionary lexicographers in Japan for the Roman alphabet was related to their possession of a printing press with Roman types.

Finally, the first extant missionary dictionary of Vietnamese, the *Dictionarium Annamiticum, Lusitanicum, et Latinum* of Alexandre de Rhodes, published in 1651, starts with Vietnamese lemmata, followed by Portuguese and finally Latin (earlier Vietnamese–Portuguese–Latin and Portuguese–Vietnamese dictionaries have been lost).²⁹ The main trilingual section is 450 double-column pages long. At the end is an index of the Latin equivalents in the dictionary, 'Index latini sermonis', which probably had a similar function to that of the Tagalog–Spanish second section of San Buenaventura's dictionary; there is no index of the Portuguese equivalents.

²⁶ Cooper, 'The *Nippo Jisho*', 418; *Vocabulario da lingoa de Iapam* (1973) is one of several facsimile editions.

²⁷ *Vocabulario de Iapon declarado primero en portugues* (1972) is a facsimile edition.

²⁸ For context, see Maruyama, 'Linguistic studies by Portuguese Jesuits', 141–2.

²⁹ For de Rhodes' dictionary and its predecessors, see Phan, *Mission and Catechesis*, 32–5.

The Jesuit lexicographical production in Japan is the most impressive of these traditions regarding the total number of entries included in the *Dictionarium Latino Lusitanicum, ac Iaponicum* and the *Vocabulario da lingua de Iapam* (around 32,000 each), but it is without any doubt the Philippine tradition that has no comparison in any other part of the world, except in Mesoamerica, regarding the total amount of lexicographical works. According to Sueiro Justel, 108 works were compiled during the colonial period until 1898, of which 30 are believed to be lost and 17 are re-editions.³⁰

Directionality, Trilingualism, and Readership

When dictionaries are bidirectional, the two sections are seldom equal in length (the same is true of Mesoamerican dictionaries: see Chapter 27). The section that translates to Latin, Spanish, or Portuguese is not just another version of the first section, where the Asian equivalents of the European-language lemmata are rearranged alphabetically. So, for instance, San Buenaventura's *Vocabulario de la lengua tagala* contains 707 pages: 618 in the first, Spanish–Tagalog, section, and 89 in the second, Tagalog–Spanish, section. (The Tagalog dictionary of Domingo de los Santos had a similar structure.)

Although San Buenaventura states that his appendix was made in order 'to help the natives learn Castilian', it has been plausibly remarked that 'San Buenaventura did not accord teaching Spanish to the natives high priority.'³¹ Likewise, although the modern editor of San Antonio's Tagalog–Spanish dictionary has suggested that it was aimed at Tagalog-speakers who wanted to learn Spanish, it is doubtful that all the specific information in the dictionary was really intended to serve such a readership. In fact, it was unusual for missionary dictionaries to suggest that monodirectional dictionaries, or sections of bidirectional dictionaries, in which Spanish was the target language were meant for the use of speakers of the source language. A century before San Buenaventura, Pedro de Alcalá had, exceptionally, claimed that his Spanish–Arabic dictionary (in which all the Arabic content was presented in Roman transliteration) was written both for Spanish-speaking Old Christians and for Arabic-speaking converts, but this seems like an unrealistic ideal: the latter would have found it very hard to use.³²

³⁰ Sueiro Justel, *Historia de la lingüística española en Filipinas*.

³¹ San Buenaventura, *Vocabulario*, 'porque los naturales puedan de prender con mas facilidad a hablar la lengua Castellana'; Wolff, 'The *Vocabulario de lengua tagala*'.

³² Alcalá, *Vocabulista arauigo*, sig. azv, for which see the entry by O. Zwartjes in *Corpus de textes linguistiques fondamentaux*.

After San Buenaventura, however, neither the monodirectional dictionaries in which Spanish was a target language (such as Vivar's 'Calepino ylocano' or, in the New World, Ruiz de Montoya's Guarani–Spanish *Tesoro de la lengua Guarani*, for which see Chapter 26), nor bidirectional dictionaries such as Méntrida's *Vocabulario*, in which the Visayan–Spanish section is six times longer than its Spanish–Visayan counterpart, repeat his ideal of teaching the native population Spanish. By contrast, the title page of the *Dictionarium Latino Lusitanicum ac Iaponicum* states that it is not only for the use of Japanese learners of Latin, but also for that of European learners of Japanese.³³

Trilingual dictionaries were compiled in East Asia, as they were in Mesoamerica. Their order differs: it is remarkable that the *Dictionarium Latino Lusitanicum ac Iaponicum* of 1595 places Latin first, whereas de Rhodes' *Dictionarium Annamiticum, Lusitanum, et Latinum* of 1651 places it last. In Diego Collado's trilingual dictionary, Latin comes first, followed by Spanish and finally Japanese (no index of Japanese is appended). Collado's work on Japanese and de Rhodes' on Vietnamese were both printed by the Propaganda Fide Press, in 1632 and 1651 respectively. This means that there did not exist any official policy from the Vatican on what kind of dictionaries had to be composed – or, rather, it was their policy to be flexible towards the manuscripts submitted for print. It seems that there were some official guidelines from Spanish authorities, prescribing the model of Elio Antonio de Nebrija for teaching grammar in New Spain. On the other hand, it seems that no guidelines were used by any religious order regarding the external structure of the dictionaries and the internal structure of their entries.

As well as these fully trilingual dictionaries, some apparently bilingual dictionaries actually include material from three languages. As noted above, the Portuguese–Chinese dictionary of Ruggieri and Ricci contains not only Portuguese but also Italian entries. The Bodleian manuscript of the Chinese–Spanish dictionary of Francisco Díaz contains a number of equivalents in Portuguese rather than Spanish, and the Jagiellonian manuscript sometimes mixes Portuguese and Spanish in the same entry.³⁴ As we shall see, Varo's Portuguese–Chinese dictionary includes a number of Portuguese–French–Chinese entries. Such manuscripts circulated among missionaries from

³³ *Dictionarium Latino Lusitanicum ac Iaponicum*, title page, 'in vsum, et gratiam Iaponicae iuventutis, quae Latino idiomati operam nauat, nec non Europeoru[m], qui Iaponicu[m] sermonem addiscunt'.

³⁴ Kraków, Jagiellonian Library, Ms Berol. Ms Sin 13. Hisz. chin., XVII, s.v. *chú*, uses Portuguese 'coração' and Spanish 'coraçon' in the same sentence.

different nations, and manuscripts often contain sections written by different scribes.³⁵

Sources

Whereas missionary lexicographers of the Mesoamerican languages often invoked the name of Nebrija, and sometimes that of Ambrogio Calepino, as a source, their counterparts in East Asia usually did not identify sources. As Byron Ellsworth Hamann has shown in *The Translations of Nebrija*, the majority of early lexicographical works produced by Spanish-speaking missionaries in the Americas were modelled on one or other of the many editions of Nebrija's Spanish–Latin dictionary, but only one such dictionary produced in Asia, San Buenaventura's dictionary of Tagalog, appears in his study. The entries in this dictionary follow Nebrija *grosso modo*, although Hamann points out that, unlike Nebrija's characteristically terse entries, San Buenaventura's take up two or more lines of type, beginning with the Spanish lemma with a Tagalog translation, followed by entire phrases in Castilian, translated to Tagalog, showing each word's use in context.³⁶ As Hamann observes, very little work has been done with San Buenaventura's *Vocabulario*.³⁷ He concludes that San Buenaventura used John Minsheu's *Dictionarie in Spanish and English* (1599), a dictionary in the Nebrija tradition, partly because both San Buenaventura and Minsheu spell certain words with initial <z> (for instance *zebolla* instead of *cebolla*). This is in fact flimsy evidence, since these letters are often confused, as in Francisco Varo's dictionaries; likewise, in Noceda and San Lúcar's much later dictionary, we find spellings as *zebo*, *zebolla*, *zepillo*, and *zicatriz*.³⁸ Joaquín García-Medall comes to a different conclusion, namely that the main source of San Buenaventura was probably Alonso Urbano's trilingual Spanish–Nahuatl–Otomi dictionary (for which see Chapter 27) and that this was also the main model for Méndrida's dictionary of Visayan.³⁹

Other missionary dictionaries of East Asian languages may be derived from Nebrija, directly or through Mesoamerican intermediaries (more research is needed on the influence on East Asian lexicography of the Nahuatl dictionary of Molina, for which see Chapter 27).⁴⁰ W. South

³⁵ See, for instance, Zwartjes, *Portuguese Missionary Grammars*, 286.

³⁶ Hamann, *Translations of Nebrija*, 81. ³⁷ Hamann, *Translations of Nebrija*, 83.

³⁸ Noceda and San Lúcar, *Vocabulario de la lengua tagala*, 190 (second sequence of pagination).

³⁹ García-Medall, *Vocabularios hispano-asiáticos*, 39, 287–95.

⁴⁰ For the point about Molina, see the suggestive remarks in Klöter, *Language of the Sangleys*, 79–81.

Coblin concludes that the Spanish–Mandarin dictionary of Varo was based on Nebrija, although he does not argue his case in detail, or point to a specific edition of Nebrija.⁴¹ Since Varo compiled his Portuguese–Mandarin dictionary before his Spanish–Mandarin one, we would expect the original source of his wordlist to be Portuguese. No Portuguese translation of Nebrija appears to be extant so, if Varo did use Nebrija, he would have had to translate the wordlist of his source into Portuguese rather than just using the wordlist of Cardoso – unless he had access to a Portuguese version of Nebrija which is now lost. More research is needed in order to trace the models used by Varo, and to explain why many lemmata of the Portuguese–Mandarin dictionary are in fact in Spanish – there are even mixed lemmata such as *bueno, bem fizeste*, which begins in Spanish and ends in Portuguese – and why so many entries are trilingual Portuguese–French–Mandarin (whereas Varo’s Spanish–Mandarin dictionary does not have Spanish–French–Mandarin entries). Similarities between the headwords of Varo’s Spanish–Mandarin dictionary and those of one of the major dictionaries of Early Manila Hokkien would likewise repay investigation.

Calepino in dictionary titles can be used with a generic meaning, as a synonym for ‘dictionary’, but the lexicographic style of a great number of works composed in the Philippines definitely follows Calepino’s. It is probable that San Buenaventura not only used Nebrija but also had access to an edition of the *Dictionarium* of Calepino, since, as we shall see, he marks accentuation with abbreviations which also occur in Calepino – and not in Nebrija. San Buenaventura’s adaptation of Nebrija moved the dictionary more towards the Calepino style in other respects, for instance the addition of examples, texts, proverbs, locutions, and metaphorical usage, which are less common in dictionaries following the style of Nebrija closely.⁴²

Calepino’s dictionary was an important source for missionary lexicographers of Japanese. The seven-language Lyon edition of 1570 was probably the main source of the *Dictionarium Latino Lusitanicum ac Iaponicum* of 1595 (for the polyglot editions of Calepino, see Chapter 14).⁴³ The Portuguese material in this dictionary is translated directly from the Latin of Calepino.⁴⁴ The Japanese–Portuguese *Vocabulario da lingua de Iapam* also follows the model of

⁴¹ W. S. Coblin, in Varo, *Vocabulario de la lengua mandarina* (2006), 11–26.

⁴² See Fernández Rodríguez, ‘Lexicografía de la lengua ilocana’, 10–21.

⁴³ Kishimoto, ‘Adaptation of the European polyglot dictionary of Calepino’, 211–14; Kishimoto, ‘Process of translation’, 19–25.

⁴⁴ Messner, ‘Ist das *Dictionarium latino lusitanicum ac iaponicum* ein Wörterbuch der portugiesischen Sprache?’

Calepino in the way it represents the order of the words, the roots, and the forms of the verb.⁴⁵

Nebrija and Calepino were not the only European sources for the work of the missionary lexicographers; for instance, a significant part of the wordlist of Ruggieri and Ricci's *Dicionário Português–Chinês* was derived from an edition of the Portuguese–Latin dictionary of Jerónimo Cardoso.⁴⁶ Other sources were the works composed by other missionaries. We have seen that the *Vocabulario de Iapon* printed in Manila in 1630 is a Spanish translation of the earlier Portuguese *Vocabulario da lingua de Iapam*. Even when direct translation did not take place, we see that missionaries took one or more earlier dictionaries as their model, as in the case of the use of the Tagalog dictionary of Domingo de los Santos in that of Noceda and San Lúcar. As has been observed, Nebrija is usually not mentioned by name, but generally, in the prologues, we find a detailed list of names of the works made by more immediate predecessors. So, for instance, Domingo de los Santos observes that he follows the grammar of Andrés Verdugo in certain technical respects, and that his dictionary is based on the works of Pedro de San Buenaventura (mainly its Spanish–Tagalog part) and Francisco de San Antonio (Tagalog–Spanish).⁴⁷ San Lúcar states in his prologue to the *Vocabulario de la lengua tagala* that he is familiar with all the grammars of Tagalog, of which he has seen thirty-seven, and that more has been written on this language than on all the living and dead languages of Europe.⁴⁸ In this prologue we can find a large list of names of predecessors – but it is always necessary to compose new dictionaries, according to San Lúcar, mainly because many words are no longer in use, or sometimes because intensive study of the language in question has led to new insights. Another good reason to compose a new dictionary, he adds, is that the earlier ones were difficult to get, and the surviving copies were often corrupt or badly damaged.⁴⁹ Dictionary prologues might not only acknowledge the work of predecessors, but also adapt it; for instance, when Vivar tells his readers that he has omitted many names of 'trees, herbs, flowers, and other mechanical things' from his Ilocano dictionary, since naming them all would be a never-ending enterprise, his words

⁴⁵ Debergh, 'Débuts des contacts linguistiques', 36. ⁴⁶ Messner, 'Primeiro dicionário', 60–4.

⁴⁷ Santos, *Vocabulario de la lengua tagala*, sig. §3r.

⁴⁸ Noceda and San Lúcar, *Vocabulario de la lengua tagala*, sig. §1v, 'se hallan en esta lengua tantos Artes, que ella sola excede en el numero (siendo tan corta su extension) al numero de Artes, que avra para las lenguas vivas, y muertas de toda Europa. Treinta, y siete he leído.'

⁴⁹ Noceda and San Lúcar, *Vocabulario de la lengua tagala*, sig. §1r.

seem to be adapted from a passage in the prologue to the dictionary of Domingo de los Santos.⁵⁰

Sometimes, this sort of borrowing may have been the result of the extensive travels of the missionary lexicographers. In 1630, as we have seen, the Spanish *Vocabulario de Iapon* was published in Manila. In 1632, Francisco Díaz arrived in Manila; Diego Collado, who had been in Manila in 1611 and in Japan in 1619, returned to Manila in 1635 and stayed there until 1638; in 1641 Alexandre de Rhodes was in Macao and in Manila; in 1648, Francisco Varo arrived in Manila; and in 1651 Martino Martini, the author of a grammar of Chinese written in Latin, was there as well; between 1649 and 1651, both Martini and Varo were in mainland China. We do not know whether, for instance, Díaz and de Rhodes met when they were both in Manila, or whether Martini and Varo met when they were both in mainland China, but the opportunities were clearly there. There are obvious examples of influence: de Rhodes' Romanization of Vietnamese was influenced by earlier Romanizations by Jesuits of Mandarin Chinese, and the Romanization of the Dominicans working in Manila (and in mainland China) was inherited from the Jesuit Nicolas Trigault, particularly as it was developed in his *Xīrú Ērmùzī* of 1626 (for which see Chapter 15). Much work remains to be done in this area.

Missionary lexicography in East Asia could not always draw on European sources: Nebrija and Calepino provided no model for the handling of Chinese characters. Although we cannot exclude the possibility that the compilers of the *Rakuyōshū* had been inspired by Chinese dictionaries written by Europeans, some of its sources appear to have belonged to Japanese traditions.⁵¹ The 'cabecilla' style of dictionary, represented by the 'Vocabulario de letra China' of Francisco Díaz – and also, presumably, by a lost seventeenth-century Hokkien dictionary from the Philippines, the 'Cabecillas, o léxico del dialecto de Emuy, o del mandarín' of Magino Ventallol – was laid out quite unlike any European dictionary.⁵²

Markedness and Grammatical Information

Some of the linguistic information provided in the East Asian missionary dictionaries was redundant: in the *Dictionarium Latino Lusitanicum ac*

⁵⁰ Vivar, 'Calepino ylocano' (2012), 495, 'se han omitido, en este nuevo Calepino, muchos nombres de arboles yervas, flores, y otras cosas mecanicas porque seria nunca acabar'; cf. Santos, *Vocabulario de la lengua tagala*, sig. §4r, 'No DUDO, hallará el Ministro en algunos Pueblos Vocablos que no estan aqui, que no es facil, el recogerlos todos, y en particular, nombres de Arboles; Yervas, Aves, y Pescados, y aun algunas Alajas de casa.'

⁵¹ Bailey, 'The *Rakuyōshū*', 323.

⁵² For Ventallol's work, see Klöter, *Language of the Sangleys*, 55.

Iaponicum of 1595, every single entry contains the abbreviation 'Lus.' for the Portuguese equivalent of the Latin headword, followed by 'Iap.' for the Japanese, which seems superfluous, since the languages are clearly recognizable. If we consider the huge number of entries (32,000), the compilers could have saved a lot of space omitting these abbreviations.

However, much significant information might also be provided. The *Vocabulario da lingua de Iapam* is an example. Its treatment of verb forms goes beyond the provision of basic, present, and past forms to offer a number of examples.⁵³ It also provides detailed information about regional varieties, particularly 'Ximo' (Kyushu) and 'Cami' (Kyoto), but also varieties from other domains, such as 'Vouari' (Owari). Women's or children's words are also labelled as such. Words related to Japanese religious practices are labelled as 'B' or 'Bup.' (*Buppo* 'Buddhist'): there are more than fifty terms related to Buddhism and Shintoism under the letter A alone, and more than a hundred under B. Other codes are 'S' (*scriptura*) for literary words and 'P' (*poesia*) for poetic words. In the prologue, we read that in this dictionary, 'the most current *Cobita* words are included', and in the corresponding entry *Cobi, uru, ita* we find what this means: to be *Cobita* is 'to be extravagant in using new and delightful words' and *cobite yū* is 'to use these extraordinary words': the reference is to the use in speech and writing of metaphors and elegant expressions and manners of speech.⁵⁴ Homonyms are distinguished, for instance in separate entries for *fana* 'nose' and *fana* 'rose or flower', and in five entries for *faxi* with different meanings.

In the Spanish–Mandarin dictionary of Varo, markedness is likewise indicated. Although the dictionary mainly concentrates on the spoken language, some literary vocabulary is included. In one of the manuscripts of this dictionary, literary expressions are underlined, and in others they are accompanied by an explanatory note, such as 'para escritura' ('for use in writing'). Other expressions are marked as 'tosco' ('crude, coarse').

In the Chinese and Vietnamese dictionaries, grammatical markedness occurs sporadically. Apparently it was less relevant for the lexicographers to tell the readers to which part of speech a certain headword belongs: in inflected languages, this information helps to determine inflectional patterns,

⁵³ For grammatical information and labelling in this dictionary, see Maruyama, 'Linguistic studies by Portuguese Jesuits', 143–5; for labelling, see also Cooper, 'The *Nippo Jisho*', 418, 423.

⁵⁴ *Vocabulario da lingua de Iapam*, prologue, 'se põe muitas palauras Cobitas mais correntes', and s.v. *Cobi, uru, ita*, 'ser *Cobita*: estrauagante em falar palauras nouas & exquisitas ... *cobite yū*: Falar estas palauras extraordinarias'; see Fonseca, *Historiografia linguística portuguesa e missionária*, 103.

but in Chinese and Vietnamese there are no inflections. An exception is the labelling of the ill-defined category of the particle, which is quite frequent in the Vietnamese dictionary of Alexandre de Rhodes.⁵⁵

San Buenaventura's dictionary of Tagalog had quite an elaborate system of abbreviations to mark grammatical information, not only the parts of speech, but also other features: a novelty is the abbreviation *ff* meaning *facere facere* ('to cause to do') for causatives. Since other works were based on San Buenaventura, we find extended forms of his list of abbreviations in them, covering inflectional, derivational, aspectual, functional, and pragmatic information.⁵⁶

San Buenaventura, and hence his followers, also marked prosodic features. His abbreviations, which he used systematically throughout his dictionary, came from Calepino: (pp) for *penultima producta* (the penultimate syllable is long and is therefore accented) and (pc) for *penultima correpta* (the penultimate syllable is short and therefore the accent falls on the antepenultimate). He also marked compound forms to show whether they were made up of two or three elements.⁵⁷ Métrida set out a slightly more elaborate system in his Visayan dictionary, and as Vivar explained in the prologue to his 'Calepino ylocano', different analyses might distinguish as many as eight or as few as two patterns of accentuation.⁵⁸

San Buenaventura also marks loanwords, labelled with the letter C, standing for *castellano* – but for 'corrupt Castilian, pronounced by the Tagalans in a horrible way'.⁵⁹ And he also marks metaphors, a practice which is much more developed in Spanish and Portuguese grammars written in Asia than those written in the New World. So, again, do his successors: in the dictionary of Domingo de los Santos, the cross with which metaphorical usage is marked occurs almost on every page.

Diatopical varieties are noted in almost every dictionary. San Buenaventura is unusual in marking them systematically with abbreviations: 'M.' (Manila), 'L.' (Laguna), 'T.' (Tinguian, and even more precise, 'between Nacarlan and the coast'), and 'S.' (Silanga). Métrida's dictionary describes

⁵⁵ Rhodes, *Dictionarium Annamiticum*, cols. 53, 110, 122, 126, 139, 209, 219, 585, 601, 634, 658, 664, and 779.

⁵⁶ See, for instance, Santos, *Vocabulario de la lengua tagala*, sig. §3r.

⁵⁷ García-Medall, *Vocabularios hispano-asiáticos*, 101–2.

⁵⁸ Métrida, *Vocabulario de la lengua bisaya* (2004), 149; Vivar, 'Calepino ylocano' (2012), 495, 'Unos señalan ocho acentos, otros dan solo cinco, otros quatro, y otros solo dos, que son penultima producta y penultima correcta [sic].'

⁵⁹ San Buenaventura, *Vocabulario de la lengua tagala*, 'significa que el vocablo es castellano y que lean ya corrompido los Tagalos a su modo horruno'.

three Visayan varieties – one indicated only as ‘Bisaya’, Hiligaynon, and ‘Haraya’ (a variety spoken on Panay) – and Cebuano, marking the last of these with a cross. Vivar’s dictionary of Ilocano marks roots with a cross, and also marks diatopical varieties using abbreviations, such as ‘Am.’ (Amianan) and ‘Abag.’ (Abagatan).⁶⁰

Finally, if we look closely at de Rhodes’ dictionary, we see that he systematically points out material related to local non-Christian beliefs as those of the *ethnici*, sometimes using this term in addition to others, such as *annamitica* and *Tunchinense*. Although de Rhodes does not use a specific abbreviation for this type of information, I consider this to be a special category of markedness, which I would like to call diaethnical markedness, since it has the purpose of informing the reader that the meaning of the term in question is specifically local and culture-bound, and that it has to be avoided in a Christian context, unless it is used as part of a neologism coined to express a Christian concept.⁶¹ Similar diaethnical marking occurs in the references to *gentios* in missionary dictionaries of Japanese.

Derivation

As occurs in the dictionaries of Tarascan in New Spain (see Chapter 27), the root was a crucial concept for Philippine grammars and dictionaries. Derivation was seen in antiquity as the formation of one part of speech from another, for instance of nouns from verbs, but in the Philippines, missionary grammarians and lexicographers regarded all parts of speech as derivations from roots: this was a great step forward in the history of lexicography. San Buenaventura observes that he will explain both the isolate Tagalog roots (which can stand by themselves) as well as the ‘bound roots’ (‘*raíces atadas*’, a forerunner of the concept of ‘bound morpheme’). A missionary grammarian of Tagalog called Juan Oliver had developed a numbering system for verbal derivations, which was taken up by San Buenaventura, and by later grammarians and lexicographers (Oliver’s original work has been lost).⁶² Derivational information did not need to be given in both halves of a bidirectional dictionary: in the dictionary of Noceda and San Lúcar, for instance, *matar* ‘kill’ is translated as *patay* in the Spanish–Tagalog section, where no derivatives are given, but when we search for *patay* in the Tagalog–Spanish section we find two columns with a large

⁶⁰ Vivar, ‘Calepino ylocano’ (2012), 494.

⁶¹ For more details see Zwartjes, ‘Some remarks on missionary linguistic documentation’.

⁶² For more details, see García-Medall, *Vocabularios hispano-asidóticos*, 74–80.

amount of information related to derivations, accompanied with many examples, some of them from religious texts, such as the Ave Maria.

As in New Spain, where Juan de Córdova explains to the users of his Zapotec dictionary that it would not make sense to include every derivative form for each verb for reasons of space, missionary lexicographers in the Philippines pointed out that a learner who was familiar with the mechanisms – or ‘juegos’ (‘games’) as they are called by Noceda and San Lúcar – of derivation could construct any form from the elements they provided. Given the word for ‘fish’, for instance, words with senses such as ‘to sell fish’, ‘to go fishing’, and ‘to ask for fish’ could be constructed; the user able to play with these words will be able to play with any other.⁶³ It has recently been remarked that ‘the Tagalog root is capable of more than five hundred affixational forms, but the idea that any portion of them should be cited for a given root did not come to mind in the case of the *Vocabulario* [of San Buenaventura] for this is in fact not a dictionary in any modern sense of a lexicon that expounds a language’.⁶⁴

In the Jesuit Japanese dictionaries, and hence in the Spanish *Vocabulario de Iapon*, we find a detailed treatment of derived forms: a single headword may be followed by sixty different derived and compound forms, covering three full pages.⁶⁵ By contrast with the Philippine tradition initiated by San Buenaventura, bound morphemes are usually not included as separate headwords in the Japanese dictionaries, although there are important exceptions, such as the inclusion of *-wa* and *-ga*.

As a logical consequence of the typological features of Chinese and Vietnamese, derivation did not play any role in the dictionaries describing these languages.

Content

In the more encyclopedia-like missionary dictionaries from East Asia, we see not only Western concepts being translated into other languages, but also attempts to grasp non-Western concepts. At one end of a spectrum of responsiveness to the local, just as lexicographers in Mesoamerica often translated Nebrija’s entry *mezquita* ‘mosque’ without any critical consideration, we see something comparable in the dictionary of Francisco Varo, who also has an entry ‘*mesquita de moros, lý páy chù*’.

⁶³ Noceda and San Lúcar, *Vocabulario de la lengua tagala*, sig. §3r.

⁶⁴ Wolff, ‘The *Vocabulario de lengua tagala*’, 40.

⁶⁵ García-Medall, *Vocabularios hispano-asiáticos*, 121–2.

When Hamann analyses the work of San Buenaventura, he observes that the entries

typically begin with simple Castilian–Tagalog word pairs, but these are followed by whole phrases in Castilian translated into Tagalog, showing each word’s use in context (and illustrating variations in the base term’s meaning). The presence of these extensive Castilian phrases means that today San Buenaventura’s *Vocabulario* is as much a guide to the meaning of terms in early modern Castilian as it is a source of information on early modern Tagalog.⁶⁶

This is just one simplified image, and although it is valid for some entries, there are so many examples where the opposite occurs. As has been demonstrated in recent scholarship, San Buenaventura’s dictionary is a treasury of information regarding Tagalog ‘methods of agricultural production, hunting and fishing, textile preparation, construction, navigation’, and more.⁶⁷ Moreover, many of the examples of usage are not even translated into Spanish at all.⁶⁸ In the Philippine tradition, we find detailed information about the people, the culture, their language, and, in particular, specific pragmatic usage. Missionary lexicographers in the Philippines paid much attention to colloquial speech, adages, teachings, proverbs, riddles, reprehensible social conduct, insults, taboo words.⁶⁹ Missionaries were never free from ethnocentrism, but some of them documented culture, beliefs, and the ethno-sociological features of daily life with encyclopedic scope and great precision, and most missionary lexicographers working in the Philippines also paid attention to intercultural misunderstandings.⁷⁰

Whatever the influence of Nebrija on dictionary entries which began with a Spanish headword, no Western model could have inspired the entries which began with a headword in an East Asian language. In Díaz’s dictionary we frequently find Mandarin lemmata translated with a generic word, such as ‘una yerba’ (herb), ‘una ave’ (bird), ‘una ave grande’ (big bird), or ‘un pajaro’ (bird), precisely because the words in question did not have Spanish equivalents. Likewise, the *Vocabulario da lingoa de Iapam* includes several technical

⁶⁶ Hamann, *Translations of Nebrija*, 81.

⁶⁷ Scott, ‘Sixteenth-century Tagalog technology’; Wolff, ‘The *Vocabulario de lengua tagala*’, 47.

⁶⁸ See the entry *escalones*, reproduced in Fernández Rodríguez, ‘Lexicography in the Philippines’, 19.

⁶⁹ García-Medall, *Vocabularios hispano-asiáticos*, 180–92; Fernández Rodríguez, ‘Lexicografía de la lengua ilocana’, 312–64.

⁷⁰ García-Medall, *Vocabularios hispano-asiáticos*, 193.

terms from Buddhism and Japanese literature, for information if not for active use:

references to Buddhism and Shinto in the *Vocabulario* are remarkably free from sectarian bias for a book of those days. Qualifications, such as 'according to the fables of the gentiles', are occasionally added, but on the whole an objective standard is maintained. For example, f. 273: '*Vmaregauari* . . . To be born again into this world in a different form or substance, as the gentiles believe'.⁷¹

There is no doubt that Alexandre de Rhodes' dictionary is a work of a missionary, and we must always bear in mind that one of the most important aims of these works was evangelization. On the other hand, the dictionary is also a source of indigenous culture, as we see from the inclusion of the Vietnamese proverb, 'Life is a journey, death a return home', carefully labelled as 'from the books of the *ethnici*' (de Rhodes used this proverb elsewhere as part of an argument against belief in the transmigration of the soul, asking 'How can the soul return home if it continues to transmigrate?').⁷² Such inclusions make the missionary dictionaries even more valuable as ethno-historical documents.

These works are also important sources regarding sociolinguistics. In the native Japanese tradition, colloquial speech was regarded as deserving no serious attention at all, and serious study of it starts only in the nineteenth century. The main purpose of the *Vocabulario da lingoa de Iapam* was much more descriptive than prescriptive and focused on the spoken language, but there is also attention to the literary and poetic styles, and there are hundreds of examples from literary texts. This does not mean that the editors recorded the several varieties without any prejudice, since they often labelled entries in terms of 'good' and 'bad', 'presumably on the grounds of elegance and politeness'.⁷³

According to Ricci's view, all missionaries in China should speak and write the language of the court known as Mandarin (*guānhuà*: see Chapter 15). This language 'is used in audiences and tribunals', and, 'if one learns this, he can use it in all provinces; in addition, even the children and women know enough of it to be able to communicate with all the people of another

⁷¹ Cooper, 'The *Nippo Jisho*', 424 n. 43; see also Debergh, 'Débuts des contacts linguistiques', 34.

⁷² Rhodes, *Dictionary Annamiticum*, col. 687, '*vita est transitus, mors est reuersio ad propriam habitationem. prouerbium desumptum ex Ethnicorum libris*'; cf. Phan, *Mission and Catechesis*, 85.

⁷³ Cooper, 'The *Nippo Jisho*', 418–19.

province'.⁷⁴ Ruggieri did not attempt to describe one of the regional vernaculars. Ricci wrote in 1582 that there are many different languages in China, but also that 'they have still another language which is, as it were, universal and common. And this is the language used by the Mandarins and the court, and to them it is like Latin to us.'⁷⁵

⁷⁴ Cited by W. S. Coblin in Varo, *Vocabulario de la lengua mandarina* (2006), 539.

⁷⁵ Cited from Schreyer, *European Discovery of Chinese*, 9.

European Traditions in India and Indonesia

TOON VAN HAL

‘Friendly reader and new missionary’, thus Manoel da Assumpçam addressed his readership in 1743, ‘I take it that you have come to Bengal with an Apostolic spirit and Apostolic charity, and with the ambition to convert the whole world to the law of JESUS Christ’. For the purpose of bringing ‘lost sheep to the bosom of the Church’ it was of great importance – so the Augustinian impressed in Portuguese upon his reader – to learn the Bengali language, ‘com fundamento’. It was for this reason that Assumpçam had written this work, in which he explained ‘the rules of the grammar’ and provided a ‘vocabulary in two parts, the first from Bengali to Portuguese, the second from Portuguese to Bengali’. He assured his reader that he would find in his work ‘all, or any rate the greater part, of the words used by the natives’.¹

It is precisely this combination of grammar and dictionary which makes Assumpçam’s work representative of the missionary linguistic works which were so abundantly produced in the early modern period. However, it ceases to be representative when one considers that the large majority of missionary wordlists, particularly those from India, were never published. Compared to grammars and religious texts translated into the native language, the place of printed wordlists, which were typically more voluminous, among the printed books of the missionaries was always relatively modest. But in comparison to

I would like to express my appreciation to Tensing Rodriguez and Christophe Vielle for sharing documentation. Reuben J. Pitts translated a first Dutch version of this text into English. Special thanks are due to the editor of this book. This research was made possible through a research project allocated by the KU Leuven Research Council and a short-running project by the Volkswagen Foundation.

¹ Assumpçam, *Vocabulário em idioma Bengalla e Portuguez*, sig. §5r, ‘Amigo leitor, e Missionario novo; como supponho que vens a Bengalla com espirito, e caridade Apostolica, e com zello de converter todo o mundo á Ley de JESUS Christo; e de trazer ao gremio da Igreja as ovelhas desgarradas do seu rebanho; e que para este fim dezejas aprender o Idioma Bengalla, com fundamento; aqui te offereço esta obra, em que acharas as regras da Grammatica deste Idioma; e hum Vocabulário em duas partes; a primeira de lingua [sic] Bengalla, e Portugueza; e a segunda de lingoa Portugueza; e Bengalla; em que acharas, senão todos ao menos a mayor parte dos Vocabulos de que uzaão os naturais.’

New Spain, where five wordlists were published in the sixteenth century alone (see Chapter 27), the balance for India is very meagre indeed.

This lack of printed sources makes it very hard to study the missionary lexicographical output in India. For the early modern missionary lexicography of New Spain, one of the greatest difficulties, according to Thomas Smith-Stark, is 'that of having a clear idea of what was actually produced since the bibliography on the subject tends to be incomplete, confusing and full of errors and inconsistencies'.² Smith-Stark's observation also holds true of India (and even more so). For instance, a number of older studies refer to a printed lexicographical work on Sinhala from 1759, a copy of which S. G. Perera never managed to locate.³ Conversely, it is likely that there are a number of archives and libraries in which currently unknown manuscripts await discovery. Furthermore, the preservation of a mere title is not always sufficient basis to determine the language which the work in question describes. The Indo-European language Konkani, for instance, was referred to by terms such as *concana*, *concanica*, *brâmana-goana*, *goana*, *canarim*, and *canarina*.⁴ However, the designations *canarim* and *canarina* were sometimes also used to refer to Kannada, a Dravidian language. In manuscript catalogues this inevitably leads to confusion, and often the issue can be resolved only by an inspection of the manuscript itself. Only a minority of the preserved manuscripts contain a preface or another kind of front matter. The digitization of the relevant source material is as yet in its infancy, while most modern facsimiles are very hard to obtain.⁵ Moreover, only a very small fraction of these wordlists have so far been studied in depth, the most thorough study being Gregory James' analysis of the first printed Tamil lexicon.⁶

The relationship between the missionary lexicographers and the people whose languages they documented was never simple. Their work provides important early documentation of some languages: for instance, as we shall see, the earliest records of colloquial Telugu are by missionaries.⁷ On the other hand, although the European impact on other cultures, from which the activity of the missionaries cannot be separated, was less destructive in South Asia than in the Americas, it may have contributed to the disappearance of a number of languages in this area as well.⁸ Furthermore, European agents, both missionaries and merchants, were primarily interested in the large, widely used

² Smith-Stark, 'Lexicography in New Spain', 21.

³ Perera, 'Some ancient grammars and dictionaries', 327.

⁴ Fonseca, 'Fontes manuscritas e impressas', 233.

⁵ *Galáxia das línguas na época da expansão*, 142. ⁶ James, 'Aspects of the structure'.

⁷ Colas, 'Cultural encounter', 78. ⁸ Collins, 'Language death in Maluku'.

languages, such as Malay, although the author of the first printed Malay wordlist was aware that ‘there are many additional indigenous languages’.⁹

Contact with European languages affected Indian languages and Indian ideas about language, introducing loanwords, influencing systems of spelling, affecting standardization, and bringing new perspectives on language.¹⁰ European perspectives on language and culture were, conversely, influenced by existing native traditions: it is too reductionist to study all missionary activities exclusively in terms of the imposition of European power on indigenous populations.¹¹ Contacts between divergent scholarly traditions may, as we shall see, create a fertile ground for the emergence of novel (‘hybrid’) knowledge.¹²

The first part of this chapter offers a broad overview of South Asian missionary lexicographical output: the known wordlists, the languages they cover, the lexicographers and the orders to which they belonged. The second part turns to the way in which the missionaries used the native lexicographical traditions. The third examines some of the choices made by the compilers of wordlists in terms of content and form. The last considers the fate of these early modern dictionaries. The lexicography of Persian, which was the official language in the Mughal empire (see Chapters 17 and 18), remains in the background in this chapter, since it was, together with the Semitic languages, regarded as one of the *linguae orientales*, which were studied at European universities from the beginning of the seventeenth century as part of biblical scholarship.¹³

Overview

Before the period of early modern interaction between Europe and South Asia, there had been some European settlement on the coasts of South India, which had not led to the making of wordlists, and the region of Indonesia had been reached by merchants from both the Middle East and China, a result of

⁹ Houtman, *Spraecck ende Woord-Boeck*, sig. ‘.2v, ‘noch vele eygene inlantsche spraken zijn’.

¹⁰ For spelling, see Chevillard, ‘Challenge of bi-directional translation’, 115; Gnanapragasam, ‘Contributions of Fr Beschi’, 349; for standardization, see James, *Colporul*, 114; for new perspectives, see Blackburn, *Print, Folklore, and Nationalism*, 26–7.

¹¹ See, e.g., Fountain, ‘Transculturation, assimilation, and appropriation’.

¹² Raj, *Relocating Modern Science*, 1–3; Winterbottom, *Hybrid Knowledge*, 5–6; Trautmann, *Madras School of Orientalism*.

¹³ By contrast, as the Orientalist August Pfeiffer remarked (*Critica sacra*, 447), languages such as Japanese, Vietnamese, Kikongo, and Malay were of no theological value other than their use in the mission field.

which had been the creation of a fifteenth-century Malay wordlist in Chinese characters.¹⁴ However, the establishment of a permanent settlement in Cochi by the Portuguese in 1500 heralded the beginning of the interaction under discussion here. On the western coast of India, Goa in particular became a crucial centre for the Portuguese, who also infiltrated coastal areas of Ceylon. The Dutch established smaller trading posts, spread along the entire Indian coastline. In the middle of the seventeenth century they would gain power primarily in Ceylon and in Indonesia. Dutch setbacks in India would make room for the British, who had established trading posts on the Indian east coast, for instance at Madras in 1639, and would eventually annex all of India. Some forty years later the French settled south of Madras, in Pondicherry. Further south still, in Tranquebar, the Danish built a fortress. So, the missions which were established in India were undertaken by various orders and backed by various states. Many missionaries, predominantly Franciscans, Dominicans, and Jesuits, in succession, were active under the protection of the Portuguese *Padroado real*. But in addition the so-called Carnatic mission, initiated in 1695 by the French king Louis XIV, played a significant role in Pondicherry and Chandernagore. The Protestants were not inactive either: in Ceylon the Catholic religion was forbidden by Dutch law from the middle of the seventeenth century. From the beginning of the eighteenth century onwards, the Danes in Tranquebar accommodated the first-ever Lutheran mission, co-ordinated from the pietistic orphanage of August Hermann Francke in Halle.¹⁵ The intercontinental movement of persons in the early modern period depended to a considerable extent on oceans, seas, and rivers: for this reason the heart of the Russian empire was a less feasible target for evangelistic or scientific fieldwork than the coastlines of Brazil, the Philippines, and India.¹⁶ As a result, by far the larger portion of missionary dictionaries from India documented languages spoken in the coastal regions. Nonetheless the Jesuit Roberto di Nobili did establish a mission from Cochi in Madurai, much further inland, 'where, as rumour had it, the local kings lived in a manner similar to ancient Mediterranean pagan despots, where learned men possessed a language as precise as Latin and where priests jealously preserved hidden sacred texts'.¹⁷ The case of

¹⁴ For early Europeans in India, see Filliozat, 'Deux cents ans d'indianisme', 87–9; for Middle Eastern and Chinese contact with Indonesia, see Winterbottom, *Hybrid Knowledge*, 55–6; for the Malay wordlist, see Echols, 'Presidential address: dictionaries and dictionary making', 14.

¹⁵ Gross, Kumaradoss, and Liebau, *Halle and the Beginning of Protestant Christianity in India*.

¹⁶ See Naarden, 'Witsen's studies of Inner Eurasia'.

¹⁷ Županov, *Disputed Mission*, 3. See also Blackburn, *Print, Folklore, and Nationalism*, 41.

Nobili shows that the missionaries gradually became aware that their success depended not only on mastery of local vernaculars but also on knowledge of the native cultures and languages of learning.

For the period 1500–1800 we currently know of about 150 different early modern wordlists describing an Indian or Indonesian language.¹⁸ As the case of Assumpçam has already indicated, many of these wordlists were supplemented by a grammar or accompanying religious texts.¹⁹ Only a fraction of these wordlists were published in printed form and, of those which remained in manuscript, around half are lost or unlocated. The extant manuscripts are for the most part in the keeping of Portuguese libraries (Lisbon, Braga, Evora, Coimbra) or in Rome, Paris, and London. In India, where Panaji (Goa) is the most important repository, it is likely that a number of interesting archives, such as the Jesuit Archives of the Madurai Province, have yet to be made available.²⁰ Some works which appear to have been lost may turn out to match one of the many anonymous works, which comprise fully one-third of the corpus. The authorship of missionary grammars and lexica is often hard to determine or ascertain, since many wordlists were not attributed to a certain author until later, and it is not always easy to judge how well substantiated such attributions are.²¹ Furthermore, authorship is by definition a relative concept. Missionaries made eager use of the work of their predecessors, and as a rule conceived of their lexicographical labours as accumulative contributions to a higher, collective project, rather than as strictly delineated intellectual property.²² Even so, in a number of cases it is known that conscious forms of plagiarism were involved: the conceited Calvinist pastor François Valentijn, for instance, ‘pirated’ the unpublished Malay dictionary of Georg Everhard Rumphius as part of ‘a most thorough plunder’ of the older man’s work.²³

Among the makers of non-anonymous wordlists the Jesuits emerge as the best-represented order. The outstanding diligence of the Jesuits in learning languages and creating linguistic tools is often emphasized in the literature.²⁴

¹⁸ For the project on which the following overview is based, see Van Hal, Peetermans, and Van Loon, ‘Presentation of the RELICTA database’.

¹⁹ Fonseca, ‘Historiographie linguistique portugaise’, 188.

²⁰ For Panaji, see Fonseca, ‘Fontes manuscritas e impressas’; for difficulties of access, see Chevillard, ‘Challenge of bi-directional translation’, 114.

²¹ Fonseca, ‘Fontes manuscritas e impressas’, 91.

²² Smith-Stark, ‘Lexicography in New Spain’, 30–1.

²³ For more or less wilful failure to acknowledge predecessors, see Xavier and Županov, *Catholic Orientalism*, 308–9; for Valentijn and Rumphius, see Beekman, *Fugitive Dreams*, 62.

²⁴ Fonseca, *Historiografia linguística portuguesa e missionária*, 81; Xavier and Županov, *Catholic Orientalism*, 213.

It must not be forgotten, however, that this order was in the first place a well-organized machine, which assured it an inordinate degree of attention in modern historiography, while others remained underrepresented to the point of invisibility. In some recent publications, for instance, the voluminous intellectual activity of the Franciscans has been brought into the spotlight, although many of its products have been lost: we know of almost fifteen wordlists of Indian languages made by the Franciscans and Capuchins, of which so far only a single Hindustani dictionary, that of François-Marie de Tours, has been located.²⁵ The order of the Carmelites was very important for South Indian lexicography as well. The lexicographical activity of the Lutherans, who were active only in Tranquebar, was also notable as was, from the second half of the eighteenth century onwards, that of the British. In general, production increases exponentially over the centuries: a large majority of wordlists date from the eighteenth, and primarily the late eighteenth, century.

As in the Americas, early modern proto-colonists noticed that in southern Asia a range of languages from different language families were spoken (principally Dravidian in the south, Indo-Aryan in the north and in Ceylon, Sino-Tibetan in Tibet, and Malayo-Polynesian in Indonesia). There were, however, three crucial differences. First, the inhabitants of the Indian sub-continent shared a framework of religious belief and practice, and Sanskrit functioned as a common scholarly language. It took some time before the Europeans realized this, since the language of the Brahmins was referred to by different names depending on the location and was recorded with different writing systems.²⁶ Secondly, in contrast to the Americas, literacy was well established in South Asia, although the intellectual culture could differ widely from place to place: in Bengal there was hardly a single book to be found, while in Tamil Nadu the poetic and literary tradition was of extreme importance.²⁷ The missionaries were confronted not only with unfamiliar writing systems and unfamiliar kinds of manuscript (notably palm leaves or *olai*), but also with unfamiliar traditions of lexicography (for these, see Chapters 4 and 7).²⁸ Thirdly, the presence in India of the followers of the Apostle Thomas, whom tradition alleged to have journeyed to India, meant

²⁵ Fonseca, *Historiografia linguística portuguesa e missionária*, 97–8.

²⁶ Van Hal, 'Protestant pioneers in Sanskrit studies', 120–1.

²⁷ Khan, 'Early history of Bengali printing', 53; Chevillard, 'Challenge of bi-directional translation', 124.

²⁸ For missionaries' encounters with palm-leaf manuscripts, see Colas, 'Cultural encounter', 73–4.

that Christianity was well represented in south-west India even before the arrival of Europeans. This meant that Syriac, the liturgical language of these Christians, was yet another language for the missionaries to consider.

Which of the native languages in India – sometimes characterized by the missionaries as ‘Indian’, but also frequently as ‘exotic’ or ‘barbarian’, a term which, incidentally, did not always imply disparagement – became the most frequent object of lexicographical activity?²⁹ Tamil emerges as the best-represented language in India, with more than a third of all preserved sources. Thanks to the work of Gregory James in particular, the early modern lexicography of Tamil has already been extensively documented.³⁰ Although Tamil is spoken on the east coast of India and the first Portuguese fortresses were, as we have seen, established on the west coast, St Francis Xavier achieved a great missionary success when a group of Tamil-speaking Parava fishermen decided to make an appeal for Portuguese support after conflicts with local Muslim groups.³¹ Not all Tamil wordlists documented the same form of the language: like other languages in the region, such as Malayalam and Malay, this language displayed not only diatopical, but also diastratic differences. ‘The Tamil language’, as Constanzo Beschi would put it in the preface to his dictionary of Common Tamil in the eighteenth century,

has a double dialect; the one is called *koduntamil* (common Tamil), the other *sentamil* (higher Tamil). They differ among themselves as much as Latin is different from, say Portuguese. Now although there are many common words in Latin and Portuguese, no one has so far dreamt on that account of giving the words of these languages mixed up in one vocabulary.³²

Apart from this well-known diglossic distinction, there were further differences dependent on caste and textual genre.³³ Early missionary studies of Tamil are lost; in 1679, the first extant dictionary of the language for missionary use, that of Antão de Proença, was published. Beschi (for whom see further below) created a number of dictionaries, widely circulated in manuscript form, of both high and spoken Tamil, in French, Latin, and Portuguese. Of the non-Jesuit Tamil experts, the most prominent was Bartholomäus Ziegenbalg, the first Lutheran missionary in Tranquebar.³⁴

²⁹ For the characterizations of Indian languages, see Fonseca, *Historiografia linguística portuguesa e missionária*, 80–1; Muru, ‘Shaping minds and cultures’, 207.

³⁰ See especially James, *Colporul*. ³¹ Blackburn, *Print, Folklore, and Nationalism*, 31–2.

³² Quoted in Gnanapragasam, ‘Contributions of Fr Beschi’, 348.

³³ Muru, ‘Shaping minds and cultures’, 206.

³⁴ Gnanapragasam, ‘Contributions of Fr Beschi’, 383–7.

Beside Tamil, the Indo-European Sinhalese was spoken on the heavily contested Ceylon (Sri Lanka). The main aim of the lexicographical production of the Calvinists in Ceylon seems to have been to provide a counter-offensive against its Catholic and Portuguese predecessors.³⁵ A letter of 1697 directed to the 'Lords Seventeen' of the Vereenigde Oostindische Compagnie (VOC; United East India Company) emphasized that Simon Cat had proven truly indispensable to Protestant missionary activity in Ceylon.³⁶ Cat's wordlist, recently published for the first time, may be the only extant Sinhalese wordlist of the Dutch period.³⁷ Some Portuguese cherished dreams of a Sri Lankan *reconquista*, and a small order – the Congregation of the Oratory of the Cross of Miracles, consisting of Brahmins who had converted to Catholicism – matched their words with deeds. These missionaries managed to keep their activities out of sight of the Dutch, because their 'color, language, customs and temperament are similar to those of the people they deal with'.³⁸ The only preserved wordlist of this order appears to be Jacome Gonçalves' *Vocabulario lusitano tamulico e chingalatico*. The majority of these authors had been raised in Goa, where Konkani was spoken. Many wordlists of this language were composed, of which only a minority have been preserved.³⁹ The number of extant Malayalam wordlists is higher, one of their leading compilers being Johann Ernst Hanxleden, who is still recalled in today's Kerala as *Arnos Padiri* ('Father Ernest'). None of his numerous dictionaries describing both common and Sanskritized Malayalam were published in the early modern period.⁴⁰ A smaller number of Telugu wordlists, mainly the work of anonymous Jesuits, survive, now concentrated in Paris; they are the first records of colloquial Telugu, which had not been of interest to lexicographers in the learned native tradition.⁴¹ Wordlists of both Hindustani and Bengali were primarily made by the British from the second half of the eighteenth century onwards (see Chapter 18), although there were some notable precedents in Portuguese and Latin, among them the Bengali dictionary of Manoel da Assumpçam. None of the eighteenth-century missionary wordlists of Nepali, Tibetan, and Burmese seem to have been preserved.

³⁵ Palm, 'Account of the Dutch Church in Ceylon', 33–4.

³⁶ Peiris, 'Sinhalese Christian Literature', 20.

³⁷ Paranavitana, *Dutch Sinhala Dictionary*, 22–3.

³⁸ Županov, 'Goan Brahmins in the Land of Promise', 178.

³⁹ Rivara, *Historical Essay on the Konkani Language*.

⁴⁰ See Vielle, 'Devotional Christianity and pre-Indology'; Vielle, 'New manuscript'.

⁴¹ Colas and Colas-Chauhan, *Manuscripts telugu*; Colas, 'Cultural encounter', 75, 78.

In Indonesia, Malay was the most important *lingua franca*.⁴² In comparison to the languages spoken in India, a veritable flood of printed wordlists for Malay is available to us, with a remarkable peak in the eighteenth century.⁴³ The driving force behind this production were the Dutch. The first printed Dutch–Malay dictionary (1603) came from the pen of Frederik de Houtman, who was not active as a missionary, but had learned the language during a period of imprisonment. A bidirectional wordlist by Caspar Wiltens, which made its way to a Dutch printing press in 1623, was the basis for a large number of strongly augmented and improved editions throughout the seventeenth century. A Latin translation by David Haex appeared in 1631 from the presses, remarkably enough, of the recently founded Catholic *congregatio de propaganda fide* – a papal initiative that would often clash with the Portuguese *Padroado real*. At the beginning of the next century, Andries Lambert Loderus published a compilation of most of the wordlists mentioned previously (*Maleische Woord-Boek Sameling*, 1707–8), in doing which he provided a striking illustration of just how efficient the circulation of lexicographical knowledge of Malay was. Meanwhile, in 1701, Thomas Bowrey had published the first bidirectional wordlist of English and Malay.⁴⁴ To this highly selective list should be added texts in which the listing of words was incidental to another purpose, such as encyclopedically oriented works in which fauna and flora are catalogued, sometimes mentioning indigenous names, such as Hendrik van Rheede’s botanical *Hortus Malabaricus* of 1678.⁴⁵

Although the Portuguese *Padroado real* was eventually forced to make way for other European powers, Portuguese retained, in comparison to other European languages, a crucial role in India. In various Asiatic trading posts, both Portuguese and Portuguese-based creoles were spoken, and most missionaries – whatever their original nationality – spoke the language very well.⁴⁶ From this perspective it is not surprising that Portuguese was the most prominent European language in the lexicographical output of India until British rule.⁴⁷ French missionaries usually used their own mother language. For Malay, and to a lesser extent Sinhalese, the metalanguage tended to be Dutch.

⁴² Winterbottom, *Hybrid Knowledge*, 56.

⁴³ Echols, ‘Presidential address: dictionaries and dictionary making’, 15.

⁴⁴ Winterbottom, *Hybrid Knowledge*, 63–5.

⁴⁵ See also the discussion of Rumphius in Leuker, ‘Knowledge transfer and cultural appropriation’.

⁴⁶ Schouten, ‘Malay and Portuguese’, 351–3; Xavier and Županov, *Catholic Orientalism*, 210–11.

⁴⁷ Fonseca, *Historiografia linguística portuguesa e missionária*; Assunção, ‘Portuguese missionary work’; Zwartjes, *Portuguese Missionary Grammars*, 271–302.

The Missionary Treatment of Indigenous Lexicographical Traditions

Most of the dictionaries referenced above are European in structure: semasiological, alphabetically ordered, interlingual. Their makers were introducing an entirely new tradition to India, contrasting with works such as the ancient onomasiologically oriented metrical dictionary of Sanskrit called *Amarakośa*, which enjoyed immense authority, and the dictionaries of Tamil written for the use of poets (for all of which see Chapter 7). Sanskrit–Tamil dictionaries did not exist, a consequence of the fact that there was no tradition of translation – only a tendency to assimilate.⁴⁸ *Amarakośa* was regarded with a degree of scorn by certain British scholars in the nineteenth century.⁴⁹ However, there were early missionaries who acquainted themselves with these native lexicographical traditions and treated them with respect.

This leads us to the topic of the controversies about Jesuit methods of accommodation. Throughout the early modern period, missionaries made energetic and often independent attempts to achieve a good knowledge of the ‘holy language’ of the Brahmans, as they were convinced that it was only through the conversion of the highest classes that it would be possible to convert the lower castes. There are various early modern testimonies that the Brahmans wished to keep this language from outsiders at all costs, although there may well have been differences depending on the period and the location.⁵⁰ Nobili’s exposure to the Vedic hymns in Madurai strengthened his conviction that the divine monotheistic doctrine had already been revealed to the native Tamil population in the distant past. He believed that the best strategy to convert the Tamil population to the true faith was to straighten out their current, distorted faith from the inside. This required the missionary to achieve a deep intimacy with the native culture: not only did Nobili familiarize himself with the native texts – there are indications that he was familiar with the native lexicographical tradition – but he also adopted the local mode of dress and became a vegetarian.⁵¹ The Jesuits adopted strategies of this nature not only in India, but also in Japan and China. For many non-Jesuits these forms of accommodation went much too far, leading to a heated discussion on the desirability of accommodation strategies. The

⁴⁸ Blackburn, *Print, Folklore, and Nationalism*, 28. ⁴⁹ James, *Colporuḷ*, 57–8.

⁵⁰ Van Hal, ‘Protestant pioneers in Sanskrit studies’, 112–13; Filliozat, ‘Deux cents ans d’indianisme’, 100. See also Xavier and Županov, *Catholic Orientalism*, 224.

⁵¹ For Nobili and the native lexicographical tradition, see Županov, ‘Aristocratic analogies’, 128–9.

discussion related to India, which has until now been less well studied than its Chinese and Japanese counterparts, would go down in history as the Malabar rites controversy.⁵²

An important later representative of the accommodation tendency was the Italian Jesuit Constanzo Beschi (for whom see also Chapter 7), who, as has just been mentioned, produced no fewer than four separate interlingual Tamil dictionaries, of which many manuscripts circulated in India. He completed his first dictionary in 1732, after about twenty years of work in the Madurai mission, a monolingual *Thesaurus linguae Tamulicae ad plenioram planioremque scriptorum Tamulensium intelligentiam . . . quatuor in partes* ('Thesaurus of the Tamil language for a fuller explanation and understanding of the writings of the Tamils, arranged in four parts'), or – with a scholarly native appellation – *Caturakarāṭi* ('Four-part dictionary').⁵³ Through his work Beschi aimed to improve the traditional *nikaṇṭu* dictionaries from the inside and to optimize them for a literary public. He did this by honouring the quadruple *nikaṇṭu* structure while implementing European principles such as full alphabetization and the use of prose within this structure, and also by increasing the portion of ordinary vocabulary considerably; his work has rightly been called 'monolingual but bicultural'.⁵⁴ Beschi's influence on the Tamil intellectual landscape was lasting, particularly after the eighteenth century.

Beschi's direct pietistic competitor, Bartholomäus Ziegenbalg, had two native dictionaries in his personal collection, the *Tivākaram* and the *Cūṭāmaṇi Nikaṇṭu*. Thanks to his catalogue, accompanied by a commentary, which is unfortunately all that remains of Ziegenbalg's extensive personal collection, we know that he wrongly assumed *Tivākaram* was written in prose. 'This book', he added, 'is studied only by those who wish to become scholars, or who interact with scholars and wish to understand their language. The common Malabarians understand not a word of it, or at least very little.'⁵⁵ Ziegenbalg made good use of the *copia verborum* of these works in the composition of his own dictionary of literary Tamil and in his researches into the Hindu pantheon.⁵⁶ An eighteenth-century catalogue of the mission

⁵² Županov, *Disputed Mission*, 3–5; Aranha, 'Vulgaris seu universalis', 333.

⁵³ James, *Colporūl*, 111; see also Gnanapragasam, 'Contributions of Fr Beschi', 347–58.

⁵⁴ Blackburn, *Print, Folklore, and Nationalism*, 59–60 (quotation at 60); James, *Colporūl*, 111–14.

⁵⁵ Ziegenbalg, *Bibliotheca Malabarica*, 50, 'Dieses Buch lernen allein diejenigen, so da wollen Gelehrte werden, oder doch solche Leute seyn, die mit Gelehrten umgehen und ihre gelehrte Sprache verstehen wollen. Die gemeinen Malabaren verstehen kein Wort aus selbigen oder doch ganz wenig.'

⁵⁶ W. Sweetman in Ziegenbalg, *Bibliotheca Malabarica*, 11, 51; his interest in the Hindu pantheon resulted in a celebrated work: Ziegenbalg, *Genealogie*.

library at Tranquebar lists copies both of the *Amarakośa* and of the Tamil dictionary *Akarāṭi Nikanṭu*.⁵⁷ A manuscript in the Vatican library bears the title *Vocabularium Sanscreticum, seu linguae Brachmanum, dictum Namalingam noesjasjanam* ('Dictionary of Sanskrit or of the language of the Brahmans, titled *Namalingam noesjasjanam*'), repeated at the end of the work, with the following clarification in Latin: 'Sjasjanam means book, Nama is name, Lingam kind. For in this work the words are distinguished by class.' The work is anonymous, but since it belonged to the Dutchman Hadrianus Relandus, it is reasonable to assume that he acquired it via his VOC connections. Both the anonymous redactor and Relandus are clearly impressed by the more than forty synonyms for 'sun' and more than twenty for 'moon'.⁵⁸

The Jesuits' interest in the native tradition was palpable as early as the seventeenth century. The Jesuit Heinrich Roth, stationed in North India, intended to compose a Sanskrit–Latin dictionary based on *Veṇīdattas Pañcattattvaprakāśa*. The plan did not materialize, although his preparatory work was preserved and has now been published as a facsimile.⁵⁹ Roth was also the author of the oldest preserved western Sanskrit grammar, which contained many native elements too. His later Jesuit colleague Jean-François Pons mentioned that there were 'eighteen traditional dictionaries, composed with divergent methods'.⁶⁰ Pons himself translated the *Amarakośa* into Latin – a work which he compared to the onomasiological *Indiculus universalis* by his countryman and confrere François-Antoine Pomey.⁶¹ Independently of Roth and Pons, their fellow Jesuit Ernst Hanxleden had made a transcript of the *Amarakośa*, which would later come into the possession of Paulinus a Sancto Bartholomaeo, the eighteenth-century Carmelite who began an Indological career in Rome after his own return from India. Based on this copy and two other transcripts, Paulinus published the first part of the *Amarakośa*, which treated the words for heaven, in 1798. Paulinus did not fail to lay strong emphasis on the fundamental importance of the native Sanskrit tradition, next to its many inherent complexities.⁶² However, he

⁵⁷ W. Sweetman in Ziegenbalg, *Bibliotheca Malabarica*, 12–13 (and, for the catalogue, 25–6).

⁵⁸ Vatican City, Bibliotheca Apostolica, Ms Vat. Ind. 8. 'Sjasjanam est liber, Nama est nomen, Lingam genus. Nam in hoc libro distinguuntur singulae voces per genera sua' (fo. 140r); for the names of the sun, see fos. 56–57, and for the moon fos. 50–51.

⁵⁹ The edition is Roth, *Sanskrit Grammar and Manuscripts* (1988).

⁶⁰ *Lettres édifiantes et curieuses*, 8. 40, 'Ils en ont dix-huit, faits sur différentes méthodes.'

⁶¹ For the translation, see Filliozat, 'Deux cents ans d'indianisme', 86; for Pomey's dictionary, see Chapter 14, and Yeo, *Encyclopaedic Visions*, 19–20; Pons wrote, in *Lettres édifiantes et curieuses*, 8. 40, 'Celui ... composé par Amarasimha, est rangé à peu près selon la méthode qu'a suivie l'auteur de l'*Indiculus Universalis*.'

⁶² Paulinus a Sancto Bartholomaeo, *Amarasimha*, v–xii.

remained a voice in the wilderness: because of his use of the Grantha script, customary in the south, as well as his blunt insults about his British colleagues, this pioneering work was disregarded, with the result that even a number of valuable textual variants and unknown verses were only recently noticed by Claus Vogel.⁶³

Choices of Content and Form in the Composition of Wordlists

Lexicographers had the option to create word-oriented wordlists, with brief entries in the spirit of Nebrija, or extensive, encyclopedic, world-oriented entries.⁶⁴ From the wordlists examined it is clear that the larger part by far limit themselves to presenting simple equivalents for each headword. Grammatical or encyclopedic information tends to be entirely absent. It is striking that the proportion of nominals in a number of non-missionary wordlists is much greater than that of verbs. An example is the Sinhalese–English wordlist which was found in the legacy of the English captain Robert Knox, who had spent years imprisoned in Ceylon.⁶⁵ It seems that only a minority of dictionaries contain entries which pay attention to collocations or phraseological expressions. Ribeiro’s Konkani dictionary turns out to be extremely well furnished in this regard, offering a large number of euphemistic expressions referring to death, such as *tachem ghoddvattem bhorlem* (‘his clock went full circle’), and other proverbs, such as *hantullnna sorso pao posruncho* (‘one should stretch the legs according to the bed’, in other words ‘one has to live within one’s means’).⁶⁶ The bidirectional Malay wordlist of Caspar Wiltens presents a separate list of ‘manners of talking’ at the end, containing expressions such as *matti de baccar* ‘to be burnt alive’.⁶⁷ The Malayalam lexicon of Hanxleden also sometimes includes phrasemes as lemmata.⁶⁸ Paulinus a Sancto Bartholomaeo distinguishes himself by the isolated publication of a hundred frequent sayings in Malayalam, which, he says, were used by pagans and Christians alike, so that even children uttered them all the time.⁶⁹

⁶³ Vogel, *Anfänge des westlichen Studiums*.

⁶⁴ Smith-Stark, ‘Lexicography in New Spain’, II–16.

⁶⁵ Ferguson, ‘Robert Knox’s Sinhalese vocabulary’.

⁶⁶ Rodrigues, ‘Glimpses of the Konkani language . . . proverbs and idioms’.

⁶⁷ Wiltens, *Vocabularium*, sigs. R1r–R4r, ‘Manieren van spreecken’; sig. R3v, ‘*matti de baccar*. levendich gebrant werden’.

⁶⁸ Guptan Nair, ‘Introduction’, 45.

⁶⁹ Paulinus, *Centum adagia malabarica*, 2, ‘Adagia quae heic afferuntur in Malabarica inter gentiles et Christianos S. Thomae nuncupatos obtinent, eaque adeo omnibus sunt familiaria, ut etiam pueri ea semper in ore habeant’.

All the same, a number of extant wordlists do have a more encyclopedic outlook. In this regard Gregory James has uncovered a fascinating dynamic in the publication of de Proença's Tamil dictionary in 1679. In the printed version one finds, under the lemma *Pañcākṣaram*, a technical term for a pentasyllabic mantra, the somewhat obscure Portuguese interpretation 'Five letters positioned in a particular pattern, which represent the human body, on which is based the whole science of the yogis'.⁷⁰ Manuscript M-34 of the State Central Library in Panaji (Goa), however, contains a much more extensive entry, which was even furnished with a diagram, the first illustration in the history of Tamil lexicography.⁷¹ Ecclesiastical censorship would probably have prevented the appearance of such material in print.

In recent research, much attention has been given to the way missionaries rendered typically Christian concepts, such as the trinity, baptism, and heaven, in the local languages. As Cristina Muru has noted, most of them 'made an effort to find a form in the receptor language . . . which could have had a similar referent in the source language, but with a different culturally oriented significance between the two'.⁷² She provides the example of the concept 'hell', which, in the corpus of Tamil dictionaries and religious texts examined by her, is invariably represented by *naraka*, a Sanskrit word which refers to the 'waiting room of the dead': by no means the same thing. The missionaries appear to have trusted that the native population would eventually internalize the Christian meaning. A comparative approach might provide some further interesting insights. It is remarkable, for instance, that the same Sanskrit word *naraka* was chosen as the equivalent for *hell* in Hanxleden's Malayalam wordlist (alongside the alternative *durgati*, 'misfortune'), Simon Cat's Sinhala wordlist, and Assumpçam's Bengali wordlist. Was the choice of this Sanskrit word from Bengal all the way down to Sri Lanka so self-evident, or could there have been mutual influence between the wordlist traditions of the various languages?

Taboo words or sexually suggestive words presented missionaries with a paradox. The indecent words that were included – apparently because the lexicographer decided that missionaries needed to be aware of their existence – were sometimes provided with a Latin translation, which seemed to suggest a greater distance than Portuguese.⁷³ In his Malayalam–Portuguese

⁷⁰ Proença, *Vocabulário Tamulico* (1966), fo. 102r, 'Sinco letras postas em certa figura, que arremeda[m] a o corpo humano, em que se fu[n]da toda a sie[n]cia dos jogues.'

⁷¹ James, 'Manuscript and printed versions', 3–5; the diagram is also reproduced in James, *Colporul*, 98.

⁷² Muru, 'Shaping minds and cultures', 215. ⁷³ James, *Colporul*, 97.

dictionary, Hanxleden translates the word *cunṇi* ‘penis’, presumably cognate with the word *cunṇa* ‘elephant trunk’, which was not taken up in the dictionary, in what might be called semi-Latin, as ‘membro viril turpiter’ (‘male member, vulgar’). The previous lemma *cunṇān(pu)* may refer either to *chunam*, a kind of lime that is chewed with betel, or to coitus. The first signification is given in Portuguese, the second in Latin.⁷⁴ *Membro viril* occurs again as an entry in Hanxleden’s Portuguese–Malayalam dictionary, where it is followed by some seven possible Malayalam translations. Gregory James has shown how the ingenious de Proença solved the problem of certain taboo words by representing the headword partly in the Tamil alphabet and partly in the Latin alphabet, ‘presumably so as not to cause offence to any literate Tamil-speakers who might see the page and/or to forestall any accusations of impropriety on the part of the missionaries’.⁷⁵ Jean-Luc Chevillard has demonstrated that Latin paraphrases such as ‘membrum muliebre’ (‘female genitalia’) in de Proença’s dictionary were often very general: the same paraphrase sometimes refers to breasts, sometimes to the genitalia, and sometimes to an embryo.⁷⁶

Let us now turn our attention to a number of more formal aspects of the missionary dictionary. In the creation of a wordlist of an Indian language with its own writing culture, missionaries had various options: they could either adopt the Indian characters, or transliterate into the Roman alphabet, or both. An example of the last strategy is found in an anonymous Portuguese–Hindustani dictionary formerly in the collection of William Marsden: the Portuguese headword *Principal* is followed in the second column by the Hindi equivalent in Roman characters (*Aval*), while the third column – clearly in a different hand – offers the *devanagari* form.⁷⁷ In the reverse order two seventeenth-century Malay–Dutch wordlists in the Vatican library first offer the Malay headword in Arabic script, followed by a transliteration into Roman characters and finally a Dutch translation.⁷⁸ Wordlists entirely in Roman characters called for less specialized resources in the printing house and were accessible to a wider European readership. By far the greater part of the wordlists were ordered by the alphabet used in Europe, including the headwords represented in a native script: indeed, according to the online catalogue entry for an anonymous Malayalam–Portuguese wordlist at

⁷⁴ Hanxleden, ‘Vocabolario malabarico lusitano’, fo. 63v, ‘copula viri cum faemina’.

⁷⁵ James, ‘Aspects of the structure’, 286–7.

⁷⁶ Chevillard, ‘Challenge of bi-directional translation’, 123.

⁷⁷ London, School of Oriental and African Studies, MS 11.953.

⁷⁸ Vatican City, Bibliotheca Apostolica, Ms Vat. Ind. 6 and Ms Vat. Ind. 7.

Harvard University, the Portuguese ordering of the Malayalam headwords is explicitly mentioned on the title page.⁷⁹ This was not an inevitable decision, given that the native Indian grammatical tradition adhered to its own, strictly phonetic ordering. In his Tamil dictionary, de Proença explicitly indicated that he would not honour this ordering, due to ‘the confusion and variety of letters, so outlandish and contrary to our pronunciation’: as this shows, the Indian languages were experienced differently from languages such as Hebrew and Arabic, for which the traditional order was always respected in wordlists.⁸⁰

The importance of the layout of a dictionary should by no means be underestimated: Pierre Delsaerdts speaks in this regard of an efficient ‘machine à lire’.⁸¹ In the dictionary of Frederik de Houtman, the Dutch is printed in black-letter, the Malay in Roman, and the Malagasy in italics. The three languages are clearly separated from each other in columns divided by rules. Much less structured is the later Kurdish dictionary by Maurizio Garzoni. For instance, the lemma *Vittoria* (‘victory’) reads as follows: ‘Vittoria, Ber, *vedi il verbo Vincere*’. The Italian and Kurdish are separated only by full stops and commas, and both languages are printed in Roman type, with editorial text in italics. In some manuscripts in which more than two languages are presented, the columns are spread over facing pages, as in a Portuguese–High Malayalam–Latin dictionary by the extremely prolific Carmelite Stephanus a.S.S. Petro et Paulo.⁸²

We have very little information about the making of these missionary wordlists beyond what can be inferred from the internal evidence of the finished texts.⁸³ Beschi was exceptional in that, in some of his prefaces, he disclosed some details of how he proceeded.⁸⁴ As we have seen, however, the missionary wordlists usually lack prefaces. Further study of the correspondence of the missionaries may shed light on the question. So may comparison of wordlists (and grammars) and the translations of key religious texts, such as catechisms, which were made before them and were presumably a basis for them.⁸⁵ In terms of material organization and alphabetization,

⁷⁹ Cambridge, MA, Houghton Library of Harvard University, MS Port 12, ‘Ordem das letras conforme ao estilo portuguez’.

⁸⁰ Proença, *Vocabulário Tamulico* (1966), fo. [vi]v, ‘o embaraço, e uarietade de letras tam co[n]junctas, e barbaras a nossa pronúcia’; translation by James, ‘Aspects of the structure’, 276.

⁸¹ Delsaerdts, ‘Kennismachines van Cornelis Kiliaan’.

⁸² Vatican City, Bibliotheca Apostolica, Ms Vat. Ind. 16.

⁸³ James, *Colporul*, 126; cf. Smith-Stark, ‘Lexicography in New Spain’, 27–33.

⁸⁴ Gnanapragasam, ‘Contributions of Fr Beschi’, 350.

⁸⁵ Xavier and Županov, *Catholic Orientalism*, 293; Smith-Stark, ‘Lexicography in New Spain’, 32.

we know that Bartholomäus Ziegenbalg had been working on *olai* before his death, which were copied onto paper thereafter.⁸⁶ One aspect of the procedure of the missionary lexicographers is certain: there can be no doubt that they were strongly dependent on local informants, despite the fact that this co-operation was rarely brought to the attention of the reader.⁸⁷ Paulinus a Sancto Bartholomaeo himself mentions how the native scholar Francisco Texeira helped him to get his hands on a copy of the *Amarakoša* in 1778 (we find the same name in a manuscript of the *Amarakoša* in the Vatican library in which the year 1765 is mentioned).⁸⁸ These invisible, subaltern voices did have an important impact on the contents and structure of dictionaries; the native scholars whose aid was solicited must not be dismissed as merely mechanical and passive ‘helplines’.⁸⁹

‘Knowledge Management’ and the Fate of Wordlists

In 1556, Portuguese Jesuits founded a printing press in Goa, which could only print Roman letters. For the development of Tamil types they had to wait until the 1570s, which meant that the pioneering missionary grammarian Henrique Henriques could not immediately publish his work.⁹⁰ As far as we know, the Portuguese printed no more than eight publications in Tamil, during two brief periods: from 1577 to 1586 they printed mainly in Cochin from types cast in India, and from 1677 to 1680 in Ambalacatta with types imported from Rome.⁹¹ Although only 19 books were printed in India in the sixteenth century, and only twice that number in the seventeenth century, in the eighteenth century production increased exponentially, with more than 1,700 new books. In 1712, on the initiative of Bartholomäus Ziegenbalg, a printing press was installed in Tranquebar. As a result of this, printing activity moved to the east coast (the Danish Tranquebar, the English Madras, the French Pondicherry, and the Dutch Colombo).⁹² Since the Jesuits in India did not have access to a printing press in the eighteenth century, Beschi

⁸⁶ Gnanapragasam, ‘Contributions of Fr Beschi’, 383.

⁸⁷ Smith-Stark, ‘Lexicography in New Spain’, 31–2; Liebau, *Indischen Mitarbeiter*, 194; Xavier and Županov, *Catholic Orientalism*, xxxi, 287.

⁸⁸ Vogel, *Anfänge des westlichen Studiums*, 15–16; Vatican City, Bibliotheca Apostolica, MS Borg. Ind. 19.

⁸⁹ Županov, ‘Prosélytisme et pluralisme religieux’, 42; Liebau, *Cultural Encounters in India*, 6.

⁹⁰ Blackburn, *Print, Folklore, and Nationalism*, 33–4.

⁹¹ Shaw, ‘Copenhagen copy of Henriques’ *Flos Sanctorum*’, 39–40.

⁹² Blackburn, *Print, Folklore, and Nationalism*, 43–4.

witnessed the appearance in print of only a single one of his numerous linguistic works in his lifetime, and that was from the press at Tranquebar, property of the Lutherans, with whom he was constantly at war.⁹³ When he noticed that his Tamil grammar had been furnished with a 'pedantic' appendix by Christoph Theodosius Walther, he refused to allow his *caturakarāṭi* to be printed in Tranquebar as well.⁹⁴ Of the 368 works which made their way to the printing presses of Calcutta between 1778 and 1800, 7 were lexicographical (3 Persian, 2 Bengali, and 2 Hindustani), and 11 grammatical (most of the printed works were almanacs and government publications).⁹⁵ The fact that printed lexicographical publications would boom in the nineteenth century may easily be inferred from systematic biographies like that by J. T. Zenker.⁹⁶

Some historians have suggested that many missionary grammars and dictionaries failed to reach the press because the missionaries themselves regarded their works as unfinished products, or as continually in progress and thus intrinsically incomplete.⁹⁷ But some lexicographers did explicitly request the publication of the dictionaries they had made. In 1703 the French Capuchin François-Marie de Tours even undertook a mission to Rome, one of the goals of which was to ensure the publication of his Hindustani grammar and dictionary by the Congregatio de Propaganda Fide.⁹⁸ His attempts proved to be in vain.

The case of de Proença's dictionary illustrates how many obstacles a dictionary had to surmount before it could pass through the press. Every manuscript prepared by a member of the order had to be judged on no fewer than three separate levels. Although three older Jesuit fathers, all of them acquainted with Tamil, had declared that the work was not in contradiction with the holy doctrine of the faith and provided positive advice, it had to wait more than a year for permission from the inquisition of Goa, which was known for its strictness and had for this reason taken extra advice from a secular priest. Even after this, the inquisition insisted that the printed copy should be compared with the approved manuscript, to ensure that no extra information had found its way in: as we have seen, some manuscripts of the dictionary included material which was not in the approved printed

⁹³ Blackburn, *Print, Folklore, and Nationalism*, 49.

⁹⁴ Blackburn, *Print, Folklore, and Nationalism*, 60.

⁹⁵ Shaw, *Printing in Calcutta to 1800*, 41; cf. Khan, 'Early history of Bengali printing', 51, 54.

⁹⁶ Zenker, *Bibliotheca orientalis* (1861). See also Xavier and Županov, *Catholic Orientalism*, 316–17.

⁹⁷ Xavier and Županov, *Catholic Orientalism*, 220, 309–10. See also Guptan Nair, 'Introduction', 37.

⁹⁸ Aranha, '*Vulgaris seu universalis*', 333–5.

version.⁹⁹ The bidirectional dictionary of Portuguese and Bengali by Manoel da Assumpção moved through the approval process a little more rapidly: of the eight censures and licences among its prelims, the first is dated 6 December 1742, and the last 5 November 1743.

Besides such ideological obstacles, the material difficulties with which early modern missionary printing projects were confronted (shortage of paper, no suitable types or worn ones, no competent printers) should not be underestimated.¹⁰⁰ It took more than fifty years before Ziegenbalg's initial project of composing a Tamil dictionary produced a printed edition (Johann Philipp Fabricius' bidirectional Tamil–English dictionary) in the last quarter of the eighteenth century.¹⁰¹ In the meantime, Ziegenbalg's successors had been encouraged to copy the dictionary, 'which at the same time gives them an exercise in the language'.¹⁰² The dictionaries which were created against the background of the East Indian trade companies, such as the VOC and the British East India Company, might enjoy financial backing.¹⁰³ But all our information suggests that the VOC directors never gave instructions to finance linguistic studies or publish linguistic works, even though fellowships were sometimes allocated for botanical studies.¹⁰⁴ It is well known that Nicolaas Witsen, one of the VOC directors as well as a linguistic enthusiast, complained about the general lack of 'scholarly curiosity in India' – 'No Sir, it is only money and no knowledge which our people are after, which is regrettable' – and about the lack of interest in Herbert de Jager's (now partly destroyed) linguistic legacy in particular.¹⁰⁵

And when a dictionary was finally printed, its successful distribution was far from assured. Only a single copy of de Proença's dictionary is known today; it lies in the Vatican library, although Maria do Céu Fonseca mentions the existence of a copy in the national library of Lisbon.¹⁰⁶ Balthasar da Costa's Portuguese–Tamil dictionary was supposedly printed on the same press as de Proença's, but no copy is known to be extant.¹⁰⁷ And it is said of

⁹⁹ James, 'Aspects of the structure', 277, 290–5.

¹⁰⁰ Blackburn, *Print, Folklore, and Nationalism*, 35–6, 40. See also Xavier and Županov, *Catholic Orientalism*, 294.

¹⁰¹ James, *Colporul*, 157. ¹⁰² Quoted in Gnanapragasam, 'Contributions of Fr Beschi', 388.

¹⁰³ Winterbottom, *Hybrid Knowledge*, 59.

¹⁰⁴ Den Besten, 'Badly harvested field', 267; Baas, 'De VOC in Flora's luthoven', 200.

¹⁰⁵ Quoted in Pytlowany and Van Hal, 'Merchants, scholars, and languages', 23 n. 4, 'Wat vraegt UwelEd. na de geleerde curieusheyt van Indiën? Neen Heer, het is alleen gelt en geen wetenschap die onse luyden zoeken aldaer, 't gunt is te beklagen', translated in Huigen, 'Introduction', 8 (see also Rietbergen, 'Wie verre reizen doet', 164); for De Jager's legacy, see Peters, *De wijze koopman*, 229–30.

¹⁰⁶ Fonseca, *Historiografia linguística portuguesa e missionária*, 338. ¹⁰⁷ James, *Colporul*, 96.

John Fergusson's *Dictionary of the Hindostan Language* of 1773 that almost no copies have been preserved, due to the fact that the edition was destroyed, almost in its entirety, in a shipwreck.¹⁰⁸

Obviously, the transmission of manuscripts was even more precarious. A number of librarians in Europe, such as the abbé Bignon in Paris and Mathurin Veyssière de la Croze in Berlin, played a crucial role in their preservation.¹⁰⁹ Some missionaries were puzzled by the European interest in their work – some works published by the Propaganda Fide, although not written for a scholarly readership, were swiftly reviewed in eighteenth-century magazines – of what use, they wondered, could dictionaries of a wholly exotic language be to European scholars?¹¹⁰ Other missionaries did eagerly respond to this demand. Between 1729 and 1735, French Jesuits operating in India transmitted more than 150 manuscript volumes to the Royal Library, thus creating the backbone of its Indian collection.¹¹¹ The Jesuits and the Congregatio de Propaganda Fide accelerated the import of linguistic tools of Indian languages: by the end of the sixteenth century, two generals of the Society of Jesus decided that dictionaries and grammars should be sent to Rome from India to allow future missionaries to prepare themselves optimally for their task.¹¹²

In India it transpired that European paper did not do well in the warmth and the heat; worse still, Bartholomäus Ziegenbalg's personal library was not only consumed by insects, but sometimes even served as fuel for the fire.¹¹³ The research library of the Discalced Carmelite order in Verapoly may have been destroyed in a flood.¹¹⁴ But it was not only in India that water and fire caused destruction. The Lisbon earthquake of 1755 not only resulted in the loss of life of thousands, but also destroyed many archives and libraries in their entirety.¹¹⁵ A large-scale project was ongoing at the time, with the aim of archiving the knowledge which had been accumulated in the colonies and putting it to writing. It is in this light that the *Bibliotheca Lusitana* (1741–59) should be understood: these four volumes aim to provide an overview of the

¹⁰⁸ Zenker, *Bibliotheca orientalis*, I.234.

¹⁰⁹ Filliozat, 'Deux cents ans d'indianisme', 106; Van Hal, 'Protestant pioneers in Sanskrit studies', 104–10.

¹¹⁰ For the puzzlement, see, for example, Filliozat, 'Deux cents ans d'indianisme', 106; for the reviews, see Xavier and Županov, *Catholic Orientalism*, 296.

¹¹¹ Colas, 'Cultural encounter', 69; Xavier and Županov, *Catholic Orientalism*, 303–8.

¹¹² Županov, 'Language and culture of the Jesuit "early modernity" in India', 89.

¹¹³ W. Sweetman in Ziegenbalg, *Bibliotheca Malabarica*, 21.

¹¹⁴ Xavier and Županov, *Catholic Orientalism*, xxx. See also Thankappan Nair, 'Father Hanxleden', 51.

¹¹⁵ Xavier and Županov, *Catholic Orientalism*, xxiii.

varied cultural and scientific work which had been done by the Portuguese, including an overview of the grammars and dictionaries composed by Portuguese missionaries. Many of the titles mentioned here were lost in one fell swoop in the earthquake.

Natural disasters were not the only reason for the loss of much of the early modern knowledge related to Indian languages which the Europeans had accumulated. We have already seen that some of the agents in knowledge management were less efficient than the Jesuits, including the Franciscans and the officials of the VOC.¹¹⁶ Even the Jesuits failed to systematize their philological knowledge.¹¹⁷ For Europeans, the many borders within the Indian subcontinent were permeable to agents of various orders, religious denominations, and nationalities. Hence, the fruits of scholarship became very nomadic, and consequently vulnerable, for as soon as the lexicographic work of a Catholic missionary fell into the hands of an outsider – whether he belonged to a different Catholic order or was a Dutch or British merchant – the danger of appropriation and oblivion was real.¹¹⁸ One modern author describes lexicographical manuscripts as ‘sometimes in a condition of criminal neglect’.¹¹⁹ Manuscripts of ‘outsiders’ were, for instance, not always registered in archival catalogues of certain missionary orders.¹²⁰

It was mainly Paulinus a Sancto Bartholomaeo who saw the urgency of recording the Indological achievements of his Catholic predecessors. Publications of his are our only source for some of our knowledge of, in particular, South Indian missionary lexicography.¹²¹ His substantial bibliographical survey ‘Bibliotheca Indica’ (1803), with much lexicographical detail, remains unpublished.¹²²

Paulinus’ bibliographical diligence may well have been a result of his anxiety about the activities of the British in Calcutta, whom he frequently criticized in his writings. Ângela Barreto Xavier and Ines G. Županov in particular argue that many successful nineteenth-century British Orientalists

¹¹⁶ Indeed, the VOC might positively suppress knowledge: see Pytlowany and Van Hal, ‘Merchants, scholars, and languages’, 26–7.

¹¹⁷ Xavier and Županov, *Catholic Orientalism*, 308; see also Županov, ‘Compromise: India’, 367.

¹¹⁸ Xavier and Županov, *Catholic Orientalism*, xxxii–xxxiv; 291.

¹¹⁹ Rodrigues, ‘Glimpses of the Konkani language ... religious concepts and rituals’, 19.

¹²⁰ See for an example Van Hal, ‘À la recherche d’une grammaire perdue: Johann Ernst Hanxleden’s *Grammatica Grandonica* retrieved’.

¹²¹ See the discussions of his work in Županov, ‘Orientalist museum’, and Xavier and Županov, *Catholic Orientalism*.

¹²² For it, and other unpublished works by Paulinus, see Jauk-Pinhak, ‘Some notes on the pioneer Indologist Filip Vesdin’, 137 n. 37; an edition is in preparation.

made surreptitious use of the products of early modern Catholic – and to a lesser extent also Protestant – Orientalism. The twentieth-century historiography of nineteenth-century British Orientalism is increasingly criticized: it is being claimed that historiographers have too often focused unduly on British agents alone, without asking what their sources were, and even that the exclusive attention given to William Jones in the twentieth century is an indication of British arrogance.¹²³ How can we test this claim? Was the existence of Assumpçam's Bengali dictionary of 1743 more or less consciously erased from early histories on scholarly activity in Bengal, or was the work overlooked simply because it was hard to come by?¹²⁴ The presence of considerable, but relatively ignored, collections of early modern documents of Catholic missionary origin, including lexicographical works, in British archives and libraries, indicates that the British certainly had access to a large amount of material (some of which has been dispersed, so that for instance a seventeenth-century Konkani–Portuguese manuscript wordlist, now in Lisbon, was bought in London in the twentieth century).¹²⁵

A number of nineteenth-century Indological lexicographers explicitly indicate that they consulted and used early modern manuscripts. The most famous nineteenth-century Malayalam scholar, Hermann Gundert, made eager use of the missionary dictionaries which he found in Verapoly.¹²⁶ The lexicographical works of the Jesuit Louis Noël de Bourzès, missionary in Madurai from 1710 to 1735, were used by nineteenth-century French lexicographers.¹²⁷ The late publication of Beschi's *Caturakarāti* in 1824 had wide-ranging results and influenced many Tamil scholars.¹²⁸ A number of lexicographical works that were published in Europe seem to have enjoyed a more limited circulation. Immanuel Kant gave lectures on 'Physische Geographie' in Königsberg for forty years, into which he integrated new insights from geography, chemistry, and biology every year. In his discussion of the Kurds he refers to Garzoni, 'who has been living there for eighteen years', and had written 'a small book on this country in the Kurdish language, which will remain the most up-to-date from this region for a long time'.¹²⁹

¹²³ Rietbergen, *Europa's India*; see also Xavier and Županov, *Catholic Orientalism*, 206.

¹²⁴ Cf. Xavier and Županov, *Catholic Orientalism*, 316–17.

¹²⁵ Xavier and Županov, *Catholic Orientalism*, 289, 325–6; *Vocabulario da lingoa Canarina* (1973), unpaginated introduction, section 1.

¹²⁶ Guptan Nair, 'Introduction', 39. ¹²⁷ Filliozat, 'Deux cents ans d'indianisme', 108.

¹²⁸ Blackburn, *Print, Folklore, and Nationalism*, 60.

¹²⁹ Kant, *Physische Geographie*, II.1, 166, 'Garzoni, ein römischer Missionair, der sich dort 18 Jahre aufgehalten hat, hat ein Werkchen von diesem Lande in der Kiürdischen Sprache geschrieben, welches wohl noch lange das Neueste aus dieser Gegend enthalten wird.'

The less than accurate description of Garzoni's work prompts one to doubt if Kant ever saw it.

Conclusions and Outlook

The early modern European lexicography of the Indian languages was in the first place instrumental. The missionaries wanted to save indigenous souls through knowledge of their language, while the trade companies saw a chance to increase their profits. In an encomium at the end of De Houtman's lexicon the poet expresses his explicit hope that 'through this beautiful language, profits will increase considerably'.¹³⁰ Merchants had more access to interpreters and translators than missionaries did. However, in relying on interpreters, the merchants ran the risk of being cheated by these intermediaries. This was why the commanders of the VOC of Fort Golconda decided that they should have sufficient mastery of Telugu to be able to express short messages themselves and to assess the work of the translators.¹³¹ Nevertheless the lexicographical activity carried out in the context of the VOC seems to have been limited to Calvinistic pastors, whose missionary activities have been underestimated.¹³² The primarily instrumental nature of missionary language study did not prevent the most talented missionaries, at times against the fundraisers' will, from developing an authentic interest in indigenous cultures and rites, which proved in its turn to be instrumental and helpful in converting souls. Tellingly, some of the most gifted lexicographers, such as Beschi, Hanxleden, and Gonçalves, produced poetry in the indigenous languages.¹³³

One can wonder to what extent the missionaries' lexicographical activity contributed to the general nineteenth-century Indomania, which led to the creation of numerous chairs in Indology and (Indo-European) linguistics.¹³⁴ The connection between the two seems to have been relatively restricted. In the first place, the letters of these missionaries were much more accessible to European scholars than their lexicographical works, of which the circulation, as we have seen, remained very limited. In the second place, the similarities between the European languages and Sanskrit do not appear to have greatly

¹³⁰ In Houtman, *Spraecck ende Woord-Boeck*, [223], 'door dees schoone Tael, 'tprofijs seer sal vermeeren'.

¹³¹ Van Dam, *Beschryvinge van de Oostindische Compagnie*, II.2, 164.

¹³² Meuwese, 'Language, literacy, and education', 107.

¹³³ James, *Colporul*, 106; Guptan Nair, 'Introduction', 38.

¹³⁴ See, for example, Rabault-Ferhahn, *Archives of Origins*.

surprised all early modern missionaries, since one of the prevailing interpretations of the Babel story indicated that all the world's languages did preserve a few elements of the *lingua Adamica*, which was often equated with Hebrew. As a rule, no attention was given to linguistic kinship in the major wordlists. Some of the missionaries' letters, however, comprised small comparative wordlists, in which, for instance, Tamil was traced back to Hebrew.¹³⁵ Walther's pietist colleague Benjamin Schultze made a comparative wordlist of more than ten Indian languages, which was published in the middle of the eighteenth century.¹³⁶

Digital editions of missionary lexicographic undertakings, at least two of which are currently underway, will allow us to examine in a more systematic vein the connections between missionary wordlists of one single Indian language, between missionary wordlists of different Indian languages, and between the contents of missionary grammars, dictionaries, and religious texts in general.¹³⁷

¹³⁵ For instance a letter of 1725 from C. T. Walther to J. W. Zierold, now Halle, Franckesche Stiftungen, AFSt/M 1 B 2: 28.

¹³⁶ [Schultze and Fritz], *Orientalisch- und Occidentalisches Sprachmeister*, between 212 and 213; Abel, *Symphona symphona*, is an unacknowledged reprint.

¹³⁷ For Tamil, see Chevillard, 'Challenge of bi-directional translation', 126–7; Heinz Werner Wessler is supervising the research project 'Hindi lexicography and the cosmopolitan in the encounter between Europe and India around 1700'.

Missionary and Subsequent Traditions in Africa

GONÇALO FERNANDES

According to the latest calculation of *Ethnologue*, the African continent has more than 2,000 languages. The lexicography of this vast body of languages until the nineteenth century was very modest. Their codification in grammars and dictionaries increased dramatically after the Berlin Conference of 1884–5 formalized the process of the European colonization of Africa, and the history of linguistics of the African languages became a subject of scholarly attention in the twentieth century.¹ Apart from the establishment of the Xhosa Dictionary Project in 1968, only in the last decade of the twentieth century were there considerable developments in African lexicography: in 1991 the journal *Lexikos* – the only journal in Africa devoted to lexicography – was launched; in 1992, the African Language Lexical Project (ALLEX) was formed (between the Universities of Oslo and Zimbabwe); in 1995, the African Association for Lexicography (AFRILEX) was founded, with *Lexikos* as its official journal from 1996; and, in 2000, the African Languages Research Institute (ALRI) of the University of Zimbabwe was established.²

This chapter presents an overview of selected dictionaries of African languages, underlining the role of missionaries, both Catholic and Protestant, and native speakers since the beginning of European colonization. It is not a comprehensive analysis of the lexicography of all African languages: many have been excluded because of lack of space, and others are treated, or at least mentioned briefly, in other chapters, for instance Ancient Egyptian and Coptic (Chapter 2), Arabic (Chapters 8 and 19), Nubian (Chapter 11), Ge'ez (Chapter 14), and Afrikaans (Chapter 22). Nor does it attempt to treat all the dictionaries of each language. The main objective is to complement the research of other scholars – Clement Doke's *Bantu* deserves

¹ Doke, 'Early Bantu literature'; Doke, 'Bantu language pioneers'.

² Hartmann and James, *Dictionary of Lexicography*, 4 (entries 'African Association for Lexicography' and 'AFRILEX'), 87 (entry '*Lexikos*'); Chabata, 'Lexicography of Shona', 949.

special mention here – and to point out the pioneering and lesser-known milestones of African lexicographic works.

The languages are divided into three main geographical zones: western-central, eastern, and southern Africa. (The western and central zones could have been separated, and assigning a given language to one zone rather than another is not always straightforward, but the division imposes some order on the languages.) No chronological subdivision has been attempted; for most languages, such a subdivision would hardly be appropriate, because only a few dictionaries or vocabularies were written before 1860.³

Western-Central Africa

We begin with the languages of western-central Africa, in the order of their first dictionaries: Kikongo, of which there is an extant dictionary of 1648, together with Kimbundu and Umbundu; then Wolof, of which the first dictionary is of 1825; then Yoruba, of which the first dictionary is of 1843; and finally Hausa, of which the first dictionary is also of 1843.⁴

Kikongo

The oldest dictionary of a Bantu language which is known to survive is the *Vocabularium Latinum, Hispanicum, et Congense ad usum missionariorum transmittendorum ad regni Congi missiones* ('Latin, Spanish, and Kikongo vocabulary for the use of the missionaries sent to the missions in the Kingdom of Kongo'), finished in mid 1648.⁵ It has been called 'the oldest Bantu dictionary' *tout court*, but there is reason to believe that earlier wordlists have been lost.⁶ It is preserved in a manuscript at the Italian National Library in Rome, written in 1652, and there is an edition of 1928.⁷ The latter cuts and rearranges

³ For 1860 as a date separating pioneering and later work, see Doke, *Bantu*, 74 and 83.

⁴ The following account is supplemented by, for instance, Busane, 'Lexicography in central Africa', and Mavoungou, 'Lexicography of the languages of central Africa'; see also Afane-Otsaga, 'Lexicography of Fang'.

⁵ For it, see D'Alençon, 'Essai de bibliographie', 41–2; Hildebrand, *Le martyr Georges de Geel*, 260–3; Buenaventura de Carroceria, 'Los capuchinos', 219–26; Cole, 'History of African linguistics', 2; Bonvini, 'Repères pour une histoire', 140; Zwartjes, *Portuguese Missionary Grammars*, 297; De Kind, de Schryver, and Bostoen, 'Pushing back the origin of Bantu lexicography', 183; Bortolami, *I Bakongo*, 32.

⁶ See Doke, 'Early Bantu literature', 96 – 'the claim of the editors that this is "the oldest Bantu dictionary" is difficult to substantiate. Brusciotto's 1650 dictionary, not now extant, is the oldest of which we have a record' – and Zwartjes, *Portuguese Missionary Grammars*, 212.

⁷ Rome, Biblioteca Nazionale Vittorio Emanuele, Fondo Minori 1896, MS Varia 274, edited as *Le plus ancien dictionnaire bantu* (1928).

to turn the Latin–Spanish–Kikongo original into a Kikongo–French–Flemish work which is a misleading representation of the original, and has been forcefully criticized.⁸ The manuscript is 121 folios in extent, and includes approximately 7,000 lemmata.⁹ The dictionary has been attributed to the Fleming Adriaen Willems (who later adopted the name Joris van Gheel or, in Spanish, Jorge de Gela), who was the copyist of the extant manuscript. In fact, the main compiler was very likely a Portuguese- and Kikongo-speaker, Manuel de Roboredo (later Francisco de São Salvador), whose father and mother were, respectively, a Portuguese nobleman and a relative of the king of Kongo. It is possible that the *Vocabularium* and the lost Latin–Spanish–Italian–Kikongo work of the Capuchin missionary Antonio Teruel, which was sent to the Congregatio de Propaganda Fide in Rome in 1662, share a common ancestor, itself now lost.¹⁰

In 1805, the Italian Capuchin Bernardo Maria da Canicattì published a grammar of the ‘Bunda or Angolan language’ and, at the end, added a Portuguese–Latin–Kikongo–Kimbundu vocabulary with approximately 1,000 entries. It was much criticized because it confused forms from at least three native languages: Kimbundu, Kikongo, and Umbundu; however, the English missionary William Holman Bentley acknowledged Canicattì’s philological contribution, saying, for instance, that ‘there are many mistakes, and many words which it is impossible to trace; but as he acknowledges his imperfect acquaintance with Kongo, and only gives his list as philological study, we must not criticize, but be thankful for his contribution’.¹¹ In fact, Canicattì knew his own limitations and the shortcomings of his vocabulary. He complained that he had no access to dictionaries, grammars, or other linguistic resources apart from two seventeenth-century missionary works: the catechism by Marcos Jorge, translated into Kikongo by Mattheus Cardoso, and (for a short period of time) the grammar book by Giacinto Brugiotti da Vetralla. Thus, he concluded that his knowledge was very limited.¹² Canicattì tried to describe the ‘Sonho’ dialect, which does not

⁸ For its editorial procedure, see *Le plus ancien dictionnaire bantu* (1928), xvi; for criticism, see Doke, ‘Early Bantu literature’, 96 (‘such a method of handling the manuscript is the opposite of scientific’); Hildebrand, *Le martyr Georges de Geel*, 269; and Zwartjes, *Portuguese Missionary Grammars*, 297.

⁹ Zwartjes, *Portuguese Missionary Grammars*, 297, follows an unreliable secondary source for his figure of 169 folios.

¹⁰ For Teruel’s work, see Brásio, *Monumenta Missionaria Africana*, XII.369–70, and Zwartjes, *Portuguese Missionary Grammars*, 298–9; for its possible source, see Doke, ‘Early Bantu literature’, 96; Buenaventura de Carrocera, ‘Los capuchinos’, 226–9.

¹¹ Bentley, *Dictionary and Grammar*, xii.

¹² Canicattì, ‘Diccionario abbreviado da lingua Congueza’, 151–3.

mean the Kikongo language from San Salvador, the capital, but from the coast, Sonyo or Saint Antonio, 'on the left bank at the mouth of the river'.¹³ He did not separate the prefixes from the roots or stems, concluding that the fundamental structures of Kikongo are indeed the 'initial letters or syllables'. They 'govern and distinguish the words, as in Kimbundu, and not the terminations'. This similarity proves, for him, that 'both nations had the same origin'.¹⁴

A dictionary of Kikongo which is still in use is the English–Kikongo and Kikongo–English *Dictionary and Grammar of the Kongo Language* (1887, with an appendix of 1895) by Bentley, who was assisted by a Congolese Christian called Nlemvo, acknowledging his work in an interesting passage:

In this translation and linguistic work, Nlemvo, who rendered such valuable assistance in the preparation of what was published in 1887, has still continued his aid, rendered all the more efficient by these fourteen years of work, which have trained and developed his great natural aptitude. This gives the uniformity which is of such great importance.¹⁵

In the Preface, Bentley summarizes the history of the presence of Europeans in the Congo region, and explains how the Baptist Missionary Society made San Salvador (currently M'banza Kongo, Angola) the base of their operations in 1879.¹⁶ In addition, he highlighted the absence of serious dialectal variants of, as he calls the language, 'Kixi-Kongo', saying, for instance, that the differences between the language of San Salvador and of Stanley Pool in the interior (the current Lake Nkunda or Pool Malebo, which separates Kinshasha and Brazzaville) were smaller than those between the English of southern England and 'broad Scotch'. Thus, for Bentley, the best medium of communication was the Kikongo language as spoken in the old capital, San Salvador.¹⁷

The dictionary as published in 1887 included 'some 10,000 Kongo words, omitting as far as possible the thousands of derivative words, which, being formed from the root-words according to simple rules, needed no special note', and the appendix of 1895 offered 4,000 more Kikongo words, 'on the same principle, which include, as far as possible, all words or roots which are used in the Kongo literature of the English Baptist Mission published up to the present'.¹⁸ The appendix used a revised orthography, which Bentley

¹³ Bentley, *Dictionary and Grammar*, xii.

¹⁴ Canicatti, 'Diccionario abreviado da lingua Congueza', 154.

¹⁵ Bentley, *Appendix*, v; cf. the account of Nlemvo in Bentley, *Dictionary and Grammar*, xviii.

¹⁶ Bentley, *Dictionary and Grammar*, ix–xvi. ¹⁷ Bentley, *Dictionary and Grammar*, xi.

¹⁸ Bentley, *Appendix*, v.

admitted ‘may cause some difficulty . . . since *simba* appears as *ximba* under X in the Dictionary, and as *simba* under S in the Appendix’; but, he continued, ‘that lack of uniformity is of small moment, compared with the importance of the attainment of a permanent form at the earliest possible date, and the wider usefulness of our literary productions’.¹⁹ These literary productions, particularly the translation of the New Testament and a magazine named *Se kukianga* (‘The dawn is breaking’) meant, for Bentley, that by 1895, Kikongo could ‘no longer be spoken of as an unwritten language’: the publication of the 1887 dictionary had apparently not in itself given the language written status.²⁰ For Doke, ‘Bentley’s “dictionary” is very reliable, but what he has recorded might have been done more concisely: he has an irritating way of entering the same word repeatedly for each separate meaning, instead of listing the various meanings under a common entry.’²¹

Kimbundu

During the colonial period, Kimbundu was undoubtedly the African language most widely studied by missionaries under the *Padroado real*.²² Nevertheless, the first-known printed dictionary was published only in 1804: the 720-page Portuguese–Latin–Kimbundu *Diccionario da lingua Bunda ou Angolense* by Bernardo Maria Canicatti, whose wordlist of Kikongo and Kimbundu of 1805 I have just discussed. In his address to the reader, Canicatti explains the importance of the dictionary to missionary activity, and the difficulties experienced when dealing with the interpreters, native people with little or no instruction, knowing only a few Portuguese words. He also explains its importance in improving economic or commercial relations and the application of justice. Canicatti named the local language the ‘general language of Angola’, because, according to him, it was spoken in many native kingdoms throughout the whole of what he called the Kingdom of Angola or the Kingdom of the Mbundu people.²³ He also mentions his lack of knowledge of other previous dictionaries and, thus, the imperfections of his work, because he felt he was the first to penetrate that ‘dark labyrinth’, adding that defects and imperfections are almost inseparable from the earliest works of this genre.²⁴

¹⁹ Bentley, *Appendix*, vi. ²⁰ Bentley, *Appendix*, v. ²¹ Doke, *Bantu*, 18–19.

²² For their work on Bantu linguistics, see, e.g., Louwrens, ‘Contributions made by the Portuguese’.

²³ Canicatti, *Diccionario da lingua Bunda*, vii–ix.

²⁴ Canicatti, *Diccionario da lingua Bunda*, v; cf. Chatelain, *Folk-Tales of Angola*, 23, on its ‘incorrectness, confused spelling, and erroneous renderings of words’.

In 1893 was published the first dictionary of Kimbundu by an Angolan, titled *Ensaio de Dicionario Kimbundu-Portuguez*. Its author, the poet and journalist Joaquim Dias Cordeiro da Matta, did not have any religious education (only primary school) and died prematurely, but is considered the ‘father of Angolan literature’. Chatelain, probably the world’s leading expert on Kimbundu in the nineteenth century, stated that Matta’s *Ensaio* was ‘the best vocabulary of Ki-mbundu yet published’.²⁵ Interestingly, Matta said Kimbundu had maintained its independence from Portuguese and it was the same language as that spoken a hundred years ago, contrasting with the creoles of other countries.²⁶ He also stated, however, that it would be an error to exclude neologisms from other languages, because new social realities need to have new vocabularies. He gives the example *ngálúfu* (plural *jingálúfu*), from the Portuguese *garfo* ‘fork’, to express the new social reality. Until the arrival of the Europeans, speakers of Kimbundu just used a knife to eat and, for that, they had the noun *pôku*, but new vocabulary had come with new ways of living.²⁷ Unfortunately, Cordeiro da Matta did not realize that Kimbundu was a prefixal language – or, at least, did not realize the relevance of the fact for the structure of a dictionary – and so he did not separate the classifier prefixes for the purposes of alphabetization. For instance, he disconnected *Ambundu* from *Kimbundu* and *Mbundu*, and put all infinitive verbs with the prefix *ku* together.²⁸

Finally, the principal dictionary now in use is the *Dicionário Etimológico Bundo-Português* (1951) by the Spiritan father Albino Alves. It analyses the etymology of almost 20,000 words, and presents about 2,000 proverbs and 200 riddles. It was published in two heavy volumes, approximately 2,000 pages; the orthography follows the guidelines published as *Practical Orthography of African Languages* by the International Institute of African Languages and Cultures in 1927 and 1930. In the prologue, Alves stated that, when writing the dictionary, he used nearly sixty books and opuscles and his own knowledge, acquired during his fourteen years dedicated to learning Kimbundu, through direct contact with the native people, but that he had had no help in the etymological descriptions, because there were no treatises on this subject.²⁹ He began with a very imaginative hypothesis: all Kimbundu

²⁵ Chatelain, *Folk-Tales of Angola*, 25. ²⁶ Matta, *Ensaio de dicionario*, x.

²⁷ Matta, *Ensaio de dicionario*, xi, citing Sousa e Oliveira and Francina, *Elementos grammaticaes da lingua Nbandu*, xi.

²⁸ Matta, *Ensaio de dicionario*, 1 (*Ambundu*), 24 (*Kimbundu*), 98 (*Mbundu*), 33–85 (infinitives in *ku*).

²⁹ Alves, *Dicionário Etimológico*, 6.

words derive from two onomatopoeic sounds, *ta* and *hu*. The monosyllable *ta* referred originally to the idea of throwing up an object and *hu* meant the interior of a person.³⁰ Alves explained also that he did not repeat all words formed with a known affix, because it would be tedious to enumerate all verbs ending in *wa*, *isa*, and *ila* or *ela*, or all nouns beginning with *u* and ending in *i* or beginning with *ukwa*. From all verbs one can make new words, adding these and similar prefixes and suffixes; thus, it would not be worth multiplying the volume of the dictionary by four or six.³¹ Alves was the first author to explain that Kimbundu has four tones. He used impressive designations – ‘alto . . . normal . . . baixo . . . plangente’ (‘high, normal, low, and plangent’) – and indicated the tones with Portuguese diacritics: the acute accent for the high tone, the grave accent for the low tone, and the diaeresis for the plangent tone, ‘remembering’, as he put it, ‘the tearful eyes of the passionate fado singer on the vowel of the lower tone syllable’.³² Nevertheless, he added that many European people could not distinguish the four tones, and their practical importance was not so great, since words which could only be distinguished by the contrast of tones in isolation can usually be distinguished in practice by their conversational context.³³

Umbundu

Umbundu is another language of Angola, often mistaken for Kimbundu, which is spoken further north. The first printed Umbundu dictionary was published in Boston, Massachusetts, in 1885, by two American missionaries, William Henry Sanders and William Edwards Fay. It lists approximately 3,000 words in 76 pages, but, unfortunately, it has no linguistic explanation: there is no introductory chapter or preface, and the words are not even divided into prefixes and roots, although there are entries in the Umbundu–English section for the infinitive and negative prefixes *oku* and *ka*, and the suffixes *-vo* ‘also’ and *-ñõ* ‘only’ or ‘just’, and the last of these is given as an equivalent for *just* and for *only* in the English–Umbundu section. Based mainly on this vocabulary, a layman, José Pereira do Nascimento, a medical officer in the Portuguese Royal Navy, appended Portuguese–Umbundu and Umbundu–Portuguese wordlists to his grammar of Umbundu, published in 1894. These have approximately the same entries as Sanders and Fay’s dictionary, but with a few improvements: Nascimento

³⁰ Alves, *Dicionário Etimológico*, 10–12. ³¹ Alves, *Dicionário Etimológico*, 13.

³² Alves, *Dicionário Etimológico*, 6. ³³ Alves, *Dicionário Etimológico*, 6.

separated all prefixes, roots and suffixes, and he quoted very modern Africanists, as well as older grammarians and lexicographers.

The dictionary by the French Spiritan father Grégoire Le Guennec and his Portuguese confrere José Francisco Valente, which was published in 1972, also deserves to be highlighted. Le Guennec died twelve years before the publication of the dictionary, but the work which he had left in manuscript was, in fact, its main source. Le Guennec had dedicated twenty years to the composition of the dictionary, but the Portuguese language was not his mother tongue, and he was not comfortable with it. Valente introduced many neologisms, which, in his opinion, enriched the language and represented the 'ação civilizadora' ('civilizing achievement') of the missionaries. The dictionary adopts phonetic orthography, slightly modified, and presents only the root of the verbs, for example, by omitting the verbal prefixes.

Wolof

Our survey now moves northwards up the Atlantic coast to Wolof, which is now spoken in Senegal, Gambia, and Mauritania. The first proper Wolof dictionary was published only in 1825: the French–Wolof–Bambara and Wolof–French *Dictionnaire français–wolof et français–bambara, suivi du dictionnaire wolof–français* by Jean Dard, the founder of the first public school for black people in Senegal. According to Bonvini, the main concerns of Dard were to prove that the black people, as well as their language, belonged fully to humanity.³⁴ In fact, Dard argued that African languages are primitive, but regular, although not yet written:

this dialect is regular, uniform, governed by fixed principles; among its imperfections, advantages are to be found which do not belong to other languages at all; in conclusion, it has its own characteristics, and others which it shares with the languages of east Africa, even of Europe.³⁵

Indeed, he believed that African languages are more regular than European languages, because they are primitive in the sense of not being derived from other languages.³⁶

³⁴ Bonvini, 'Les deux premières grammaires françaises du Wolof', 115.

³⁵ Dard, *Dictionnaire français–wolof*, vi, 'ce dialecte est régulier, uniforme, assujetti à des principes fixes: au milieu de ses imperfections, on découvre des avantages qui n'appartiennent point aux autres langues; enfin il a des caractères propres, et d'autres qui lui sont communs avec les langues de l'Afrique orientale, même de l'Europe'.

³⁶ Dard, *Dictionnaire français–wolof*, xxii.

Other Wolof dictionaries followed.³⁷ The most important Wolof dictionary today is the *Dictionnaire wolof-franais et franais-wolof* of Jean-Lopold Diouf, of the Center of Applied Linguistics of Dakar, published in 2003. It has 10,105 lemmata in the Wolof–French part (370 pages) and 4,647 in the French–Wolof part (177 pages). It is a dictionary of contemporary Wolof. Each entry has a phonetic transcription, classification of the parts of speech, translation equivalents, dialectal variants, etymology, meaning, examples, synonyms, antonyms, and analogy.³⁸

Yoruba

The last two languages of western-central Africa to be discussed in this section are Yoruba and Hausa, both spoken in what is now Nigeria. The first Yoruba dictionary was published in 1843 by a native speaker of the language called Samuel Crowther. He had been born in Osogun, in what is now south-western Nigeria, where his name was Ajayi. In 1821 he was captured as a slave, and in the following year he was purchased by Portuguese slave traders for servitude on the Brazilian plantations, but the ship on which he was to be transported was intercepted by two ships of the British navy, and the slaves were rescued. He became Christian in 1825, taking the name of Samuel Crowther at his baptism, and was ordained for ministry in the Anglican church in 1843, becoming ‘bishop of the countries of Western Africa beyond the Queen’s dominions’ in 1864 – the first African bishop in the Anglican church – and dedicating his life to the Niger mission.³⁹

Crowther’s work on the lexicography of Yoruba appears to have had its origins in his participation, on behalf of the Church Missionary Society, in an expedition in 1841 ‘consisting of three Steamers of the Royal Navy . . . sent up the Niger by her Majesty’s Government’ for a mixture of missionary, diplomatic, and commercial purposes.⁴⁰ A printed outline vocabulary was supplied to the expedition, to be filled in with words from the languages which its members encountered, and Crowther began work on this from his own knowledge as soon as the expedition had begun.⁴¹ The first edition of Crowther’s *Vocabulary of the Yoruba Language* goes far beyond the notes he made on the Niger expedition. It begins with an account of the history of the Yoruba people, from their origin myths to the present day, very much from

³⁷ See, e.g., Calhoun, ‘Reading paratexts’, on a Wolof–French and French–Wolof dictionary of 1855.

³⁸ See Mavoungou, ‘Lexicography of the languages of western Africa’, 966.

³⁹ See Page, *Black Bishop*, and Walls, ‘Crowther’ for his biography.

⁴⁰ Schn and Crowther, *Journals*, i. ⁴¹ Schn and Crowther, *Journals*, 260, 263

a Yoruba point of view, and then presents a 48-page grammar; 176 pages of wordlists, in single columns, divided fairly evenly between English–Yoruba and Yoruba–English; and 20 pages of texts in English and Yoruba.⁴²

The second edition was published nine years later, in 1852. It was sometimes issued with a revised grammar, as *A Grammar and Vocabulary of the Yoruba Language*, but the two were also issued separately.⁴³ The *Grammar and Vocabulary* began with a brief anonymous introduction, which drew attention to the proverbs and idioms collected by Crowther ‘from the lips of his countrymen in the course of common conversation’ and to his use of the ‘system of phonography’ proposed in the ‘Rules for reducing unwritten languages to alphabetical writing in Roman characters, with reference especially to the languages spoken in Africa’ which had been issued by the Church Missionary Society in 1848 (see below).⁴⁴ A much longer preface followed, written by Owen Vidal, the newly consecrated Anglican bishop of Sierra Leone, and notable for its warm appreciation of the Yoruba language, which Vidal presumably knew only from Crowther’s dictionary: like Dard, he commented on the ‘beautiful completeness and perfect regularity which characterize its formative process’; he remarked that ‘the Yoruba is no ordinary language’; and he compared the Yoruba proverbs which Crowther had gathered to the poetry of the Hebrew Scriptures.⁴⁵ Next came an introduction to the grammar; the grammar itself; and, finally, a Yoruba–English wordlist of 287 single-column pages and 3,000 or more entries.⁴⁶ There was no English–Yoruba wordlist.

A third edition then followed, undated, but probably from the period between 1867 and 1872. It was issued with minimal preliminaries: the title page reads simply ‘A vocabulary of the Yoruba language, &c. &c.’, without author or imprint, and a few remarks about the marking of pronunciation and morphology appear on its verso (they had previously appeared in the *Grammar* and the *Vocabulary* of 1852).⁴⁷ There are two wordlists, in double

⁴² Crowther, *Vocabulary of the Yoruba Language* (1843), i–vii (history), 1–48 (grammar), 1–83 (English–Yoruba), 84–176 (Yoruba–English), 177–96 (texts).

⁴³ Crowther, *Grammar of the Yoruba Language*, third unnumbered preliminary page, calls the revised grammar ‘substantially a new work’.

⁴⁴ Crowther, *Grammar and Vocabulary of the Yoruba Language*, iv–v; the ‘Rules’ are reproduced in Spencer, ‘S. W. Koelle and the problem of notation for African languages’, 89–91.

⁴⁵ In Crowther, *Grammar and Vocabulary of the Yoruba Language*, 6, 17, 20.

⁴⁶ Crowther, *Grammar and Vocabulary of the Yoruba Language*, iii–v (introduction), 1–38 (Vidal’s preface), i–vii (introduction to the grammar), 1–52 (grammar), 1–291 (Yoruba–English).

⁴⁷ The anonymous title page and remarks had appeared in Crowther, *Vocabulary of the Yoruba Language* (1852), directly after Vidal’s preface; the remarks had also appeared in his *Grammar of the Yoruba Language*, 4th unnumbered preliminary page.

columns, 144 pages of English–Yoruba and 254 pages of Yoruba–English. This was presumably an advance release of the wordlists that had been printed for a more elaborate third edition which never appeared. Nor was this the end of the career of Crowther’s Yoruba dictionary, for it was drawn on by twentieth-century lexicographers.

Hausa

Hausa has a large mother-tongue speaker population, and is also widely used as a *lingua franca*; as the maker of its first dictionary pointed out, ‘The Hausa is one of the most extensive Languages of Central Africa. An acquaintance with it will open a door of communication with an immense population, and over a vast tract of country.’⁴⁸ This dictionary was the *Vocabulary of the Hausa Language* published in 1843 by the German-born Anglican missionary James Frederick Sch6n, who, according to Mavoungou, ‘can rightfully be regarded as the father of Hausa studies’.⁴⁹ He had, like Crowther, taken part in the Niger expedition, and the two men would remain close. In its course, he ‘directed his attention to the acquisition of the Ibo and Hausa Languages, and has collected extensive Vocabularies and Phrases in both . . . But he has thought proper, for various reasons, to postpone the publication of the Ibo for a future period.’⁵⁰ In the dictionary’s introduction, Sch6n explained that he learned the language by speaking directly with native people, including slaves and kings, in many different places in central Africa, concluding that ‘it is rich in words; and its grammatical structure is easy and beautiful’.⁵¹ He found etymological similarities between Hausa and some European languages, such as ‘Celtic’, English, and ‘Scotch’, and other African languages, like Fula (now classed in the Niger-Congo phylum, and hence seen as completely unrelated to the Afroasiatic Hausa). The book is structured like Crowther’s *Vocabulary of the Yoruba Language*, comprising a grammatical treatise; the vocabularies themselves, English–Hausa (c. 3,200 lemmata) and Hausa–English (c. 3,800 lemmata); collections of phrases, many of them for use in medical consultations; and English and Hausa texts for translation.⁵²

⁴⁸ Sch6n, *Vocabulary of the Hausa Language*, i.

⁴⁹ Mavoungou, ‘Lexicography of the languages of western Africa’, 964.

⁵⁰ Sch6n, *Vocabulary of the Hausa Language*, iii–iv (and cf. Sch6n and Crowther, *Journals*, 116); he did indeed publish on Igbo in years to come, and the title page of the English–Igbo second part of Crowther’s *Vocabulary of the Ibo Language*, published in 1883, states that it was prepared by him.

⁵¹ Sch6n, *Vocabulary of the Hausa Language*, ii.

⁵² Sch6n, *Vocabulary of the Hausa Language*, 1–30 (grammar), 1–102 (English–Hausa), 103–157 (Hausa–English), 158–69 (phrases), 170–90 (texts).

According to Roxana Ma Newman and Paul Newman, ‘in spite of its age, it is surprisingly modern in its approach and grammatical analysis, describing the major grammatical morphemes, pronoun classes, and derivational suffixes’.⁵³ Schön published a second Hausa dictionary (he thought of it as a new work rather than a revised edition of the *Vocabulary*) in 1876, at the instigation of Samuel Crowther.⁵⁴

Two other Hausa bilingual dictionaries with other European languages, namely French and German, should be pointed out.⁵⁵ On the one hand, the *Essai de dictionnaire français–haoussa et haoussa–français*, published in 1886 by the French captain of infantry and head of the Arab office of Bou Saada (Algeria) Jean-Marie Le Roux, is the first dictionary of Hausa written in both Ajami (Arabic script) and Boko (Roman script). However, ‘the fact that many entries are actually Arabic rather than Hausa indicates that it was likely based on the speech of a Hausa-speaker living in North Africa who had an incomplete command of his native language. The work therefore remains more a curiosity in the history of Hausa lexicography than a work of scholarly significance.’⁵⁶ On the other hand, the *Wörterbuch der Hausasprache* by the German missionary Adam Mischlich, published in 1906, has almost 700 pages and 7,000 Hausa lemmata, also written in both Arabic and Roman scripts. Newman and Newman observe that ‘Mischlich’s dictionary reflects a deeper analysis of Hausa grammatical structure and understanding of derivational morphology.’⁵⁷

Eastern Africa

We now turn to the languages of our second main geographical division, namely eastern Africa, beginning with Swahili, which was documented in a polyglot dictionary of 1850, and then turning to Ronga, of which there was a dictionary in 1856; and to Nyungwe, of which a dictionary was completed in 1889. The lexicography by Europeans of the languages of Ethiopia is not treated here: its story begins with learned dictionaries of the early modern period, and is distinct from the main story of the missionary, colonial, and

⁵³ Newman and Newman, ‘Hausa lexicographic tradition’, 267.

⁵⁴ Schön, *Dictionary of the Hausa Language*, ii, quotes a letter of Crowther’s to him, dated 1874: ‘One of the most important helps that we now need is the publication of your Hausa Vocabulary now in manuscript.’

⁵⁵ For others, see Newman, ‘Century and a half’, and Newman and Newman, ‘Hausa lexicographic tradition’, 267–73.

⁵⁶ Newman and Newman, ‘Hausa lexicographic tradition’, 268.

⁵⁷ Newman and Newman, ‘Hausa lexicographic tradition’, 269.

post-colonial lexicography of the languages of Africa. However, one dictionary from Ethiopia does call for brief mention because it was used as a basis for the first dictionary of any of the east African languages: it is the Oromo–English *Vocabulary of the Galla Language*, completed in 1841 by the German missionary Johann Ludwig Krapf, and published in the following year. Krapf was prevented from continuing his missionary work among the Oromo people after 1842, and sailed for the Swahili-speaking sultanate of Zanzibar in the following year.

Swahili

The first Swahili wordlist was published in 1850 as part of Krapf's 64-page *Vocabulary of Six East-African Languages*. This dictionary is laid out in seven columns: the alphabetical English wordlist of the Oromo dictionary is on the left, Oromo equivalents (in a different dialect from the one which Krapf had documented in 1842) are on the right, and between them are equivalents for the English words in five Bantu languages: Swahili, Nyika, Kamba, Kipokomo, and Yao.⁵⁸ However, his main work is *A Dictionary of the Suahili Language*, which was published posthumously in 1882, having been completed in 1860 and circulated in manuscript form since at least 1864.⁵⁹ In its published form, it consists of a 39-page grammar and a 431-page Swahili–English dictionary, both printed in double columns. Krapf describes mainly the Mvita (or Kimvita) variety spoken at Mombasa, on the coast of Kenya, commenting on the other main Swahili variety, the Unguja (or Kiunguja) variety spoken in Zanzibar, which is now part of Tanzania, that

the Kisuahili spoken at Zanzibar has a very large infusion of Arabic and other foreign words. The Mombassians, therefore, consider the dialect of Zanzibar as the ‘manéno ya Kijingajinga’, i.e., the language of ignorant people, or of newly arrived slaves and other foreigners.⁶⁰

Krapf admitted that ‘the Zanzibar dialect was not without usefulness, as it is spoken by a very large number of people along the coast, and also affords to the translator the resource of being able to adopt at will an Arabic word when in difficulty for a proper expression in Kisuahili’.⁶¹ He considered that Swahili should become the standard language of evangelization in east Africa, and

⁵⁸ For its compilation, see Krapf, *Vocabulary of Six East-African Languages*, iii–x.

⁵⁹ See Polomé, ‘Lexicography of the Niger-Kordofanian languages’, 2647; Steere, *Handbook of the Swahili Language* (1870), iv.

⁶⁰ Krapf, *Dictionary of the Suahili Language*, xi.

⁶¹ Krapf, *Dictionary of the Suahili Language*, xii; cf. Steere, *Handbook of the Swahili Language* (1870), iii.

should be taught, at least as ‘their literary language’, to speakers of other languages: ‘as the Kisuhili is the most cultivated of the dialects in this part of Africa, and is, moreover, spoken from the equator southwards to the Portuguese settlements of Mozambique, it should be made to supersede, as much as possible, the minor dialects inland which are spoken by only a small population’.⁶² Krapf used the Roman alphabet, adapting the *Standard Alphabet* of Lepsius (see below), but said, in the published dictionary of 1882, that he regretted

not having chosen the Amharic Alphabet for the great family of languages to the south of the Equator. As I was the first European who reduced Suahili to writing, and as there was then no universal alphabet compiled, I might easily have chosen . . . the Amharic character, which would evidently suit the Suahili better than the Roman . . . However . . . I have never regretted having rejected the Arabic mode of writing, which is too imperfect and too ambiguous for writing Suahili in a correct manner.⁶³

For Polomé, ‘what distinguishes this work from its predecessors is the deliberate effort of its author to provide detailed information on the usage of the terms, their background and their sociocultural context’.⁶⁴ Indeed, each entry of the dictionary has conversational examples, and sociological and anthropological explanations.

The pioneer lexicographical study of the Unguja variety of Swahili spoken at Zanzibar, *A Handbook of the Swahili Language as spoken at Zanzibar* (1870), was written by the Englishman Edward Steere, who had done missionary work in east Africa in the 1860s, and was consecrated as the Anglican bishop of central Africa in 1874.⁶⁵ The book has two main parts: a grammar of 225 pages and a Swahili–English dictionary of 148, followed by an appendix of texts and phrases. Steere’s work was continued by Arthur Cornwallis Madan, who had left a teaching position at Christ Church, Oxford, to work as a lay missionary in east Africa. He published revised and enlarged versions of Steere’s *Handbook* from 1884 onwards, as well as his own *English–Swahili Dictionary* (1894) and *Swahili–English Dictionary* (1903). Doke assessed Madan’s *English–Swahili Dictionary* as ‘an admirable work dealing with the Zanzibar form’, which ‘has been of inestimable value for many years’.⁶⁶

⁶² Krapf, *Dictionary of the Suahili Language*, xi.

⁶³ Krapf, *Dictionary of the Suahili Language*, xiv.

⁶⁴ Polomé, ‘Lexicography of the Niger-Kordofanian languages’, 2647.

⁶⁵ Steere, *Handbook of the Swahili Language* (1870), iv–x, sets out the background of his work, including his debt to Krapf.

⁶⁶ Doke, *Bantu*, 57.

There have been other bilingual dictionaries of Swahili, including some in German and French.⁶⁷ The first monolingual dictionary of an African vernacular language must be pointed out, the *Kamusi ya kiswahili yaani kitabu cha maneno ya kiswahili* ('Swahili–Swahili dictionary') of 1935, by Frederick Johnson, who published various revised versions of Madan's dictionaries. 'This', according to Doke, 'is a veritable milestone in Bantu lexicographical studies.'⁶⁸ According to Benson, Johnson published the *Kamusi ya kiswahili* for 'smoothing the way . . . for the non-Swahili adult literate eager to understand such Swahili literature as came his way'.⁶⁹

Ronga

Like that of Swahili, the first wordlist of Ronga, a language of what is now Mozambique, appeared in a polyglot work, *The Languages of Mosambique*, published in 1856 by the German comparative linguist Wilhelm Heinrich Immanuel Bleek, known as the 'father of Bantu philology'. It was founded, as Bleek made clear, on manuscript wordlists by the explorer and naturalist Wilhelm Karl Hartwig Peters, who had written them in a copy of the outline vocabulary printed for the Niger expedition (see above). Bleek also – but 'only at a late stage in his own work' – knew an earlier version of *Polyglotta Africana* (1854) by the German-born missionary Sigismund Wilhelm Koelle, which was advertised on its title page as 'a comparative vocabulary of nearly three hundred words and phrases, in more than one hundred distinct African languages'.⁷⁰ Polomé remarks, by the way, that the *Polyglotta Africana* was an 'important milestone in Niger-Kordofanian lexicography', but adds that 'whatever its merits and the value of the information it provides, it is, however, no substitute for a regular dictionary'.⁷¹

Over its 399 pages, *The Languages of Mosambique* presented 1,742 English words; their equivalents in nine languages, written in parallel columns, in a modified version of the alphabet of Lepsius; and a column of miscellaneous observations.⁷² The nine varieties named on Bleek's title page correspond to modern languages still spoken in Mozambique: 'the Dialects of Lourenço Marques, Inhambane, Sofala, Tette, Sena, Quelimane, Mosambique, Cape

⁶⁷ See Doke, *Bantu*, 55–64, and van Spaandonck, *Practical and Systematical Swahili Bibliography*.

⁶⁸ Doke, *Bantu*, 64. Unfortunately, I have not been able to study the text of this book directly.

⁶⁹ Benson, 'Century of Bantu lexicography', 72.

⁷⁰ Janson, 'Languages and language names', 308.

⁷¹ Polomé, 'Lexicography of the Niger-Kordofanian languages', 2647.

⁷² For the use of Lepsius, see Bleek, *Languages of Mosambique*, xvii.

Delgado, Anjoane, The Maravi, [and] Mudsau' are now called, respectively, Ronga, Gitonga, Ndaou or Cindau (for which see below), Nyungwe (for which see below), Sena or Cisená, Echuwabo, Emakhuwa or Makuwa, Kimwani, and Cinyanja.⁷³

The *Grammaire ronga* (1896) by Henri-Alexandre Junod has a supplementary Ronga–Portuguese–French–English vocabulary.⁷⁴ However, the free-standing lexicographical treatments of Ronga are fairly recent, such as the *Dicionários Xironga* by the secular priest José Luís Quintão (1951). Noteworthy among them is the *Dicionário Ronga-Português*, published in 1960 by a layman, the Portuguese linguist and phoneticist Rodrigo de Sá Nogueira. His main purpose was to apply, at the same time, a scientific method in the description of Ronga, based on work with native informants, and a didactic plan, in order to be clearer for the readers than earlier dictionaries.⁷⁵ That is why he did not adopt entirely the alphabet of Lepsius. Instead of a purely phonetic system of transcription, Nogueira chose a system which used dashes and the apostrophe to divide the semic units and the various morphological elements of the word, such as the roots, prefixes, and suffixes.⁷⁶ He said explicitly that phonetic transcription should be used only by phoneticists and that it, as an orthographic system for everyday use by all people, is absolutely unacceptable.⁷⁷ A striking feature of his dictionary is the suggestion that the Bantu and Indo-European languages shared a common origin. For him, this was only a suspicion, for which he saw certain lexical similarities as evidence, but always cautiously, in a hypothetical way. His hope was that one day, in the future, someone could prove the point.⁷⁸

Nyungwe

One of the earliest lexicographical descriptions of any Mozambican language was written by a French missionary at the service of the *Padroado real*, Victor-Joseph Courtois, Father Superior of the Bembe Mission and a primary teacher for the native children from 1882 until his death in 1894. His Portuguese–Nyungwe and Nyungwe–Portuguese dictionaries were published a few years after his death by the University of Coimbra, in 1899 and 1900 respectively, but they were finished and prepared for the printer in 1889.⁷⁹ His main

⁷³ Janson, 'Languages and language names', 298–307; see also Sitoe, *Lexicografia da língua Tsonga*, and Fortune, '75 years of writing in Shona'.

⁷⁴ Junod, *Grammaire ronga*, 65–90. ⁷⁵ Nogueira, *Dicionário Ronga-Português*, vii.

⁷⁶ Nogueira, *Dicionário Ronga-Português*, xv. ⁷⁷ Nogueira, *Dicionário Ronga-Português*, xvi.

⁷⁸ Nogueira, *Dicionário Ronga-Português*, x–xi.

⁷⁹ Courtois, *Diccionario Portuguez-Cafre-Tetense*, 484; Courtois, *Diccionario Cafre-Tetense-Portuguez*, v.

dictionary, the Portuguese–Nyungwe, which runs to 424 pages, is a translation of the Portuguese dictionary by Jos  da Fonseca, improved by Jos  Ignacio Roquete, first published in Paris in 1848. When he had some difficulties in the translation of a Portuguese word into Nyungwe, he used periphrasis in order to give a sense of its meaning.⁸⁰ The Nyungwe–Portuguese dictionary is a shorter complement to the Portuguese–Nyungwe dictionary (it has only eighty-one pages) and a kind of *vade mecum* for those who are beginning to study the language; as Courtois said, no similar work existed.⁸¹

Southern Africa

The languages of southern Africa treated here are the Shona languages, one of which was documented in 1856 by Bleek (see above); Zulu, of which the first dictionary was published in 1861; and Xhosa, of which the first significant dictionary appeared in 1872.

The Shona Languages

The first printed wordlist of a Shona language was that of Ndau in Bleek’s *Languages of Mosambique* (see above). Emmanuel Chabata has observed that, in this respect, the title of Bleek’s work ‘is misleading since its vocabulary [of Ndau] was drawn from the Sofala dialect area, along the Pungwe–Save rivers of Zimbabwe’.⁸² Bleek does not give much information concerning this wordlist or the language it documents, saying only that it was ‘taken from a fairly written manuscript, which, I believe, was drawn up by an old Chinese or Indian, who had spent the greater part of his life in that country. He also gave Dr Peters the copy of a description of the interior in Portuguese verses, composed by himself’.⁸³

The first dictionary dedicated to Shona was the *English–Mashona Dictionary* published by the Austrian-born Jesuit missionary Andrew M. Hartmann in 1894. Father Hartmann was chaplain to the Pioneer Column of the British South Africa Company for the British government, sent ‘into the indeterminate borderlands between the Ndebele and the Shona in mid-1890’, and served at a mission at Empandeni, in what is now Zimbabwe.⁸⁴ George Glynn Fortune, the greatest Shona expert in modern times, stated that Father

⁸⁰ Courtois, *Diccionario Portuguez–Cafre-Tetense*, v.

⁸¹ Courtois, *Diccionario Cafre-Tetense–Portuguez*, v. ⁸² Chabata, ‘Lexicography of Shona’, 948.

⁸³ Bleek, *Languages of Mosambique*, xi–xii.

⁸⁴ For the Pioneer Column, see Marks, ‘Southern and central Africa’, 445.

Hartmann described Zezuru, a Shona language variety spoken in an area around Salisbury (now Harare).⁸⁵ He added that Hartmann ‘notes dialectal differences but calls them verbal, not grammatical . . . His spelling seems strange to us today as he confuses many voiceless and voiced phonemes . . . thus betraying the influence of his German background. But his ear in other respects was remarkably acute . . . As a result, he used a system of word division remarkably close to our present-day practice.’⁸⁶

The two major African languages spoken in the area which is now Zimbabwe are Shona and Ndebele, and so the English missionary William Allan Elliott published a 441-page English–Ndebele–Shona dictionary in 1897, as *Dictionary of the Tebele & Shuna Languages*. Elliott said that Ndebele was ‘of course only a variety of the Zulu’ (of which there were already dictionaries: see below) and not a language itself, but that Shona was ‘practically an unknown tongue’.⁸⁷ The Shona portion of the *Dictionary* was based mainly on Hartmann’s *Dictionary*. Elliot adopted, however, a different orthography ‘in some respects from that used by Fr Hartmann’, explaining that ‘after fourteen years’ residence in Matebele Land, I had fixed my spelling before his was printed, and I see no reason to make alteration now’.⁸⁸ He added that he had tried ‘to present a written basis for the Shuna language as a whole, from which the peculiarities of the different dialects may be observed’.⁸⁹

Further dictionaries of Shona followed.⁹⁰ In 1996, Herbert Chimhundu published the first monolingual Shona dictionary, *Duramazwi ReChiShona* (‘Dictionary of Shona’) (DRC), as a consequence of the ALLEX Project and the ALRI (see above). It ‘is a synchronic, general purpose, medium-sized, monolingual dictionary mainly meant to cater for the needs of lower secondary school learners’.⁹¹ It has a 34-page introduction and 504 pages of entries, and gives phonological and grammatical information, dialectal variants, and synonyms. An enlarged and extended edition was published in 2011, *Duramazwi Guru ReChiShona* (‘Advanced dictionary of Shona’) (DGC). Chabata concludes that

DGC is more advanced . . . in terms of its size and the presentation of its meanings. The headword and sense selection for this dictionary was more

⁸⁵ Fortune, ‘75 years of writing in Shona’, 57.

⁸⁶ Fortune, ‘75 years of writing in Shona’, 57–8.

⁸⁷ Elliott, *Dictionary of the Tebele & Shuna Languages*, v.

⁸⁸ Elliott, *Dictionary of the Tebele & Shuna Languages*, vi.

⁸⁹ Elliott, *Dictionary of the Tebele & Shuna Languages*, vi.

⁹⁰ For other dictionaries of Shona, see Doke, *Bantu*.

⁹¹ Chabata, ‘Lexicography of Shona’, 949; unfortunately, I was unable to study either of Chimhundu’s dictionaries directly.

comprehensive for it includes language used in all spheres of life. DGC is also the first dictionary in the history of Shona lexicography to include phrasal headwords such as proverbs, idioms and pithy sayings. The improvement . . . can be accounted for by the use of the Shona corpus as well as the experience the editors gained from having worked on DRC.⁹²

Zulu

The first substantial Zulu–English and English–Zulu dictionaries were published by James Perrin in 1855, in London and Pietermaritzburg respectively.⁹³ According to Eric Hermanson, both of Perrin’s dictionaries were edited under the supervision of John William Colenso, the first Anglican bishop of Natal, who was also the author of the ‘advertisement’ to each.⁹⁴ Colenso said that ‘in the compilation of this, and of the Kafir–English Dictionary, Mr Perrin has derived considerable assistance from the Vocabularies prepared by some of the American Missionaries, to which they very kindly gave him free access.’⁹⁵ The Kafir–English is shorter (166 pages) than the English–Kafir dictionary (225 pages), but it has very interesting clarifications. In just two pages, Perrin explained some of the basics of Zulu phonetics, the orthography, and the method used in the presentation of the lemmata. For example, ‘The root of the word is always placed first. After it the incipient particle or prefix. As the sign of the Infinitive is not given in English, it is not expressed in Zulu; thus, HLA, “eat,” instead of UKUHLA, “to eat”’.⁹⁶

Colenso published his *Zulu–English Dictionary* in 1861. It is much bigger than Perrin’s dictionaries, with 548 pages of entries. He starts the ‘advertisement’ clarifying the points that ‘should be noticed by the Student’. Among these, he pointed out the focus of the dictionary on ‘pure Zulu’ (without words from neighbouring languages such as Xhosa); its exclusion of regional differences within Zulu; its marking of words which ‘have been formed by corruption from the English or Dutch languages, or have been coined by Missionaries’; and its orthographic treatment of Zulu sounds.⁹⁷ More than a century later, Colenso’s dictionary could still be described as ‘very sound’, although of course ‘now out of date’.⁹⁸

⁹² Chabata, ‘Advanced dictionary’, 115.

⁹³ For earlier vocabularies of Zulu, see Gauton, ‘Lexicography of the Nguni languages’, 912. Polomé, ‘Lexicography of the Niger-Kordofanian languages’, 2648, saw the Zulu tradition as beginning only with Colenso in 1861.

⁹⁴ Hermanson, ‘Colenso’s first attempt’, 17–18. ⁹⁵ Perrin, *English–Kafir Dictionary*, [iii].

⁹⁶ Perrin, *Kafir–English Dictionary*, iv. ⁹⁷ Colenso, *Zulu–English Dictionary*, iii.

⁹⁸ Doke and Vilakazi, *Zulu–English Dictionary*, 2nd edn, vii.

In 1905, Alfred Thomas Bryant, a British Roman Catholic secular priest, who taught at the first Catholic mission in what was then Zululand, published an important *Zulu–English Dictionary*, with prefatory material on Zulu history, comparative philology, and grammar; a wordlist of approximately 22,000 lemmata with supplements such as ‘a vocabulary of the *hlonipa* language of the Zulu women’; and ‘an appendix containing additional words, improvements, corrections, etc.’.⁹⁹ His *magnum opus* signified ‘a real advance on previous work in several directions. Words were arranged alphabetically according to the stem, and verbal derivatives were reduced to those which had some special significance beyond the normal. What was most important was the inclusion of aids to pronunciation’.¹⁰⁰ Doke called it an ‘outstanding work’.¹⁰¹

A number of Zulu dictionaries have followed.¹⁰² The most important Zulu–English dictionary at present is the *Zulu–English Dictionary* published in 1948 by Clement Doke and the Zulu writer and educator Benedict Wallet Vilakazi. It was re-edited in the following year, revised in 1953, and re-edited many times thereafter. Polomé calls it ‘excellent’, and Benson adds that it ‘contains a very full body of citations, tone patterns for every word, examples of rare usage, details of etymology, and full treatment of verbal extensions’.¹⁰³ In 1958, Doke, Daniel McKinnon Malcolm, and Jonathan Mandlenkosi A. Sikhakhana edited the *English–Zulu Dictionary*, as a companion to the *Zulu–English Dictionary*. Since 1990, they have been published together. They are, indeed, a standard lexicographical work made by leading scholars and Zulu linguists. They are intended for Zulu native speakers, Zulu-speakers who want to study English, and linguists who study Zulu.

The first monolingual Zulu dictionary, titled *Inqolobane yesizwe*, was published in 1966 by Cyril Sibusiso Nyembezi and Otty Nxumalo. Mark Sanders describes this dictionary as an ‘invaluable compendium of Zulu vocabulary, figures of speech, and proverbs’.¹⁰⁴

⁹⁹ Bryant, *Zulu–English Dictionary*, 12*–66* (history), 67*–82* (comparative philology), 83*–108* (grammar), 1–737 (main wordlist), 738–61 (supplementary wordlists), 762–78 (addenda).

¹⁰⁰ Doke, *Bantu*, 75–6.

¹⁰¹ Doke and Vilakazi, *Zulu–English Dictionary*, 2nd edn, vii; see also Polomé, ‘Lexicography of the Niger-Kordofanian languages’, 2648.

¹⁰² See Gauton, ‘Lexicography of the Nguni languages’, 914, and Doke, *Bantu*, 74–81.

¹⁰³ Polomé, ‘Lexicography of the Niger-Kordofanian languages’, 2648; Benson, ‘Century of Bantu lexicography’, 75.

¹⁰⁴ Sanders, *Learning Zulu*, 121; I was unable to study this dictionary directly.

Xhosa

For Doke, ‘the earliest real dictionary of Xhosa’, published in 1872, was *A Dictionary of the Kaffir Language, Including the Xosa and Zulu Dialects*, compiled by the Wesleyan Methodist missionary William Davis.¹⁰⁵ The second part, the *English–Kaffir Dictionary*, was published five years later. Although Davis referred to a ‘Kaffir language’ and ‘Xosa and Zulu dialects’ in his title (*Kaffir*, ultimately from an Arabic word meaning ‘unbeliever’, had come to be used of the Nguni peoples; in an entry of 2016, *OED* labels it as ‘Now *hist.* and *offensive*’), he perfectly distinguished both languages as ‘the XOSA KAFFIR, spoken by the Amaxosa tribes, who live in Kaffirland beyond the Eastern Frontier of the Cape Colony’ and ‘the ZULU KAFFIR spoken by the Zulu tribes in the Natal Colony and the country on its borders’.¹⁰⁶ Nevertheless, he saw that they are closely related:

They have the same grammatical constructions, and a large majority of the words are the same both in form and meaning. But many words which are the same *form* in both languages differ in *signification*, and others are different both in form and meaning; and yet in many instances in which this diversity obtains, the *original root* from which these words are derived is evidently one and the same; and in the case of others, the *root* will be found in one language and the derivatives in the other.¹⁰⁷

The *Dictionary of the Kaffir Language* is organized etymologically by the roots of Xhosa (or Zulu) words, the prefixes being placed before the root. According to Doke, ‘meanings are carefully discussed and numerous illustrative sentences are included’.¹⁰⁸

In 1899, the German-born Lutheran missionary Johann Heinrich Albert Kropf (not to be confused with Johann Ludwig Krapf, the lexicographer of Swahili) published *A Kaffir–English Dictionary*, which has been called ‘masterly’ and ‘one of the best bilingual dictionaries in isiXhosa’.¹⁰⁹ Kropf’s dictionary is organized into two columns, has almost 500 pages, and was a work of almost all his lifetime, at least since 1845 when he arrived in Xhosa-speaking southern Africa.¹¹⁰ He did not follow the *Standard Alphabet* by Lepsius, deciding with some reluctance that he should favour the orthography of the existing translation of the Bible into Xhosa.¹¹¹ Kropf encountered lexicographical challenges, and his handling of certain prefixed forms has

¹⁰⁵ Doke, *Bantu*, 83; for earlier lists of Xhosa words, see Mtuzi, ‘Critical survey’, 167.

¹⁰⁶ Davis, *Dictionary of the Kaffir Language*, v.

¹⁰⁷ Davis, *Dictionary of the Kaffir Language*, v. ¹⁰⁸ Doke, *Bantu*, 83.

¹⁰⁹ Doke, *Bantu*, 85; Moropa and Kruger, ‘Mistranslation’, 73.

¹¹⁰ Kropf, *Kaffir–English Dictionary*, iii. ¹¹¹ See Kropf, *Kaffir–English Dictionary*, v.

been criticized.¹¹² Nevertheless, for the editors of *The Greater Dictionary of Xhosa*, 'Kropf's dictionary is a masterly and scholarly work that has stood the test of more than three-quarters of a century. Kropf appears to be the first lexicographer to have fathomed and thoroughly mastered the intricacies of Xhosa phonology, including the distinction between radical, aspirated and ejective sounds.'¹¹³

In 1968, the University of Fort Hare, in Alice, South Africa, created the Xhosa Dictionary Project (now IsiXhosa National Lexicography Unit), under the direction of Herbert Walter Pahl, which culminated with the publication of the Xhosa–English–Afrikaans *Greater Dictionary of Xhosa*. The volumes appeared in reverse order: in 1989, the third volume was published, comprising the letters Q–Z; the second, in 2003 (K–P); and the first, in 2006 (A–J). A successor of Pahl's as editor-in-chief of the dictionary, Peter Tshobisa Mtuze, observes that

The Greater Dictionary of Xhosa does not confine itself to . . . the dialect first reduced to writing by the missionaries. It also includes other dialectal and regional connotations as well as *hlonipha* (language of respect used by married women and the newly initiated boys). Many other variations in language usage are accommodated in the dictionary, some bordering on what could be stigmatized as colloquialism. In this way, the dictionary tries not to be prescriptive but to be as descriptive as possible.¹¹⁴

The use of Xhosa as the lemmatizing language of the dictionary 'was a major break from the tradition whereby all information about the language was hitherto given through another language, implying that those who did not know such a language, i.e. English or Afrikaans, could not benefit'.¹¹⁵ Since its publication, a monolingual dictionary, *Isichazi-magama SesiXhosa* (2008), has appeared.

Orthography

A theme which has run through this chapter is the problem of orthography.¹¹⁶ The first lexicographers and many of their successors used their own systems, based mainly on the orthography of each lexicographer's mother tongue. In 1848, the Church Missionary Society published 'Rules for reducing unwritten languages to alphabetical writing in Roman characters, with reference

¹¹² Mtuze, 'Critical survey', 170. ¹¹³ Pahl et al., *Greater Dictionary of Xhosa*, III.xxxviii.

¹¹⁴ Mtuze, 'Critical survey', 170. ¹¹⁵ Nkomo and Wababa, 'IsiXhosa lexicography', 356.

¹¹⁶ For an more detailed overview, see Bendor-Samuel, 'Adaptations'.

especially to the languages spoken in Africa', which, as we have seen, were used by Crowther in his *Grammar and Vocabulary of the Yoruba Language* of 1852. In 1855, they replaced this with a *Standard Alphabet for Reducing Unwritten Languages and Foreign Graphic Systems to a Uniform Orthography in European Letters, translated from the Allgemeine linguistische Alphabet of 1854* by Karl Richard Lepsius, who had first created it to transcribe the Egyptian hieroglyphs and then adapted it to living African languages, mixing morphologic and phonetic principles. We have seen responses to the alphabet of Lepsius in Bleek's *The Languages of Mosambique* of 1856; in Krapf's Swahili dictionary, published in 1882; in Kropf's Xhosa dictionary of 1899; and in Nogueira's Ronga dictionary of 1960. Finally, in 1927, the International Institute of African Languages and Cultures, founded in London in the previous year, published a *Practical Orthography of African Languages*, revised in 1930; I have noted its use in Alves' Kimbundu dictionary of 1951. Currently, the orthographical solutions are various: a given lexicographer may use his or her own orthographical system, or that of the institution to which they belong (as in the case of the IsiXhosa National Lexicography Unit), or a national system (as in the case of the *SiSwati Orthography* of the South African Department of Bantu Education).¹¹⁷

Conclusions

African lexicography started during the first centuries of colonization. It was mostly undertaken by European missionaries, both Catholic and Protestant, to teach other missionaries how to speak with the indigenous people and to teach the language of the colonizers to them. The metalanguages were those of the colonizers: mainly Portuguese, English, French, and German. American missionaries made their own contributions (for instance the Umbundu dictionary published by Sanders and Fay, and the wordlists on which Perrin's Zulu lexicography was based), as other travellers in Africa may have done, like the 'old Chinese or Indian' whose records of Ndaue were used by Bleek. Black African lexicographers from Manuel de Roboredo studied African languages: I have noted the names of Crowther and Nlemvo in the nineteenth century, Vilakazi in the mid twentieth, and Chimhundu, Mtuze, Nxumalo, and Nyembezi more recently.

¹¹⁷ Nkomo and Wababa, 'IsiXhosa Lexicography', 356; Department of Bantu Education, *SiSwati Orthography*.

The oldest Bantu vocabulary is still a manuscript waiting for – and deserving – a critical edition. Many other vocabularies or dictionaries were circulated as manuscripts among the missionaries, and they are lost. The first Bantu printed dictionary was published only in 1804, and few dictionaries or vocabularies were published earlier than 1860, the majority having been published after the Berlin Conference, and indeed mainly in the twentieth century. Monolingual dictionaries have been published only in recent years: Johnson's Swahili *Kamusi ya kiswahili yaani kitabu cha maneno ya kiswahili* (London, 1935); Nyembezi and Nxumalo's Zulu *Inqolobane yesizwe* (Pietermaritzburg, 1966); Chimhundu's Shona *Duramazwi ReChiShona* (Harare, 1996); and Tshabe, Guzana, and Nokele's Xhosa *IsiChazi-magama sesiXhosa* (Pietermaritzburg, 2008).

Finally, the role of three scholars, who are truly landmarks in the field, should be emphasized in the development of African lexicography: Wilhelm Bleek, the 'father of Bantu philology'; Clement Doke, the 'father of the history of Bantu languages'; and, not for exclusively lexicographical reasons, Samuel Ajayi Crowther, who was captured as a slave when he was thirteen years old, purchased, and rescued, and became the first African bishop in the Church of England, and the author of the first Yoruba dictionary.

Missionary and Other Traditions in Australia

WILLIAM B. MCGREGOR

This chapter sets out in broad brush some of the main trends in lexicographic work on the indigenous languages of mainland Australia and nearby islands – some of the major ones being Torres Strait Islands, Bathurst Island, Melville Island, Groote Eylandt, Mornington Island, and Bentinck Island – by missionaries, from the early days of contact with Europeans to the present. Prior to contact with Europeans, some 400 languages were indigenous to this region, depending on the criteria one adopts for languagehood; the twenty-first edition of *Ethnologue* lists just under 400 languages on an assortment of criteria, while R. M. W. Dixon recognizes around 250 languages based on the linguistic criterion of mutual intelligibility.¹ These belong to twenty or thirty different families which have not to date been shown to be related to one another. Since European contact, a number of other languages have come to be spoken in Australia. These non-indigenous languages are excluded from the present chapter, except for a small number of post-contact varieties, mainly pidgins and creoles, that are currently spoken by indigenes of Australia.²

Introduction

Missionaries arguably made the most significant contributions to the lexicography of indigenous languages in Australia and the Pacific until modern times. In overall terms, perhaps, they made more substantial contributions to lexicography than to grammatical description of the languages – though they also made important contributions to grammatical description.³ In a number of regions, their work began fairly soon after initial contact with the

¹ Dixon, *Australian Languages*, 5.

² For overviews of Australian languages, see Dixon, *Languages of Australia*, and Dixon, *Australian Languages*.

³ Stockigt, 'Pama-Nyungan morphosyntax'.

indigenes. But missionaries were not always the first to undertake lexicographical and other studies of indigenous languages. Explorers sometimes gathered short wordlists from the peoples they encountered. Thus Captain James Cook and his crew compiled wordlists in a few places where they had more or less protracted contact with indigenes. For instance, Cook and Joseph Banks compiled wordlists in Guugu Yimidhirr while their ship was undergoing repairs at the Endeavour River in northern Queensland from 17 June to 4 August 1770.⁴

Almost a century before Cook and Banks, the privateer William Dampier recorded a single word of an Aboriginal language during a brief sojourn somewhere on the Dampier peninsula in 1688 while he was careening his ship. He relates that some indigenes approached the ship threateningly, and the ship's drum was sounded, at which they 'ran away as fast as they could drive, crying "Gurri, gurri" deep in the throat'.⁵ The word cited is almost certainly the Bardi word *ngaarri*, a term for a malevolent spirit.

Some early colonists and settlers also recorded wordlists of local languages. In the early years of the British penal colony at Port Jackson, wordlists of the local language Iyura (Eora) were compiled by officers. These included Watkin Tench and William Dawes, who also gathered words from nearby people during an expedition in 1791, thus revealing Australia as a multilingual continent.⁶ Settlers and others who had significant contact with Aborigines – including protectors of Aborigines, policemen, postmen, squatters, teachers, and others – continued collecting wordlists throughout the nineteenth and early twentieth centuries, as Australia was progressively colonized. In many instances, words of the indigenous languages were written according to the English spelling system, though some did use a 'uniform orthography' such as the one recommended by the Royal Geographical Society in 1885, in which each letter and digraph has ideally a distinct and consistent phonetic value.⁷ These wordlists were generally quite short, and usually provided single-word translations of English glosses. There were a few exceptions, however. In the nineteenth century, George Fletcher Moore compiled *A Descriptive Vocabulary of the Language in Common Use amongst the Aborigines of Western Australia* (1842), a dictionary of Noongar (Nyungar) varieties amounting to almost 2,000 entries, during his residence in the region of Perth from 1830 to 1841. Moore's dictionary built on earlier

⁴ See further Haviland, 'A last look'.

⁵ Quoted in O'Grady, 'Lexicographic research in Aboriginal Australia', 782.

⁶ Wilkins and Nash, 'European "discovery" of a multilingual Australia'.

⁷ 'System of orthography for native names of places'.

work by the settler Robert M. Lyon, comprising around 500 words organized into semantic domains, and by the explorer George Grey.⁸

In the late nineteenth and early twentieth centuries, a number of amateur anthropologists and linguists began to record information on Aboriginal people, customs, kinship systems, and languages. One of the most notable of these was Edward Micklethwaite Curr, who sent out a standard list of some 125 English glosses to police, magistrates, squatters, missionaries, and others, residing in all parts of the continent. The resulting 300 or so wordlists (including a range of varieties, many of which are no longer spoken) were published in his four-volume work *The Australian Race*. A little later, in 1904, Daisy Bates was appointed by the Western Australian government to collect vocabularies of the languages of the state; she continued until 1912, by which time her project had extended far beyond mere gathering of vocabularies. Bates sent out some 500 questionnaires to squatters, policemen, postmen, missionaries, and the like. Bates' questionnaire was somewhat more comprehensive than Curr's – her intention was to obtain sufficient information to permit some appreciation of the structure of the languages – and made up some 97 printed pages (including space for information) and comprised up to about 2,000 English glosses.⁹ Bates herself provided information on a number of languages she had first-hand contact with, including a number of Kimberley languages she encountered at the turn of the twentieth century. Not unexpectedly, most of the questionnaires were very incompletely filled in. The vocabularies have never been published, though much of the material is now available online.¹⁰ These vocabularies represent an important source of information on the languages of Western Australia, including a number of now extinct or highly endangered languages.

Also working around the turn of the twentieth century was the surveyor R. H. Mathews, who produced descriptions of a considerable number of languages mainly of south-eastern Australia. Mathews' descriptions followed a fairly consistent scheme, and generally included sections covering the orthography, the grammar, and a vocabulary of between 100 and 500 words organized according to semantic domain.¹¹

Prior to the turn of the twentieth century, few if any professionals contributed significantly to the lexicography of Australian languages. The

⁸ Lyon, 'Glance at the manners and language'; Grey, *Vocabulary*; see Moore, 'History of the 1842 *Descriptive Vocabulary*'.

⁹ See McGregor, 'Daisy Bates' documentations' for a detailed description.

¹⁰ See Thieberger, 'Daisy Bates in the digital world', and Thieberger (ed.), 'Digital Daisy Bates'.

¹¹ Koch, 'R. H. Mathews' schema', 181–3.

only professional to do primary research in the nineteenth century was the American linguist Horatio Hale, who visited Australia in 1839–40 as philologist for the United States Exploring Expedition. His report contains grammatical descriptions of two languages of New South Wales, Awabakal and Wiradjuri, based largely on missionary sources; just a short comparative wordlist in the two languages is included, comprising just over 200 common words, organized alphabetically on English glosses.¹² In the early years of the twentieth century the German anthropologist Hermann Klaatsch and the little-known Swedish anthropologist Yngve Laurell visited Australia and recorded some words in Aboriginal languages they encountered. Their wordlists – fairly extensive in the case of Klaatsch – were not published, but represent important lexicographic resources on now moribund languages such as Nyulnyul.

Subsequently, in the period from about 1910 to the early 1950s, professional anthropologists, including Alfred R. Radcliffe-Brown, Ursula H. McConnel, Adolphus P. E. Elkin, Norman B. Tindale, and Phyllis M. Kaberry, collected wordlists in the languages of the people they worked with, usually relatively short and typically with single-word English glosses. A number of these anthropologists had had some training in phonetics (Elkin in London, McConnel in London and at Yale, Tindale in Adelaide), and employed the IPA or some variant of it in their transcriptions of words in the languages. Tindale in particular is remarkable for the number of languages from across the continent, many now moribund, from which he gathered wordlists, in the course of anthropological expeditions in 1938–9 and 1952–4.¹³

From about 1930 the field of Aboriginal linguistics saw increasing professionalization, the appearance of better-trained investigators, and the first academic linguists. Among the academic linguists in the first wave (c. 1930–60) were Arthur Capell, Gerhardt Laves, and T. G. H. Strehlow. Beginning in the late 1920s, these linguists gathered much lexical material in a range of languages – Laves in half a dozen languages of Dampier Land, Daly River, northern New South Wales, and southern Western Australia; Capell in many languages of Arnhem Land, the Kimberley, and other regions; and Strehlow primarily in Western Arnhem. None of these men published substantial dictionaries of any language they worked on, though T. G. H. Strehlow worked for years on a dictionary of Western Arnhem,

¹² Hale, *Ethnography and Philology*, 505–10 (and cf. 479–81).

¹³ See Monaghan, 'Laying down the country', Monaghan, 'Norman B. Tindale and the Pitjantjatjara language', and Tindale, 'Compiled vocabularies'.

and there were expectations that it would be a huge work.¹⁴ It was not until the 1980s that academically trained linguists began to make a significant contribution to the lexicography of indigenous languages, and to publish substantial dictionaries, a notable example being Jeffrey Heath's *Nunggubuyu Dictionary*. Indeed, it was not until the 1990s that a significant number of large dictionaries began to appear. From the 1930s, non-academic linguists – including many missionaries – were also better trained in linguistics, undertaking studies in universities and courses run (from the early 1950s) by SIL International (formerly known as the Summer Institute of Linguistics).

Little previous work has been done on the history of lexicography in Australia, this being a somewhat underdeveloped domain in the history of Australian linguistics. Indeed, serious interest by Australianists in the history of Aboriginal linguistics is relatively recent, and is only a couple of decades old. This interest was to a considerable extent sparked by the realization that legacy materials provide significant information on many now moribund languages for which it is now impossible to gather much primary data, and that these materials need to be interpreted in relation to their times.¹⁵

Geoffrey O'Grady's book chapter 'Lexicographic research in Aboriginal Australia' deals with work from first European contact up to the late 1960s, discussing the content and organization of some representative works, and identifying aspects of the languages (phonological, semantic, and grammatical) that have given rise to difficulties for lexicographers. O'Grady draws a distinction between wordlists, comprising fewer than 1,000 lemmata, and dictionaries, comprising more than 1,000 lexical entries. He observes that only eight dictionaries of languages of Australia and the Torres Strait Islands had been published as of 1968, and that there were, to his knowledge, a further forty-eight unpublished dictionaries, all dating from the twentieth century.

Cliff Goddard and Nicholas Thieberger take the story from 1968 to 1993.¹⁶ They observe that, by the beginning of this period, lexicography in Australian languages was no longer motivated by concerns to produce documentations

¹⁴ Moore, 'T. G. H. Strehlow', 290–1.

¹⁵ Notable early treatments of the history of Aboriginal linguistics are O'Grady, Voegelin, and Voegelin, 'Languages of the world: Indo-Pacific', 2–13, and Capell, 'History of research in Australian and Tasmanian languages'. McGregor, *Encountering Aboriginal Languages*, is a collection of articles dealing with a range of issues in the history of Aboriginal linguistics, while Clara Stockigt's PhD thesis 'Pama-Nyungan morphosyntax' provides a detailed account of early descriptions of the grammar of Pama-Nyungan languages.

¹⁶ Goddard and Thieberger, 'Lexicographic research'.

of moribund languages just for the purpose of their preservation, but was also directed towards use by indigenous people, for instance in education and language-reclamation programmes. Aboriginal people themselves became more actively involved in compiling dictionaries and producing definitions.¹⁷ The advent of personal computers in the early 1980s greatly facilitated the production of dictionaries; earlier, mainframe computers had been used in the production of wordlists, such as the unpublished 'Research dictionary of the Western Desert Language of Australia' of Eric Ten Raa and Susan Tod Woenne (1970–3).¹⁸ Goddard and Thieberger also mention improved methodologies for gathering lexicographic data, going beyond the elicitation of standard lists of English glosses characteristic of early lexicographic research. They mention intensive language learning by the linguist, participant observation, and use of audio recording devices and transcribed texts, as well as collaboration with speakers of the language.

In 'The lexicography of indigenous languages', Thieberger provides a broad overview of modern lexicographic work on indigenous languages of Australia and the Pacific region from the early 1990s. He identifies the main methodologies in the creation and dissemination of dictionaries of these languages, including software and repositories for digital files, for instance the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). Since the advent of the internet in the early 1990s, online dictionaries have appeared, the first being perhaps Peter Austin and David Nathan's dictionary of Gamilaraay.¹⁹ With the increasing prevalence of portable devices such as mobile phones, dictionaries have recently been developed for those media. For instance, dictionaries of Barngarla, Yuwaalaraay, Wiradjuri, and Yorta Yorta are available for iPhones and iPads; and Jean-Christophe Verstraete has produced dictionaries of the Cape York languages Rimanggudinhma, Mbarrumbathama, Morrobolam, and Umpithamu for Android devices.

Thieberger mentions as the main publishers of dictionaries of Australian languages Pacific Linguistics at the Australian National University (since 2011 published as a series under De Gruyter Mouton, Berlin); the IAD Press (Institute for Aboriginal Development); and SILA (the Summer Institute of Linguistics Australia). Earlier, Oceania at the University of Sydney and the

¹⁷ Goddard and Thieberger, 'Lexicographic research', 181; see also Oates, 'Aboriginal recording of Aboriginal language'.

¹⁸ For personal computers, see Goddard and Thieberger, 'Lexicographic research', 181–5.

¹⁹ Austin and Nathan, *Kamilaroi/Gamilaraay Web Dictionary*; see Austin, 'Gamilaraay (Kamilaroi) language', 53–5, for a description.

then Australian Institute of Aboriginal Studies (AIAS) in Canberra (now Australian Institute of Aboriginal and Torres Strait Islander Studies, AIATSIS) published a number of works on Aboriginal languages, including notable dictionaries such as Coate and Elkin's *Ngarinjin-English Dictionary* (Oceania) and Heath's *Nunggubuyu Dictionary* (AIAS).

Following Goddard and Thieberger, this chapter uses the term *wordlist* in reference to any listing of words of an Aboriginal language accompanied by brief English glosses, typically a single word or two. The term *dictionary* is used for a compendium that provides more detailed and sophisticated semantic, grammatical, and/or usage information on the words. Length is not a consideration in these definitions, though wordlists are typically much shorter than dictionaries, and often comprise fewer than 1,000 lemmata, whereas many dictionaries include more than twice this number. Dictionaries of Australian languages are typically bilingual, the meanings of words being almost always given in English, rather than explained using the language itself as the metalanguage.

The account that follows focuses on the missionary contribution to the lexicography of Australian Aboriginal languages. It is organized according to a three-period model of research on Australian languages, which distinguishes a first or early period (1790–1930), a second or transitional period (1930–60), and a third or modern period (post-1960).²⁰ These periods blended into one another, and were not characterized by radical Kuhnian-type paradigm shifts. McGregor suggests that lexicographic research roughly fits the three-period pattern, albeit perhaps with a lag of a decade in the turning points.²¹

Early Period (1800–1930)

Many of the most extensive wordlists and dictionaries of this 130-year period were produced by missionaries, and include a number of published works as well as a perhaps larger and more significant set of unpublished ones. Some of these were stand-alone works, though many were parts of larger descriptions of the languages; only the more significant instances of the latter category are included in the discussion that follows. Many of these early works provide

²⁰ McGregor, 'Introduction' in McGregor, *Encountering Aboriginal Languages*, 9.

²¹ McGregor, 'Introduction' in McGregor, *Encountering Aboriginal Languages*, 19; compare O'Grady, 'Lexicographic research in Aboriginal Australia', which identifies a turning point in the late 1930s, and Goddard and Thieberger, 'Lexicographic research', which identifies a later one in the early 1970s.

lexical information on languages that have since gone out of use, and thus represent important legacy materials essential to the reclamation of many languages of the eastern part of the country.

Lancelot Threlkeld, who had been trained in the Congregationalist Missionary Seminary in Hampshire, England, established the first mission in Australia in 1826, on the eastern side of Lake Macquarie to the north of Sydney. He is also the author of the first description of an Australian language, now known as Awabakal, based mainly on information provided to him by a native speaker of the language, Biraban.²² His earliest publication on the language, *Specimens of a Dialect of the Aborigines of New South Wales* (1827), undertakes little grammatical analysis and does not include a wordlist. His *Australian Grammar* of 1834 contains a discussion of pronunciation and orthography (Part I), a more extensive grammar (Part II), as well as a wordlist (Part III).²³ The wordlist comprises approximately 700 entries over 26 pages. It is organized primarily according to parts of speech (personal names and toponyms, common nouns, verbs), but includes a separate section on terms for parts of the body. Awabakal words are given in the first column, mostly listed alphabetically in the sections; the second column gives brief English glosses in most cases, though sometimes – especially for culturally specific phenomena – more detailed explanations are provided. So the entries ‘Po-ri-bai, A husband’ and ‘Po-ri-kun-bai, A wife’ are followed by

Pun-ti-mai, A messenger, an ambassador. They are generally decorated with the down of the swan or hawk on their heads when on an embassy. They arrange the time, place, and manner of engagement in battle; or when punishing a supposed offender or real aggressor. They bring intelligence of the movements of hostile tribes, or the last new song and dance. When they travel at night, a firestick is always carried by them as a protection against ‘The powers of darkness’. Evil spirits of which they are in continual dread.²⁴

Threlkeld divided Awabakal words into syllables, delimited by hyphens, and employed an orthography that was strongly influenced by the system employed at the time by missionaries working on Polynesian languages; his orthography was employed by a number of subsequent missionaries. A second chapter in Part III provides a range of example sentences with interlinear glosses and free translations into English, organized so as to

²² See further Carey, ‘Lancelot Threlkeld and missionary linguistics’; Wafer and Carey, ‘Waiting for Biraban’.

²³ The wordlist is Threlkeld, *Australian Grammar*, 79–104.

²⁴ Threlkeld, *Australian Grammar*, 91.

illustrate points of the grammar. The *Australian Grammar* was republished in slightly revised form as *An Australian Language* in 1892.

In 1932, the Church Mission Society established the Wellington Valley Mission on the western side of the Great Dividing Range. The first missionaries, William Watson and Johann Christian Simon Handt, apparently produced a wordlist of some 4,000 words in Wiradjuri, which was never published.²⁵ James W. Günther arrived five years after the establishment of the mission and prepared two analyses of Wiradjuri that were probably heavily influenced by the earlier analyses by Watson and Handt. These were not published for more than fifty years, appearing eventually as an appendix to *An Australian Language*.

William Ridley was an itinerant missionary who travelled extensively through New South Wales from 1850. Unlike the majority of missionaries, he was not associated with a mission, and spent less time on particular languages than most missionaries, who spent years or decades on a language.²⁶ From 1852 to 1856 Ridley missionized to the Gamilaraay and neighbouring groups in northern New South Wales. His major linguistic work, *Kámilarói, and Other Australian Languages* (1875; 1866), includes short wordlists of Gamilaraay and a few other languages spoken in the vicinity. These were organized according to part of speech and semantic domain, with mainly single-word English headwords (in some sections listed alphabetically), followed by a representation of the word in the Australian language. The Gamilaraay wordlist amounts to some twenty pages, while the others are considerably shorter, typically no more than a couple of pages. This work also includes a comparative table of basic lexemes in fifteen languages of eastern Australia.

Lutheran missions were established in South Australia soon after the official settlement of the colony in 1836. The earliest Lutheran missionaries were ordained by the Evangelisch-Lutherischen Missions-Gesellschaft zu Dresden (Evangelical Lutheran Mission Society of Dresden), and produced a number of descriptions of languages of the Adelaide region.²⁷ Christian Gottlob Teichelmann and Clamor Wilhelm Schürmann published *Outlines of a Grammar, Vocabulary, and Phraseology, of the Aboriginal Language of South Australia* (1840), the first substantial description of a South Australian language. This work included a grammatical description of Kaurna, the language of the Adelaide region, as well as a vocabulary of some 2,000 words,

²⁵ Stockigt, 'Pama-Nyungan morphosyntax', 134.

²⁶ Stockigt, 'Pama-Nyungan morphosyntax', 480.

²⁷ Stockigt, 'Pama-Nyungan morphosyntax', 177.

listed alphabetically. English glosses are typically short, though part of speech category is indicated. Occasionally more detailed information including example phrases or sentences is provided, as in this example:

Pia, or Bia, *adv.* This important and frequently used particle denotes that the person using it has a doubt, or is ignorant, or at least uncertain, in regard to the subject spoken of – perhaps; may be; as *Pia wa ngaityo yunga? Madli bia pa.* – Where may my brother be? Perhaps he has died. *Wa ngaityo tando wa?* – Where is my bag, where? *Wa alya bia?* – Where may it be? (meaning, I don't know.)²⁸

Teichelmann subsequently sent a more extensive and comprehensive vocabulary of the language to George Grey in South Africa in 1857.²⁹

In 1844, Clamor Schürmann published a wordlist of Barngarla, amounting to some 3,000 words. This work begins with an outline of the grammar of the language, but the bulk comprises an alphabetical Barngarla–English wordlist. Mostly brief English glosses are provided of each word, though part of speech indication is also given, and occasionally example sentences with free translations (unfortunately they are usually not glossed), as follows:

mai-mintyuru, particular with food, stingy
mai-munta, liberal with food
mai-pinta, greedy, close
mai-pinta worta, a close or greedy fellow
maitya, *adv.* expressing assurance, indeed, then, now; as *maitya ngannaru*, why then (viz should it be) *maitya kulbarri ikkardna*, they are many indeed³⁰

In 1843 Heinrich Meyer published a dictionary of Ramindjeri, one of the Ngarrindjeri varieties, which was also prefaced by a grammatical description. The dictionary was again alphabetically organized according to the Ramindjeri headwords. It includes a number of verbs describing cultural practices, as well as many terms for fish and shellfish. This dictionary had somewhat fewer headwords than the dictionaries of Kaurna and Barngarla. However, it also presented a number of bound morphemes as headwords, and provided somewhat more detailed information on their meanings and uses, with a considerable number of example phrases and sentences which usually included interlinear glosses as well as free translations. So, for

²⁸ Teichelmann and Schürmann, *Outlines of a Grammar*, 38.

²⁹ Simpson et al., 'I could have saved you linguists a lot of time and trouble', 97; Stockigt, 'Pama-Nyungan morphosyntax', 189.

³⁰ Schürmann, *Vocabulary of the Parnkalla Language*, 25.

instance, ‘Ka, *particle* expressive of doubt or interrogation; whether’ is illustrated with two glossed examples,

Kam- ma- itye porl-?
 it your it child?
 Ka- itye yamm- ur ell- ai
 whether he another had been being?

explained respectively as ‘Is that your child?’ and ‘I thought he had been another.’³¹

Subsequently, the Congregationalist missionary George Taplin worked on another Ngarrindjeri variety at Point McLeay mission. He produced a manuscript wordlist of perhaps around 1,000 entries alphabetically organized on the English headwords. Soon afterwards he compiled a comparative vocabulary of seventy-one basic meanings in some twenty Aboriginal languages from various parts of the continent. Unlike his Lutheran predecessors, Taplin recognized lamino-dental stops.

The year 1866 saw the establishment of the Bethesda Mission, also known as Killalpaninna mission, in north-east South Australia to the east of Lake Eyre, also by Lutherans. They worked on the local languages, including Diyari, from the beginning. Four grammars of Diyari were written by Bethesda missionaries from 1868 to the turn of the twentieth century, all unpublished manuscripts written in German.³² Carl H. M. Schoknecht’s grammar of 1872 included a vocabulary of more than 1,000 Diyari–German entries, followed by German–Diyari, organized alphabetically. An English translation of this work was prepared in 1947 and published in 1997. Johann Flierl compiled a comparative Diyari–Wangkangurru vocabulary, now in the State Library of New South Wales.³³ The linguistic and anthropological work of the Bethesda missionaries culminated in a thirteen-volume unpublished manuscript by Johann G. Reuther, who served at the mission from 1888 to 1906. Volumes I–IV are a Diyari dictionary, which comprises more than 4,000 entries, copiously illustrated with example sentences. Volume V provides comparative wordlists of up to 1,744 entries in Diyari and neighbouring languages (Arabana, Yawarrawarrka, Wangkangurru, Kuyani, Ngamini, Thirrari, and Yandruwandha), with single English glosses and no illustrative

³¹ Meyer, *Vocabulary*, 67.

³² They are described in Stockigt, ‘Pama-Nyungan morphosyntax’, 299, 372.

³³ Flierl, ‘German vocabulary of native tribes’; see also Stockigt, ‘Pama-Nyungan morphosyntax’, 326–7.

sentences.³⁴ Volume VII lists place-names (2,449 entries). In 1904, Henry Hillier, a layman who taught at the Bethesda mission school, produced a map including some 2,468 place-names in north-eastern South Australia.

A decade later, in 1877, another Lutheran mission, Hermannsburg mission, was established in central Australia on the Finke River, west of Alice Springs. Six grammatical descriptions of the local language, Arrernte, were produced by the missionaries or written on the basis of information provided by them.³⁵ Best-known of the Hermannsburg missionaries was Carl Strehlow, who arrived at the mission in 1892. Strehlow is most famous for his seven-volume ethnography of the Arrernte.³⁶ In addition, he also produced three grammars (just one of which was published), literacy materials, translations of religious texts, and an unpublished dictionary of Western Arrernte amounting to some 223 pages of typescript, probably compiled between 1900 and 1907 under the influence and guidance of Baron Moritz von Leonhardi, his European mentor and editor. All of these materials were written in German. The wordlist was one of the largest and most comprehensive of an Australian language, comprising some 7,600 Arrernte lexemes, 6,800 Luritja words, and 1,200 Diyari words. It is organized into four columns, Western Arrernte, German, Luritja, and Diyari. The German glosses in the second column are mostly brief, although sometimes multiple glosses are given to indicate the range of an Arrernte term, and sometimes more substantial explanations of meanings are provided. Only sporadic grammatical information is provided. Strehlow employed an orthography which he dubbed 'continental' because of the values assigned to the vowels; he attempted to represent sounds consistently, though his was a broad phonetic rather than phonemic orthography that included a number of under- and overdifferentiations. The last Lutheran mission established in South Australia was in 1901 at Koonibba on the west coast of the state. The Revd August Hoff, appointed superintendent in 1920, recorded vocabularies in some of the local languages. These, however, were not published until 2004.

The first missionary wordlist in Western Australia seems to have been compiled by Rosendo Salvado, a Spanish Benedictine monk, founder and first abbot of a mission at New Norcia north of Perth, established in 1846. His

³⁴ See Hercus, 'Looking at some details', for a detailed discussion of Reuther's wordlists, and Lucas and Fergie, 'Pulcaracuranie', on his ethnographic and linguistic work generally.

³⁵ See Stockigt, 'Pama-Nyungan morphosyntax', 373–411, for a detailed description and evaluation.

³⁶ See especially Kenny, 'From missionary to frontier scholar', and Kenny, *The Aranda's Pepa*; see also Kenny, 'Early ethnographic work'.

Memorie storiche of 1851, a historical and descriptive account of Australia and particularly of Benedictine missionary work there, includes a wordlist in the Noongar variety Wadjuk with Italian equivalents.³⁷

In 1890, a mission was established by the Trappist order at Beagle Bay, towards the northern end of the Dampier peninsula in the far north of Western Australia. The monks immediately began to learn the local language, Nyulnyul. Within a few years they had produced a substantial French–Nyulnyul lexicon, as well as a short grammatical description written in French.³⁸ The authorship of these works is uncertain, and versions exist in different hands. The monks most likely each worked on their own copy and shared their findings; almost certainly the superintendent of the mission, Fr Alphonse Tachon, was involved in the compilation of both works. This interpretation seems to be supported by a letter from Daisy Bates to the Lord Abbot of the Trappist mother house, Abbaye Notre Dame de Sept-Fons in Dompierre-sur-Besbre, France, that accompanied the wordlist manuscript she was returning to the order in 1947. This manuscript had been given to her almost fifty years previously, in 1900, by the Spanish missionary Fr Nicolas Emo, who had arrived at the mission five years after its opening; it was presumably his personal copy, in his own handwriting.

The lexicon is alphabetically organized on the French translations, mostly single words. In some instances this ordering is departed from, and a headword specifies a semantic domain that is filled out by further subheadings. The Nyulnyul column often contains (partial) paradigms of the corresponding lexemes, and sometimes phrases (e.g. preverb-inflecting verb collocations for verbal meanings expressed by compound verb constructions) or partial sentences, specific translations of which are usually provided in the French column. Thus this work is more than a mere wordlist and borders on a dictionary. It is a valuable resource on a now moribund language, and deserves further investigation.

The Trappist monks did not understand the phonology of Nyulnyul very well, and wrote words largely according to French orthography. The wordlist itself betrays their missionizing purpose: there are entries for a range of concepts useful in translating religious material (e.g. ‘adulterer’, ‘adulteress’, ‘angel’, ‘beget’, ‘expiate’, ‘idleness’, ‘immortal’, ‘lewd’, ‘obscenity’). In some cases the meanings of indigenous Nyulnyul words were extended (sometimes in unexpected ways, at least to the speaker of a Standard Average

³⁷ Salvado, *Memorie storiche*, 364–75 = Salvado, *Salvado Memoirs*, 256–66.

³⁸ See McGregor, ‘Early Trappist grammar’.

European language); in other cases the new concepts have been analysed into components which are translated (revealing a working knowledge of translation practice) – for instance, the term for ‘bigamist’ is given as *wurumbang maler inier*, literally ‘having many wives’ (phonemically, *wurrumbang malirinyirr* ‘many wife-with’). In still other cases, words are glossed in terms most appropriate to the Christian religious register (for instance, *kurwol* (*kurrwal*) is given under ‘heaven’ and ‘paradise’, but not under ‘sky’). Other domains – such as flora, fauna, kinship, and artefacts – are not as well represented as would be expected for a wordlist of its size.

Other copies of the manuscript wordlist must have remained in the mission and survived the handover to the Pallottine order in 1901, since Fr Joseph Bischofs, superintendent of the Beagle Bay Mission from 1905 to about 1914, produced a typescript English–Nyulnyul wordlist that was clearly heavily based on the earlier Trappist one, with English glosses replacing the original French ones. Unfortunately, this version omits some of the more interesting details and examples. It also makes a few adjustments to the spelling of Nyulnyul words, albeit without improvement to the overall accuracy or consistency of the work.

The reformed blackbirder and pearler Sydney Hadley established a Protestant mission on Sunday Island in 1899. Both he and William Bird, the schoolteacher at the mission, filled out Daisy Bates’ questionnaire of 1904, independently and together. The questionnaires were fairly completely filled out with forms that are recognizably Nyulnyulan, and overall the quality of representation is surprisingly good. There is a high level of agreement between the wordlists of the two men. Later, in 1915, William Bird published a short and rather unremarkable alphabetical English–Jawi wordlist.

Benedictine monks from Salvado’s foundation at New Norcia established a mission at Pago in the northern Kimberley in 1908. A number of priests appear to have produced wordlists and other materials, though most documents are anonymous and undated. Fr Fulgencio Cubero produced a manuscript on the culture of the Aborigines of the mission area in 1924, subsequently translated into English by an unknown person. This work concludes with a wordlist – not included in the English translation – of just under twenty pages. This is organized alphabetically according to the headword in ‘Cuini’ – doubtless Gunin/Kwini, the language variety spoken in the mission region. The Gunin/Kwini words are spelt in a Spanish-style orthography, and many phonemic contrasts are not recognized (for instance, the apical contrast for stops, nasals, and laterals). Cubero did, however, recognize

the velar nasal (writing it <ng>), and even recorded it word-initially. Words are given brief, usually single-word, Spanish glosses.

Intermediate Period (1930–1960)

In the intermediate period between 1930 and 1960, missionaries continued to make a significant contribution to Aboriginal lexicography, though none of their dictionaries or wordlists were published in print form. The dawn of this period is marked by the beginnings of institutionalization of linguistics in Australia: the establishment of anthropology as a discipline in the University of Sydney in 1926, and the formation of the Language Committee in the University of Adelaide in 1930–1.

The Presbyterian missionary Revd James R. B. Love first visited the Kimberley in a six-month stint during 1914–15 at Port George IV mission, where he began to work on Worrorra. He returned to the mission, by then relocated to Kunmunya, in 1927, remaining there until 1940. He recommenced work on Worrorra immediately on his return, and within a year had collected a thousand-word vocabulary, compiled a forty-page grammar, and begun his translation of the gospel of Mark.³⁹ Love went on to complete an MA thesis on Worrorra grammar at the University of Adelaide under the supervision of John Aloysius FitzHerbert, Professor of Classics at the University of Adelaide, and a leading figure in the South Australian Language Committee, which designed a phonetic alphabet for use in description of Australian languages.⁴⁰ Subsequently he compiled 2 wordlists, a 136-page typescript English–Worrorra vocabulary in 1939, and a 121-page manuscript Worrorra–English vocabulary in 1941. Love was an excellent observer and commented intelligently on a number of issues relevant to lexicography, including gender assignment of nouns (including terms for new concepts), and on idioms:

It is curious to find an occasional idiom used just as in English. For example, the expression, What is it? in English may mean, What is that thing? or What do you want? The Worora word *Angudja*? has exactly the same two meanings. Again, in English, we call water that is mineralized ‘hard’. In Worora, mineralized water is called *burudu*, which means ‘hard’.

On the other hand, again, the English idiom cannot be translated literally in many cases. To speak of the heart as the seat of the emotions would not

³⁹ Burgess, ‘Kunmunya years’, 81, quoting from Love’s journal from 1928.

⁴⁰ Monaghan, ‘Norman B. Tindale and the Pitjantjatjara language’; Simpson et al., ‘I could have saved you linguists a lot of time and trouble’.

convey sense to a Worora man. Like Hebrew, which speaks of 'bowels of compassion', Worora speaks of the belly as the seat of gentle affections. A man of good disposition has a good belly. If he is righteously indignant, his belly is hot. The pancreas is regarded as the seat of anger, and nodules in the mesentery are supposed to be the seat of laughter. The ear is the home of the intellect, and one does not lay up memories in the heart, but in the ear.⁴¹

In 1941, Love left Kunmunya for the mission at Ernabella, in central Australia, remaining there until 1946. While there, Love worked on Pitjantjatjara, a Western Desert variety, producing a 42-page Pitjantjatjara–English wordlist in manuscript, and English–Pitjantjatjara wordlists, a manuscript of 61 pages and a typescript of 40.⁴²

In about 1934, Howard H. J. Coate established a base at Isdell River in the northern Kimberley, from which he ministered as an 'itinerant missionary'. Motivated by practical considerations of communication, he started to learn Ungarinyin (Ngarinjin) and began compiling a wordlist. Howard Coate had only a basic education and no training at all in language description. He sought the assistance of J. R. B. Love and from the late 1930s worked with the anthropologist A. P. Elkin and the linguist Arthur Capell, both from Sydney University. Working for these men, Coate gathered a large corpus of mythological texts in Worrorra, Ungarinyin, and Wunambal, and compiled a sizeable wordlist of some 120 pages in these languages. Coate left the Kimberley in 1948 and did not return until the 1960s. His major contribution dates to the third (post-1960) period.

In 1947, Wilfrid H. Douglas was appointed superintendent of Sunday Island Mission. During his brief spell on Sunday Island, Douglas compiled vocabularies of Bardi and Nyulnyul based on Capell's fieldnotes. He also did some of his own recording of Bardi, including lexical items, verb paradigms, and some texts.⁴³ While there, he produced 'Word gems from Iwany', an illustrated wordlist in Bardi, which was presumably intended for use by speakers of the language. He is one of the few missionary linguists to have produced literacy materials for speakers of Aboriginal languages in the twentieth century prior to 1960, though the extent to which the booklet was actually used is not known; it was not formally published for many years (see below). Only after his time on Sunday Island did Douglas undertake formal training in linguistics, as a participant in the first course taught by the Summer Institute of

⁴¹ Love, *Stone-Age Bushmen*, 49–50. ⁴² See Trudinger, 'Converting salvation'.

⁴³ Douglas, 'Alphabetising Bardi', 197.

Linguistics in Australia, in 1950. Like Coate, his main contribution was made in the third period.

One of the most remarkable dictionaries produced in the intermediate period is represented by Parts II–IV of Nekes and Worms' *Australian Languages*, published in microfilm in the series Micro-Bibliotheca Anthropos.⁴⁴ In late 1930 the Pallottine missionary Fr Ernest (Ernst) A. Worms arrived in Broome, Western Australia, and shortly thereafter began working on the languages and peoples of the region. He was joined in 1935 by his former teacher in the Limburg Seminary, Germany, Fr Hermann Nekes. The two men collaborated over the following decade or so, gathering and analysing primary data on languages of Australia, most particularly the languages of the Dampier peninsula, for their magnum opus *Australian Languages*.

Parts II–IV of Nekes and Worms' *Australian Languages* account for over 70 per cent of the thousand or so pages of this work. Part II, 'Dictionary English–native languages', is effectively an alphabetically organized finder list of words from English to Aboriginal languages, comprising some one hundred pages. It provides little information on the words in Aboriginal languages other than their source language, indicated by an abbreviation. Unfortunately, the many errors, inconsistencies, and somewhat weird choices of headword (for instance, 'incombustible wood') compromise its usability as a finder list.⁴⁵ The listing of Part II also serves to a certain extent, albeit inconsistently, a secondary function of grouping together words into semantic domains.

The much more substantial (more than 600 pages) Part III, which goes under the rather grandiose title 'Dictionary native languages–English (a paradigmatic syntax)', is an alphabetically arranged listing of words from a range of Aboriginal languages, mostly from Dampier Land and nearby parts of the Kimberley. Nekes and Worms employed a 'phonetic' orthography in their transcription of words in Aboriginal languages; in fact, it is effectively a phonemic one using letters, digraphs, and diacritics to represent the phonemes, though not all phonemic contrasts in all of the languages are recognized. They were aware of some of the main features characteristic of the phonologies of Australian languages (general lack of voicing contrasts in stops, three phonemic vowels, and contrastive retroflexion in apical consonants). There are a number of inconsistencies in their alphabetization of

⁴⁴ See McGregor, 'Frs Herman Nekes and Ernest Worms' dictionary', for a detailed description and appraisal.

⁴⁵ McGregor, 'Frs Herman Nekes and Ernest Worms' dictionary', 6.

words: for instance, diacritics are ignored, and digraphs are differently treated (some as separate items in the alphabetic list, others under the initial letter).

Part III comprises around 9,000 headwords, many associated with more than one language. Some thirty languages are represented by lexical data collected by the authors themselves. Best represented are Nyulnyulan languages (traditionally spoken in Dampier Land and nearby regions of the Kimberley); each of Bardi, Jabirrjabirr, Nimanburru, Nyikina, Nyulnyul, and Yawuru has more than 1,000 entries (often shared). Jukun, Jawi, and Ngumbarl are represented by between 1 and 22 headwords. A handful of Kimberley languages of other families, Bunuban, Jarrakan, and Worrorran, are represented by fewer than 100 headwords. A few Pama-Nyungan languages of the Kimberley are represented by between 300 and 1,500 headwords, while Pama-Nyungan languages from further afield, mainly New South Wales, are represented by about 100 or fewer headwords.⁴⁶

Whereas Part II is no more than a wordlist, Part III counts as a dictionary: although perhaps the majority of headwords are given only brief glosses, in some instances – typically where the word expresses some culturally significant meaning – more substantial semantic information is provided, and occasionally (for 10 per cent of headwords) the part of speech is indicated. A considerable number of the entries are provided with examples of usage in the form of phrases or clauses, hence the ‘paradigmatic syntax’ in the title for this part. A headword may contain information of the following type, in approximately this order (few entries are this full, however): headword; (etymology, gloss); languages; (second headword, languages); part of speech; gloss; cross-reference (to another lexeme); paradigm or other form of the headword; example phrases and/or sentences. Interestingly, not just free words are included in this part, but also bound morphemes, albeit inconsistently. Some borrowings from English are also included.

For illustration, I provide the beginning of the entry for *mōndo* ~ *mōnd*.

mōndo Nig., Y., *mōnd* B., DD., Nimb., NN. bury, spell; magic ceremony of burying the name of an enemy: the bones of a lizard are broken with a pointing stick (*wadangar*), the animal is singed and buried in a hollow tree. The name of the person to be enchanted is called during these actions; the victim will soon die. s. *Goraŋara*, A.L. VI, 217, *Darin mundi* Wundj., *mundig*, *mandagi*, *mandagidgid*, (*manta*, *mantu* Tadjane, secret-Tind., Trans. Roy. Soc. S.A., 1937, 109 ss.; Rec. S.A. Mus. Vol. VI (1939) 246).

mōndo ma-man nilawal, to bury the name,

⁴⁶ See McGregor, ‘Frs Herman Nekes and Ernest Worms’ dictionary’, 10, for figures.

mōnde ejer-am nilawal morinj wamb-en NN., burying they put name of woman men = they men buried the name of the woman[.]⁴⁷

The small Part IV, 'Comparative dictionary of Australian pygmoid languages', is a wordlist in half a dozen languages of the Cairns rainforest gathered by Fr Worms in the 1940s. This wordlist is alphabetized on English headwords. R. M. W. Dixon dismisses the Dyirbal entries in Part IV with the words 'Almost every word is mistranscribed.'⁴⁸ This may be true, but Worms worked only for a very short time on these languages, and alone; his transcriptions are generally not so bad as to render words unrecognizable.

Some fifty years after its initial appearance in *Micro-Bibliotheca Anthropos*, *Australian Languages* finally appeared in book form – at least, Parts I (grammar) and V (texts) were finally published in extensively revised form, edited by the present author. The other parts, the dictionaries, were excluded from the printed book on the basis of their size, and also because of the limited usefulness of paper representation. Instead, their contents were published on a CD-ROM accompanying the book. The dictionaries from each part may be interfaced via either the SIL International software Toolbox or HTML files. The latter interface links to facsimiles of the relevant page of the microfilm.

The Benedictine mission at Pago, mentioned above, was relocated to Kalumburu, a pool on the King Edward River, during the period 1932–7. The mission superior Fr Thomas Gil produced a wordlist and translations of religious materials in what he called the Pela language, presumably Gunin/Kwini.⁴⁹ Two wordlists, evidently in different hands, probably date from the early 1930s, judging from the orthography, which is the same as that employed by Fr Gil in his translations, and different to the system used by Fr Cubero in 1924.⁵⁰ One of these is presumed to have been compiled by Fr Gil (unfortunately I have no sample of his writing for verification). One wordlist comprises 1,546 English headwords; the other, also alphabetically organized by English headword, has 1,439 entries. There is considerable agreement between the two wordlists both in terms of headwords and spelling. An unusual feature of the orthography is that the digraph <ng>

⁴⁷ Nekes and Worms, *Australian Languages* (1953), 731. ⁴⁸ Dixon, *Dyirbal Language*, 365–6.

⁴⁹ McGregor, *Gunin/Kwini*, 13. See Capell and Coate, *Comparative Studies*, 2–4, and McGregor, *Gunin/Kwini*, 3–6, on problems in identifying languages and names in the region.

⁵⁰ A few other wordlists were compiled by Drysdale River missionaries, though it is impossible to guess at their dates; they employ orthographies different to both Cubero's and Gil's.

for the velar nasal is always followed by <h>; an explanatory note says that <h> indicates a syllable boundary, even though many words have an initial velar nasal. The shorter wordlist provides considerably more information on the meaning of the words, specifying the headwords more precisely, as well as more information on inflectional forms of verbs and prefix-taking nouns; it also contains more example sentences. The Benedictine missionary Fr Seraphim Sanz, who arrived at Kalumburu Mission in 1939, also recorded information on the language. He may have compiled unpublished wordlists during the intermediate period, but his sketch grammar and dictionary of Gunin/Kwini were not published until the twenty-first century.

Modern Period (1960–Present)

The modern period of research is roughly delimited by increased institutionalization of linguistics as a discipline of its own, in the form of courses conducted by the Summer Institute of Linguistics from 1950 and the establishment in 1961 of a branch of that organization in Australia, the establishment of the departments of linguistics in Australian universities, beginning with Monash University in 1965, and the establishment of the AIAS in the early 1960s.⁵¹ Many missionary linguists – Catholic as well as Protestant – took Summer Institute of Linguistics courses; many completed undergraduate and/or graduate studies in linguistics at universities; and a number were funded at some point by grants from the AIAS, later AIATSIS. An unprecedented number of stand-alone dictionaries of high quality were produced and published in this period by missionary linguists, who also made significant contributions to literacy in Aboriginal languages.

The lexicographic research on Pitjantjatjara begun by J. R. B. Love at Ernabella in the 1940s was taken up again at the cusp of the modern period by Wilfrid Douglas, who published *An Introduction to the Western Desert Language*, a shortish grammatical description, as well as an *Illustrated Topical Dictionary of the Western Desert Language, Warburton Ranges Dialect*. Like his previous unpublished wordlist of Bardi, this illustrated wordlist, which was organized into semantic domains, was clearly intended as a resource for speakers of the language, not purely for documentation purposes. This work is one of the first to represent words of an Australian language in a phonemic orthography (devised by the author some years

⁵¹ On the Summer Institute, see Oates, 'Summer Institute of Linguistics', and Oates, *Against the Wind*.

previously).⁵² It has also proved very popular, more so than the average dictionary, with three revised editions appearing over the subsequent half century. Douglas' illustrated wordlist of Bardi, made in the 1940s, was finally published in the early 1990s as *Bardi Language Word-Book*.

In the late 1980s, Douglas also published a more comprehensive dictionary of the Western Desert Language, primarily the Ngaanyatjarra dialect, as *An Introductory Dictionary of the Western Desert Language*. This work, comprising about 2,400 distinct lexemes, is divided into three sections, Western Desert to English, English to Western Desert, and a semantically organized section divided into 99 domains. The Western Desert–English section is largely an alphabetical listing of Western Desert lexemes with indication of part of speech and a brief explanation of the meaning of the word; reference is also made to the relevant semantic domains for each headword in the third part. This information is laid out in columns. The English–Western Desert section effectively reverses the third and first columns of the first section, and omits the references to semantic domains. Each of the semantic domains of the third section is made up of an alphabetical listing of Western Desert lexemes together with English explanations, sometimes more detailed than those given in the first two parts, especially where culturally relevant categories are concerned.

Other wordlists and dictionaries of the Ngaanyatjarra variety were produced by the next generation of missionary linguists, Amee Glass and Dorothy Hackett.⁵³ Other Western Desert varieties have also had dictionaries or wordlists compiled by missionary linguists, including a Pintupi/Luritja dictionary with about 4,000 entries, compiled by Kenneth Hansen and Lesley Hansen, which has gone into three editions; a more recent picture dictionary of Luritja by Kenneth Hansen; a dictionary of Martu Wangka by James Marsh; and a Kukatja dictionary by Hilaire Valiquette, which is based on unpublished work by the Pallottine missionary Anthony Peile.

Early in the modern period, Howard Coate returned to the Kimberley after an absence of some fifteen years, and recommenced his work on indigenous languages, mainly Worrorran. The newly established AIAS funded his research from 1963 to 1972, during which time he undertook a three-month course with the Summer Institute of Linguistics (1967) and began a grammar and dictionary of Ungarinjin. The resulting two-volume *Ngarinjin–English Dictionary*, co-authored with A. P. Elkin, was published in

⁵² It was first published in Douglas, 'Phonology'.

⁵³ Glass, *Ngaanyatjarra Word List*, and Glass and Hackett, *Ngaanyatjarra & Ngaatjatjarra to English dictionary*.

1974.⁵⁴ It is alphabetically organized on Ungarinyin headwords (which number at least 8,000), which are written in a phonemic orthography; unfortunately there is no English finder-list. This work goes far beyond providing single-word English glosses for the words, and offers more detailed explanation of the meanings of lexemes. In addition it includes grammatical information as well as numerous examples of usage, some from Coate's extensive collection of mythological texts, as in this entry:

JOL -o-, -e-, to mark, to select, to choose, to pick out, to make a mental note of; balja buma nurul djonari malngud gudedi jol wo, (lit.) go you tree shade having that mark it (you go and pick out a large shady tree); gula njunda jol wenj njarug made woro, about here pick it out and saw it (mark it about here and saw it); jol bunonerinjari, you two were the ones who viewed it; be jol aŋoni, already I viewed it; buradaŋa jol buŋoni, I viewed the mob; jol wuŋon, I mark it; jol aŋon, I mark him.⁵⁵

This dictionary remains one of the most comprehensive for an Australian language.

Also published in the 1970s in the same series, Oceania Linguistic Monographs, were two other two-volume dictionaries. The first was a dictionary of Nunggubuyu by the Revd Earl J. Hughes. This dictionary is organized alphabetically on Nunggubuyu headwords (numbering around 6,000), which appear in the first column; subsequent columns present the English meaning and grammatical information. The second was the Revd Laurie Reece's bidirectional dictionary of Warlpiri, with an estimated 2,200 Warlpiri headwords. Neither of these works includes illustrative phrases or clauses.

A decade prior to the publication of Coate and Elkin's *Ngarinjin-English Dictionary*, SIL missionary linguists William J. Oates and Lynette Oates had compiled a dictionary of Gugu-Yalanji, which also included example phrases and sentences. This was expanded in the 1980s by other SIL missionaries, Hank Hershberger and Ruth Hershberger. Many other missionary linguists working under the auspices of SIL made significant contributions to Australian Aboriginal lexicography in the post-1960s period. These include, among others, dictionaries or wordlists in the Pama-Nyungan languages Gidabal, by Brian Geytenbeek and Helen Geytenbeek (1971); Walmajarri, by Eirlys Richards and Joyce Hudson (1990); Wik Munkan, by Christine

⁵⁴ See McGregor, 'Introduction', in McGregor, *Studies in Kimberley Languages*, 5–6, for further details.

⁵⁵ Coate and Elkin, *Ngarinjin-English Dictionary*, 299.

Kilham, Mabel Pamulkan, Jennifer Pootchemunka, and Topsy Wolmby (1986); and Muruwari, by Lynette Oates (1992).⁵⁶ Works on non-Pama-Nyungan languages include Kathleen Glasgow's dictionary of Burarra-Gun-Nartpa (1994) and Velma Leeding's unpublished dictionary of Anindilyakwa/Enindhilyagwa (1977).

Missionaries of the Pallottine order continued linguistic investigations into the modern period, although they attempted nothing on the grand scale of their predecessors Frs Nekes and Worms. Most notable is perhaps Fr Kevin McKelson, who arrived in the Kimberley in 1954, and continued working with Aboriginal people and their languages over the following fifty years, for thirty of which he served at La Grange Mission (now Bidyadanga community). McKelson made a significant contribution to Australian linguistics, recording information on a number of poorly documented Kimberley languages, most notably five Pama-Nyungan languages of the southern coastal region near La Grange Mission: Nyangumarta, Karajarri, Yulparija, Mangala, and Juwaliny. He produced primers in a number of these languages (mostly unpublished), as well as wordlists and dictionaries, one of the most significant being his *Topical Vocabulary in Northern Nyangumarta* (1989).

Aside from the work realized in the Kukatja dictionary published by Valiquette, McKelson's contemporary Anthony Peile did some truly innovative lexicographical work in two main semantic domains: the botanical, and the human body. He did not produce dictionaries of these domains, but rather provided extensive discussion of the meanings of lexemes from them, which he linked via example sentences to the Kukatja views and concepts of the anatomy and functioning of plants and the human body. His work was not published until after his death, and in edited form. Peile's 'Kukatja botanical terms' remains quite faithful to the original text, but his *Body and Soul* is highly edited, and much of Peile's discussion and exemplification of Kukatja terms has been expunged, compromising its lexicographic contribution.⁵⁷

Until quite recently, missionary linguists, like the majority of academic linguists, were interested exclusively in the indigenous languages of Australia. Post-contact varieties such as pidgins and creoles were not generally highly regarded, and were not considered to be worthy of study. Love expressed a typical sentiment for his times in his evaluation of Pidgin English as a 'travesty of a language' taught to Aborigines by ignorant whites, and his

⁵⁶ Richards and Hudson's dictionary had been preceded by Hudson and Richards, *The Walmatjari*.

⁵⁷ See further McGregor, 'Kukatja ethno-physiology and medicine'.

conclusion that, if exposed to 'proper' English, Aborigines would learn it.⁵⁸ It is thus unsurprising that missionaries contributed little to the lexicography of post-contact varieties until the 1970s, when John Sandefur and Joy Sandefur began working on Northern Territory Kriol and a little later Joyce Hudson began working on the Fitzroy Crossing variety. The Sandefurs produced grammatical descriptions of Kriol, as well as a dictionary. Subsequently, the Summer Institute of Linguistics published a more comprehensive *Kriol Dictionary, with Kriol to English and English to Kriol* (1986).

Conclusions

The present overview is somewhat selective and does not pretend to cover all contributions of missionaries to the lexicography of the indigenous languages of Australia. In particular, it focuses on published and unpublished dictionaries, and does not provide a comprehensive account of the numerous wordlists produced by missionaries. Very short wordlists accompanying descriptions have been largely ignored. There is also a focus on works that I consider to be most notable. I have attempted to cover the main trends in missionary lexicography, and to situate the work of missionaries in the context of the lexicography of Australian languages generally. It is clear that the missionary contribution to the lexicography of indigenous languages has been, and continues to be, substantial, and the wordlists and dictionaries produced by missionaries since the 1830s provide an important source of lexicographical information on the languages, many of which are highly endangered or moribund.

⁵⁸ Love, *Stone-Age Bushmen*, 41.

Appendix 1 The Language Varieties

This is a list of the language varieties the wordlists of which are treated in this volume, with very basic taxonomic information, and a note on the lexicographical record. Individual language varieties which appear in the list are identified in SMALL CAPITALS in headings and cross-references, and language groups or families which appear in the list are identified in *italics* in headings and cross-references.

ABENAKI. *Algonquian* language with a few speakers in Quebec and New England; early dictionaries are those of S. Rale (a. 1724) and J. Aubery (a. 1756).

ACHAGUA. *Arawakan* language spoken in Colombia; a grammar and wordlist by A. de Neira and J. Rivero (begun no later than 1703) is extant in manuscript.

AFRIKAANS. West *Germanic* language based upon the dialects Dutch colonists took to the Cape region of South Africa; today one of eleven official languages of the Republic of South Africa (others include ENGLISH, XHOSA, and ZULU). A first wordlist of 1844, by A. Changuion, treated it as a deviant form of DUTCH; a normative dictionary (*Patriot woordeboek/Patriot Dictionary*) appeared in 1902–4, and a multi-volume scholarly *Woordeboek van die Afrikaanse taal* has been in progress since 1950.

Afroasiatic. Major language family or phylum of west Asia and parts of Africa (cf. *Niger-Congo* and *Nilo-Saharan*), including COPTIC and the varieties of *Egyptian*, HAUSA, OROMO, and the *Semitic* languages.

AHOM. Tai-Kadai language formerly spoken in Assam, north-eastern India, where it was the court language of the Ahom kingdom; no longer used natively, but currently undergoing some revitalization; there was a written tradition from the thirteenth century, and the first lexicographical use of ASSAMESE was in a dictionary of Ahom of 1795; there was a dictionary with Assamese and ENGLISH in 1920.

AKKADIAN. *Semitic* language spoken in ancient Mesopotamia from the third to the first millennium BC; replaced by a variety of *Aramaic* from c. 800 BC onwards; there are SUMERIAN–Akkadian wordlists from c. 2350 BC and monolingual Akkadian synonym lists from the first millennium BC.

ALBANIAN. The sole living member of a branch of the *Indo-European* language family, spoken in Albania and neighbouring regions. The first free-standing dictionary was by F. Bardhi (1635).

Algonquian. Family of languages spoken (or, in some cases, formerly spoken) in northern North America, including ABENAKI, ALGONQUIN, CHEYENNE, CREE, the *Delaware* languages, MAHICAN, MASSACHUSETT, MIAMI-ILLINOIS, MONTAGNAIS, NARRAGANSETT, OJIBWA, and PASSAMAQUODDY.

ALGONQUIN (Algonkin). Variety of OJIBWA.

ALLENTIAC. *Huarpean* language formerly spoken in what is now Argentina; it is documented in a grammar and wordlist of 1607 by L. de Valdivia.

Ancient South Arabian. Group of *Semitic* languages attested in inscriptions from the eighth century BC to the sixth century AD (and distinct from ARABIC); the alphabetical order developed in the writing of these languages appears to have had an influence on the lexicography of EGYPTIAN.

ANDAQUI. Paezan language formerly spoken in what is now Colombia; eighteenth-century manuscript Andaqui-SPANISH wordlists are transcribed as part of the *Colección Mutis*.

ANGLO-NORMAN. Variety of FRENCH which developed in England after the Norman Conquest of 1066, declined in vernacular use from the mid thirteenth century, but was a source of the tradition of using French in certain English legal records until the late seventeenth. Short lexicographical texts were produced from the early twelfth century onwards, and dictionaries were published as late as R. Kelham's of 1779.

ANINDILYAKWA (Enindhilyagwa). *Gunwinjguan* language spoken in the Northern Territory of Australia; there is an unpublished dictionary of 1977 by V. Leeding.

ARABANA. *Pama-Nyungan* language with a few speakers in South Australia; a wordlist was included in J. Reuther's comparative work of the early twentieth century.

ARABIC. *Semitic* language of the Middle East and North Africa, recorded in inscriptions from antiquity (but not to be confused with the *Ancient South Arabian* languages), and in a long literary tradition of which the Qur'ān (seventh century) is the key text; now an official language in the twenty-two countries of the Arab League and several others; considerably diversified in spoken use. Rich and complex lexicographical tradition from the eighth century onwards.

Aramaic. Group of *Semitic* languages widely spoken natively and as *lingue franche* from the eastern Mediterranean coast as far east as the Persian empires in the first millennium BC and into the first millennium AD (with the result that the script developed for them came to be used for *Iranian* languages and even the Turkic UYGHUR), and used in Jewish and Christian literature including the Talmud, Bible versions called the Targums, and a large body of Christian texts (the language of the latter is sometimes called SYRIAC); varieties are still spoken in

Iraq and elsewhere, some by declining populations. Aramaic vocabulary appears in HEBREW dictionaries from the tenth century; the first dictionary devoted to Targumic Aramaic was by E. Levita (1541).

Arawakan (Maipurean). Language family with members distributed across South America and, formerly, some of the Caribbean islands, including ACHAGUA, LOKONO, MOXO, and the language variety spoken on Dominica and documented by R. Breton (1665).

ARMENIAN. *Indo-European* language with official status in the republic of Armenia, formerly a *lingua franca* in parts of the Mongol and Ottoman empires. There are medieval wordlists, including one from Yemen; the first printed dictionary was by F. Rivola (1621).

ARRERENTE. *Pama-Nyungan* language, or group of languages, spoken in the Northern Territory of Australia; there are unpublished dictionaries of Western Arrernte, including an early twentieth-century dictionary by C. Strehlow and a mid-twentieth-century one by T. Strehlow.

ASSAMESE. New *Indo-Aryan* language, primarily spoken in Assam in north-eastern India. Used in the late eighteenth century to gloss AHOM terms in vocabularies in manuscript; the earliest dictionary is by M. Bronson (1839), and there were substantial dictionaries by native speakers from the late nineteenth century onwards.

Athabaskan (Dene). Family of languages spoken in western North America including CARRIER, KOYUKON, NAVAJO, and WESTERN APACHE; several languages in the northern branch of the family were covered in an early dictionary by E. Petitot (1876).

AVESTAN. *Iranian* language spoken in the ancient world but now only attested in manuscripts of the thirteenth century AD onwards, mostly relating to the Zoroastrian religion, but also including a wordlist with Middle PERSIAN.

AWABAKAL. *Pama-Nyungan* language formerly spoken in New South Wales, currently undergoing some revitalization; it was the subject of the first description of an Australian language, by L. Threlkeld, with the assistance of Biraban (1827, expanded to include a wordlist in 1834).

AYMARA. Aymaran language spoken in Bolivia and Peru; a dictionary by L. Bertonio was published in 1612.

Baltic. Branch of the *Indo-European* language family, including LATVIAN, LITHUANIAN, and OLD PRUSSIAN.

BAMBARA (Bamanankan). *Niger-Congo* language widely spoken in Mali; the first dictionary was by J. Dard (1825).

Bantu. Language family of central and southern Africa, being part of the larger *Niger-Congo* group, and including CINYANJA, ECHUWABO, EMAKHUWA, GITONGA, KAMBA, KIKONGO, KIMBUNDU, KIMWANI, KIPOKOMO, NDEBELE, NYIKA, NYUNGWE, RONGA, SENA, the *Shona* languages, SWAHILI, UMBUNDU, XHOSA, YAO, and ZULU.

- BARDI.** *Nyulnyulan* language with a few speakers in Western Australia; a single word was written down by a privateer in the seventeenth century, and several wordlists were made in the twentieth century.
- BARNGARLA.** *Pama-Nyungan* language formerly spoken in South Australia, currently undergoing some revitalization; there is a dictionary app based on the nineteenth-century dictionary of C. Schürmann.
- BASQUE.** Language isolate spoken in provinces straddling the French–Spanish border and recorded from the eleventh century. There are sparse medieval lexicographical records, and bilingual wordlists (usually Basque–SPANISH) from the sixteenth century, a landmark printed dictionary being M. Larramendi’s of 1745.
- BENGALI (Bangla).** New *Indo-Aryan* language spoken by a very large population in Bangladesh and India; the earliest extant literary texts are from before the twelfth century AD. A bidirectional dictionary with PORTUGUESE by M. da Assumpçam was published in 1743.
- BICOL.** Group of *Malayo-Polynesian* languages spoken in the Philippines and forming a continuum with VISAYAN; the first extant wordlist, by M. de Lisboa, was compiled before 1628 but published only in 1754.
- BODO.** *Sino-Tibetan* language spoken in and beyond north-eastern India, unwritten until the nineteenth century, as was its close relative GARO; there is a dictionary with ASSAMESE (1987).
- BOSNIAN.** *Slavic* language variety spoken in what is now Bosnia and Herzegovina; a manuscript Bosnian–TURKISH dictionary, presenting both languages in Arabic script, was compiled by Muḥamed (Bosnevī) Hevāī Uskufī in 1631.
- BRAJ (Braj Bhāṣā).** Variety of Hindi (see HINDI–URDU); a brief Braj–PERSIAN glossary appended to a late seventeenth-century text circulated in the eighteenth century.
- BRETON.** *Celtic* language spoken in the formerly independent duchy of Brittany (from 1532 the north-western part of France), recorded in glosses and onomastic material until the fifteenth century, with a literary tradition thenceforward. Dictionaries from J. Lagadeuc’s Breton–FRENCH–LATIN *Catholicon* of 1464 onwards.
- BULGARIAN.** South *Slavic* language spoken in and beyond Bulgaria. An early form of the language was the basis for CHURCH SLAVIC, and there has been a continuous literary tradition since the tenth century; the first comprehensive dictionary was by N. Gerov (1895–1904; supplement by T. Pančev, 1908).
- Bunuban (Bunaban).** Small family of languages with a few speakers in Western Australia; Bunuban languages are represented briefly in H. Nekes and E. Worms’ *Australian Languages* (1953).
- BURARRA-GUN-NARTPA.** *Gunwinjguan* language spoken in the Northern Territory of Australia; there is a dictionary by K. Glasgow (1994).

ČAKAVIAN. Variety of CROATIAN.

CARIB. *Cariban* language spoken in the north-eastern part of South America; P. Pelleprat, *Introduction a la langue des Galibis* (1655), includes a short subject-ordered FRENCH–Carib wordlist.

Cariban. Language family spoken in the northern part of South America, including CARIB and CUMANAGOTO.

CARRIER. *Athabaskan* language spoken in British Columbia; a large dictionary by A.-G. Morice was published in 1932.

CATALAN. *Romance* language spoken in the Spanish province of Catalonia, and elsewhere in Spain, Andorra, and France. The first dictionaries are of the fifteenth century, but the relationship of Catalan to the dominant SPANISH delayed the appearance of monolingual dictionaries, the first to enjoy widespread prestige being the *Diccionari general de la llengua catalana* of P. Fabra (1931–2).

CEBUANO (Sugbú). *Visayan* language widely spoken in the Philippines; an ITALIAN–Cebuano wordlist was made by A. Pigafetta in the early 1520s, and Cebuano forms are distinguished in the *Visayan* wordlist of A. de Méntrida (1637).

Celtic. Branch of the *Indo-European* language family including BRETON, CORNISH, GAULISH, IRISH, and WELSH.

CHAGATAY. An artificial graphic *koine* form introduced for *Turkic* dialects widely spoken in Central Asia during the Timurid era; remained in use as a literary language until the early twentieth century and served as a *lingua franca* throughout Inner Asia. Texts from the end of the fourteenth century or beginning of the fifteenth; the earliest dictionary (Chagatay–PERSIAN) was compiled in the fifteenth century but its only surviving manuscript dates back to the beginning of the eighteenth.

CHEYENNE. *Algonquian* language spoken in Montana and Oklahoma; a dictionary by R. Petter was published in 1915.

CHIBCHA. Chibchan language formerly spoken in what is now Colombia; a SPANISH–Chibcha wordlist dated 1612 is transcribed as part of the *Colección Mutis*.

CHINESE. Group of *Sino-Tibetan* language varieties including HAKKA, the MANDARIN TOPOLECTS (*guānhuà fāngyán*), Northern and Southern MÍN, WÚ, and YÜÈ, spoken in the People's Republic of China, Taiwan, and Singapore, and by diasporic communities in many countries. These varieties are sometimes referred to as *topolects*, a word which avoids the suggestion that they are all mutually intelligible (they are not), but acknowledges that their shared written form, and the perception of their speakers, makes it inappropriate to call them separate languages. There has been literacy since the late second millennium BC, and the language of several discontinuous periods of activity can be periodized as follows: Old Chinese refers to the language of the eleventh century BC to the seventh century BC, in which the earliest literary classics were written

(Classical Chinese, also called Literary Chinese and Literary Sinitic, continued to be used as a literary language, and is the variety documented in many pre-modern Chinese dictionaries); Middle Chinese is the language of the seventh through twelfth centuries AD; Early Modern Chinese is the language of the fourteenth–eighteenth centuries. The standard language of the People's Republic of China, and the major language of modern Chinese lexicography, is known as MANDARIN (*pǔtōnghuà*). A continuous, and extremely rich and complex, lexicographical tradition began about the fifth century BC.

CHIKUITANO. Language isolate spoken in Bolivia; a wordlist is in an anonymous eighteenth-century *Arte y vocabulario de la lengua chiquita*.

CHOCTAW. *Muskogean* language spoken in the south-eastern United States; a dictionary was compiled by C. Byington in the nineteenth century.

CHURCH SLAVIC. The name given to the literary language of the oldest *Slavic* texts, originating with ninth-century evangelization, and primarily based on South Slavic spoken varieties that have features in common with present-day Bulgarian and Macedonian; there were printed wordlists translating Church Slavic words into the East Slavic variety called RUTHENIAN from L. Zyzanij's *Leksis* (1596), and major scholarly dictionaries such as that of F. Miklošič in the nineteenth century; Church Slavic is still used as a liturgical language.

CINYANJA (Chewa). *Bantu* language spoken in Malawi, Zambia, Mozambique, and Zimbabwe; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856).

CLASSICAL EGYPTIAN (Neo-Middle Egyptian). Variety of *Egyptian*, originally the language of literature of the Middle Kingdom (c. 2119–1794 BC); during the New Kingdom (1550–1069 BC), although already obsolete in normal speech and writing, it became the idiom of official and, especially, religious literature, and remained in use together with LATE EGYPTIAN and DEMOTIC until the end of pharaonic culture in the third century AD.

COPTIC. Variety of *Egyptian* which developed from DEMOTIC, spoken in Egypt from the third or fourth century AD until the fourteenth, when it was succeeded by ARABIC, and preserved in liturgical use thereafter. There were glossaries with GREEK and, to a lesser extent, LATIN in antiquity, and with Arabic in the Middle Ages; the first modern printed dictionary was part of A. Kircher's *Lingua aegyptiaca restituta* (1643).

CORNISH. *Celtic* language spoken in the extreme south-west of England, extinct by the end of the eighteenth century, with a limited late medieval literary tradition. There were early glosses, and Cornish–ENGLISH wordlists were made from the seventeenth century onwards.

CREE. *Algonquian* language spoken across much of Canada; the first dictionaries were produced in the third quarter of the nineteenth century.

- CROATIAN.** South *Slavic* language, spoken in parts of the Balkans formerly under (Austro-)Hungarian, Turkish, and Venetian rule, and now the official language of the Republic of Croatia, classified into ŠTOKAVIAN, ČAKAVIAN, and KAJKAVIAN dialects depending on the main forms of the interrogative pronoun meaning ‘what?’ First recorded in medieval texts, in which there is no hard distinction between Croatian and SERBIAN. There is a fifteenth-century manuscript wordlist, and there are printed wordlists from the ITALIAN–Croatian *Opera nuova che insegna a parlare la lingua schiavonesca* of P. Valentiano (1527) onwards; there is a 23-volume dictionary of the nineteenth and twentieth centuries, *Rječnik hrvatskoga ili srpskoga jezika*.
- CUMAN.** Kipchak *Turkic* language formerly spoken in the Khanate of the Golden Horde to the north of the Black Sea; there are wordlists of late thirteenth-century origin (with LATIN, PERSIAN, and GERMAN) in the fourteenth-century Codex Comanicus.
- CUMANAGOTO.** *Cariban* language spoken in Venezuela; a SPANISH–Cumanagoto wordlist is printed in F. de Tauste’s *Arte y vocabulario de la lengua de los indios chaymas, cumanaotos, cores, parias, y otros diversos* (1680).
- CZECH.** West *Slavic* language attested in short texts from the tenth century onwards and in literature and Bible translations from the later Middle Ages, and now a co-official language of the Czech Republic. Native lexicographical traditions developed in the Middle Ages; the first printed LATIN–Czech dictionary was by J. Vodňanský (1511); the nine-volume *Příruční slovník jazyka českého* was published 1935–57.
- DAKOTA.** *Siouan* language with a few speakers in and beyond North and South Dakota; dictionaries, usually with ENGLISH but including the HEBREW–Dakota dictionary of S. Pond, were made by missionaries from the mid nineteenth century onwards.
- DANISH.** North *Germanic* language spoken in Denmark (where it is the official language) and its overseas territories, with substantial texts from the fourteenth century onwards. Dictionaries from the sixteenth century; a major multi-volume scholarly dictionary, *Ordbog over det danske Sprog*, was published 1919–56, with supplement 1992–2005.
- Delaware (Lenape).** Pair of *Algonquian* languages, Unami and Munsee, both formerly spoken in what is now the north-eastern United States (Munsee has a few speakers, and is currently undergoing some revitalization); dictionaries were made by Moravian missionaries in the eighteenth century.
- DEMOTIC.** Variety of *Egyptian* and name of its cursive writing system; used primarily for *belles lettres* and documentary texts, later also for official and religious texts (for which CLASSICAL EGYPTIAN also continued to be used, calling for wordlists glossing Classical Egyptian in Demotic); preceded by LATE EGYPTIAN, and in use in speech and writing from the seventh century BC until succeeded by COPTIC in the third or fourth century AD.

DEURI. *Sino-Tibetan* language spoken by a small community in north-eastern India; there is a dictionary with **ASSAMESE** (2014).

DIYARI. *Pama-Nyungan* language with a few speakers in South Australia; a bidirectional wordlist with **GERMAN**, made by C. Schoknecht in 1872, was followed by comparative wordlists later in the nineteenth century.

Dravidian. Family of languages spoken mainly in southern India (whereas the unrelated *Indo-Aryan* languages are spoken mainly in the north), including **KANNADA**, **MALAYALAM**, **TAMIL**, and **TELUGU**.

DUTCH. West *Germanic* language spoken in the Netherlands and in parts of Belgium (where it is sometimes called Flemish). Dictionaries with **LATIN** since the Middle Ages; a multi-volume scholarly dictionary, *Woordenboek der Nederlandsche Taal*, was published from 1864 to 2001.

DYIRBAL. *Pama-Nyungan* language with a few speakers in Queensland; there is wordlist material in H. Nekes and E. Worms' *Australian Languages* (1953).

ECHUWABO (or Chuwabu). *Bantu* language spoken in Mozambique; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856).

Egyptian. Series of *Afroasiatic* language varieties (not to be confused with Egyptian **ARABIC**, the modern vernacular of Egypt): **OLD AND MIDDLE EGYPTIAN** (see also **CLASSICAL EGYPTIAN**) was succeeded by **LATE EGYPTIAN**, followed first by **DEMOTIC** and then by **COPTIC**.

EMAKHUWA (or Makuwa). The most widely spoken *Bantu* language of Mozambique; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856).

ENGLISH. West *Germanic* language closely related to **FRISIAN**, which developed from continental *Germanic* language varieties in the southern British Isles in and after the fifth century, and was written from the late seventh century onwards. It had spread across much of the British Isles by the sixteenth century and across the whole of the British Isles, the United States, and an extensive empire by the mid twentieth century, with marked diversification; it is now very widely used in international communication, with official status in fifty-four countries. There were glossaries of Old English from the late seventh century to the eleventh, and dictionaries of Middle English in the fifteenth; a strong, continuous tradition of English–**LATIN** and then monolingual lexicography developed from the sixteenth century onwards, the first free-standing monolingual dictionary being by R. Cawdrey (1604). The pre-eminent historical dictionary is the *Oxford English Dictionary* (1884–1928); wordlists of regional varieties were made from the eighteenth century onwards.

Eskimo-Aleut. Family of languages spoken in the Arctic from the eastern tip of Siberia, across Alaska (including the Aleutian Islands) and Arctic Canada, to Greenland, including **GREENLANDIC**, *Inuktitut*, **INUPIAQ**, and **YUP'IK**.

ESTONIAN. *Finno-Ugric* language spoken in Estonia; the first dictionary was published as part of H. Stahl's *Anführung zu der Esthnischen Sprach* in 1637.

EUDEVE. *Uto-Aztecan* language formerly spoken in northern Mexico; an early wordlist probably made by Jesuit missionaries was published in the twentieth century.

FAROESE. North *Germanic* language spoken in the Faroe Islands. The first dictionary was completed by J. Svabo in 1773 but remained in manuscript until 1966; a printed glossary of 1891 by J. Jakobsen has been followed by a twentieth-century dictionary tradition.

FINNISH. *Finno-Ugric* language spoken in Finland; the first printed dictionary was E. Schroderus' *LATIN-SWEDISH-GERMAN-Finnish Lexicon Latino-Scondicum* (1637).

Finno-Ugric. Language family including **ESTONIAN**, **FINNISH**, and **HUNGARIAN**.

FRANCOPROVENÇAL. Group of *Romance* language varieties spoken in parts of Italy, Switzerland, and France; its Helvetic varieties are being described in the *Glossaire des patois de la Suisse romande* (1924–).

FRENCH. *Romance* language with written texts distinct from **LATIN** from the ninth century; spoken and written in France and parts of neighbouring European countries, and in the overseas empire of France; also used as a language of international high culture and diplomacy; an official language in many countries. There has been a strong and complex lexicographical tradition from the Middle Ages onwards.

FRISIAN. West *Germanic* language closely related to **ENGLISH**, spoken on part of the North Sea coast of Europe and its offshore islands; a western variety is spoken in the Netherlands, and a northern variety in Germany. No substantial dictionary until the nineteenth century; 25-volume historical dictionary, *Wurdboek fan de Fryske Taal*, 1984–2011.

FRIULIAN. *Romance* language variety spoken in north-eastern Italy, with a written tradition from the fourteenth century onwards; Zamboni, *Dizionario etimologico storico friulano* (1984–7) is a scholarly dictionary.

GALICIAN. *Romance* language spoken in north-western Spain; the first wordlist (with **SPANISH**) dates from the sixteenth century, and there have been printed dictionaries since the nineteenth.

GAMILARAAY (Kamilaroi). *Pama-Nyungan* language with a few speakers in New South Wales and Queensland, currently undergoing some revitalization (see also **YUWAALARAAY**); a dictionary by W. Ridley was published in 1866, and an online dictionary by P. Austin and D. Nathan was launched in 1998.

GARO. *Sino-Tibetan* language spoken in and beyond north-eastern India, unwritten until the nineteenth century, as was its close relative **BODO**; there is a dictionary with **ASSAMESE** and **ENGLISH** (2000).

GASCON. *Romance* language variety spoken in south-western France; Old Gascon is treated in scholarly dictionaries, and there is a substantial dictionary of the modern spoken language.

- GAULISH.** *Celtic* language spoken in what is now France and northern Italy, extinct since late antiquity. An early medieval wordlist includes some onomastic material; there were learned wordlists from the sixteenth century.
- GE'EZ.** *Semitic* language of Ethiopia, extinct except in liturgical use. The first printed dictionary, J. Wemmers' *Lexicon aethiopicum*, was published in Rome in 1638.
- GERMAN.** *Germanic* language spoken in several varieties (one of which, Low German, became clearly distinct from DUTCH only in the sixteenth century) in Germany, Austria, and Liechtenstein, in all of which it is the official language, and in Switzerland, Luxemburg, and parts of Belgium, in all of which it has co-official status; used widely by German traders and settlers in east central and eastern Europe until 1945, and in some overseas possessions and diasporic settlements; written from the eighth century. A tradition of glossaries from the eighth century to the thirteenth was followed by a separate dictionary tradition from the fourteenth century onwards.
- Germanic.** Branch of the *Indo-European* language family, divided into North Germanic (DANISH, FAROESE, ICELANDIC, NORWEGIAN, SWEDISH), West Germanic (DUTCH, ENGLISH, FRISIAN, GERMAN, LUXEMBURGISH, SCOTS, YIDDISH), and East Germanic (GOTHIC) sub-branches.
- GIDABAL.** *Pama-Nyungan* language with a few speakers in New South Wales; a grammar and dictionary by B. Geytenbeek and H. Geytenbeek were published in 1971.
- GITONGA.** *Bantu* language spoken in Mozambique; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856).
- GOTHIC.** East *Germanic* language, spoken in several Gothic kingdoms in late antiquity, and on the Crimean peninsula as late as the sixteenth century. Crimean Gothic was the subject of a wordlist published by O. de Busbecq in 1589, and the Gothic of a late antique Bible translation was the subject of wordlists from F. Junius' 'Glossarium gothicum' of 1665.
- GREEK.** A branch in its own right of the *Indo-European* language family, spoken in ancient Greece and its colonies across the Mediterranean and Black Sea, then very widely in the eastern Roman empire and its successor, the Byzantine empire, and now in and beyond Greece and Cyprus. Ancient lexicographical works (the earliest extant are from the third century BC) were succeeded by a Byzantine tradition and then by learned Greek–LATIN dictionaries and their vernacular successors; wordlists of the modern language begin in the sixteenth century.
- GREENLANDIC.** *Eskimo-Aleut* language spoken in Greenland; the lexicographical tradition begins with the work of P. Egede (1750).
- GUARANI.** *Tupi-Guarani* language spoken in and beyond Paraguay; dictionaries and a grammar by A. Ruiz de Montoya were published in 1639–40.

- GUGU-YALANJI (Kuku-Yalanji).** *Pama-Nyungan* language spoken in Queensland; the first wordlist was by W. Oates and L. Oates (1964).
- GUJARATI.** New *Indo-Aryan* language of western India and southern Pakistan. Numerous bilingual English dictionaries prepared in the nineteenth century; nine-volume encyclopedic dictionary 1944–55.
- GUNIN/KWINI.** *Worrorran* language with a few speakers in Western Australia; there are twentieth- and twenty-first-century wordlists.
- Gunwinjguan.** Group of languages spoken in the Northern Territory of Australia, including ANINDILYAKWA, BURARRA-GUN-NARPTA, and NUNGGUBUYU.
- GUUGU YIMIDHIRR.** *Pama-Nyungan* language spoken in Queensland: the first wordlists were made by J. Banks and J. Cook in 1770.
- HAKKA (客語).** Variety of CHINESE spoken in different regions of south-eastern China, the island of Taiwan, and in ethnic Chinese communities in south-east Asia, Europe, and the Americas; no major dictionary before the nineteenth century.
- HARAYA.** Name given by A. de Ménétrida (1637) to a *Visayan* language (perhaps the modern Karay-a) spoken in the Philippines; the first extant wordlist was by Ménétrida.
- HAUSA.** *Afroasiatic* language spoken as a mother tongue mostly in Nigeria, Niger, Benin, Chad, Cameroon, Côte d'Ivoire, and Sudan, and widely used as a *lingua franca* in central Africa; the first dictionary was J. Schön's *Vocabulary of the Hausa Language* (1843).
- HEBREW.** *Semitic* language spoken by Jews in the eastern Mediterranean and a wider diaspora until late antiquity and thereafter used as a written language of the Jewish scriptures and scriptural scholarship; regained its societal functions at the turn of the twentieth century in the present-day Israel and became one of three official languages of the British Mandate for Palestine in 1922 (with ENGLISH and ARABIC), and one of two official languages of Israel in 1948 (with Arabic until 2018). Learned native lexicographical tradition from the ninth century onwards, supplemented from the sixteenth century by learned Hebrew–LATIN dictionaries made by and for gentiles, and by a new tradition of monolingual and bilingual dictionaries from the beginning of the twentieth century.
- HILIGAYNON.** *Visayan* language spoken in the Philippines; the first extant wordlist was by A. de Ménétrida (1637).
- HINDI-URDU (formerly called Hindustani).** New *Indo-Aryan* language, originally known as *hindavī*, denoting a spectrum of grammatically diverse forms spoken across northern India. Urdu (written in the Perso-Arabic script) is the official language of Pakistan, and Modern Standard Hindi (written in the *devanagari* script) an official language of India. Both derive their grammars from the spoken language of the Delhi region. The earliest dictionaries to include Hindi-Urdu

terms are from the fifteenth century; multilingual vocabularies in verse may have emerged as early as the fourteenth century. Four-volume comprehensive Hindi dictionary, 1916–28, eleven-volume revised edition 1965–75; 22-volume historical dictionary of Urdu 1973–2010.

HITTITE. *Indo-European* language formerly spoken in Anatolia and attested in texts from the sixteenth to the thirteenth century BC; occasional SUMERIAN–AKKADIAN lexicographical texts have an additional translation column in Hittite.

HOPÍ. *Uto-Aztecan* language spoken in Arizona, New Mexico, and Utah; the *Hopi Dictionary/Hopikwa Laváytutuveni* (1998) produced by the Hopi Dictionary Project is the most impressive and scholarly looking of all twentieth-century North American indigenous language dictionaries.

Huarpean. Family of extinct languages formerly spoken in what is now Argentina and Chile, including ALLENTIAC and MILLCAYAC.

HUNGARIAN. *Finno-Ugric* language, the official language of Hungary, and spoken by significant communities in neighbouring countries such as Romania. Treated in a polyglot dictionary of 1585 in the Calepino tradition, and in free-standing dictionaries from the *Dictionarium Latino–Ungaricum* of S. Molnár (1604) onwards; there is a seven-volume dictionary of modern usage, as well as historical, etymological, and regional dictionaries.

HURON. *Iroquoian* language formerly spoken in eastern North America; the first wordlist was published in G. Sagard, *Grand voyage du pays des Hurons* (1632).

HURRIAN. Hurro-Urartian language spoken in northern Mesopotamia in the second and first millennia BC; occasional SUMERIAN–AKKADIAN lexicographical texts have an additional translation column in Hurrian.

IBANAG. *Malayo-Polynesian* language spoken in the Philippines; the first extant wordlist, compiled by J. Bugarín no later than 1676, was published in 1854.

ICELANDIC. *Germanic* language which developed from continental Germanic varieties after the settlement of Iceland in the ninth century, the first extant manuscripts of a distinctively Icelandic variety dating from the twelfth century; now the official language of Iceland. Glosses are extant from c. 1200; the first printed dictionaries were Magnús Ólafsson's *Specimen lexici runici* (of the medieval language; compiled before 1634 and published 1650–1) and Guðmundur Andrússon's *Lexicon Islandicum* (of the medieval and contemporary language; completed 1654 and published 1683).

IGBO. *Niger-Congo* language spoken mostly in Nigeria; a few words were included in the comparative wordlist of C. Oldendorp (1777), and the first substantial dictionary was that of S. Crowther, assisted by J. Schön, published in 1882–3.

ILOCANO. *Malayo-Polynesian* language spoken in the Philippines. There was a literate tradition before European contact; the first wordlist was compiled by P. de Vivar, no later than 1771.

Indo-Aryan (Indic). Branch of the *Indo-European* language family spoken in India (but the unrelated *Dravidian* languages are spoken in much of southern India), Bangladesh, Pakistan, Nepal, and Sri Lanka, and in diasporic communities. Ancient traditions of literacy make it possible to describe (overlapping) chronological stages: Old Indo-Aryan (SANSKRIT); Middle Indo-Aryan (PALI and the PRAKRIT varieties); New Indo-Aryan (including ASSAMESE, BENGALI, GUJARATI, HINDI-URDU, KONKANI, MARATHI, ROMANI, and SINHALESE).

Indo-European. Major language family of western Eurasia and parts of South Asia, including the *Baltic*, *Celtic*, *Germanic*, *Indo-Aryan*, *Iranian*, *Italic*, *Romance*, and *Slavic* branches, as well as ALBANIAN, ARMENIAN, GREEK, and HITTITE.

Inuktitut. Group of *Eskimo-Aleut* languages spoken in the Arctic north of Canada and in north-west Greenland: Eastern Canadian Inuktitut and Western Canadian Inuktitut are distinct. Dictionaries have been produced since the nineteenth century.

INUPIAQ. *Eskimo-Aleut* language spoken in Alaska and the Northwest Territories; there is a dictionary of 1970.

Iranian. Branch of the *Indo-European* language family including AVESTAN, PASHTO, PERSIAN, TAJIK, and YASSIC.

IRISH. *Celtic* language spoken in Ireland, and with official status in the Republic of Ireland; inscriptions from late antiquity, and a rich literary tradition from the late sixth century onwards. The LATIN–Irish glossographical tradition begins in the seventh century, and there are very early monolingual glossaries; the first printed dictionary is M. Ó Cléirigh's *Foclóir nó sanasán nua* (1643).

Iroquoian. Language family spoken (though many of its members are now extinct or moribund) in eastern North America, including HURON, MOHAWK, and ONONDAGA.

ITALIAN. Group of *Romance* language varieties, now spoken in the republics of Italy and San Marino and in Switzerland, with official status in all three; the first written texts clearly distinct from LATIN are of the tenth century, and, although the lasting prestige of thirteenth- and fourteenth-century literary texts in the TUSCAN variety helped establish this as a literary standard, many other Italo-Romance varieties have been the subjects of dictionaries, for instance FRIULIAN and LADIN; although Sardinia is politically part of Italy, SARDIC is a distinct language. After some earlier glosses and short wordlists, the main lexicographical tradition began in the fourteenth century, a landmark being the *Vocabolario degli Accademici della Crusca* (1612).

Italic. Branch of the *Indo-European* language family including LATIN and other extinct languages of what is now the Italian peninsula.

IYURA (Eora). *Pama-Nyungan* language formerly spoken in New South Wales; the first wordlists were made in the late eighteenth century.

JABIRRJABIRR. *Nyulnyulan* language formerly spoken in Western Australia; a wordlist is included in H. Nekes and E. Worms' *Australian Languages* (1953).

JAPANESE. Japonic language spoken by a large population in Japan. The writing system uses Chinese characters, or sinographs (*kanji*), and syllabic characters derived from them (*kana*). Lexicographical activity is reported as early as the seventh century, and extant CHINESE dictionaries were compiled in Japan in the ninth century, with the first Japanese glosses by the end of the century; Japanese dictionaries were being compiled by the eleventh century, in a tradition which until the nineteenth century was strongly interested in the reading of sinographs; this tradition was supplemented by missionary dictionaries with PORTUGUESE and LATIN in the late sixteenth and early seventeenth centuries.

Jarrakan (Djeragan). Small family of languages with a few speakers in Western Australia; Jarrakan languages are represented briefly in H. Nekes and E. Worms' *Australian Languages* (1953).

JAWI. *Nyulnyulan* language with a few speakers in Western Australia; there is a wordlist by W. Bird (1915).

JUKUN. *Nyulnyulan* language formerly spoken in Western Australia; represented briefly in H. Nekes and E. Worms' *Australian Languages* (1953).

JURCHEN. Tungusic language related to MANCHU, formerly spoken by the Jurchen people, to whom the emperors of the Jin dynasty which ruled China from 1114 to 1234 belonged; bilingual dictionaries with CHINESE were produced under the M'ing dynasty (1368–1644), but appear not to have influenced the subsequent lexicography of Manchu.

JUWALINY. Variety of WALMAJARRI; a wordlist was made in the twentieth century by K. McKelson.

KAJKAVIAN. Variety of CROATIAN.

KALISPEL. Salishan language with a few speakers in Montana and Washington state; a dictionary by J. Giorda and other Jesuit missionaries was published in 1877–9.

KAMBA (or Kikamba). *Bantu* language spoken in Kenya; it was documented in J. Krapf's *Vocabulary of Six East-African Languages* (1850).

KANNADA. *Dravidian* language spoken in southern India; there are texts from the first millennium AD, and texts from before 1500 explain old Kannada words and SANSKRIT words; there are large modern dictionaries.

KAQCHIKEL. *Mayan* language spoken in what is now Guatemala, closely related to K'ICHE' and TZ'UTUJIL; the first wordlist is that of D. de Vico, made no later than 1555.

KARAJARRI. *Pama-Nyungan* language with a few speakers in Western Australia; a wordlist was made in the twentieth century by K. McKelson.

KARBI. *Sino-Tibetan* language spoken in north-eastern India; there were missionary wordlists in the late nineteenth century, and there is a dictionary with ASSAMESE and ENGLISH (2002).

- KAURNA.** *Pama-Nyungan* language formerly spoken in South Australia, currently undergoing some revitalization; a grammar and vocabulary were published by C. Teichelmann and C. Schürmann in 1840.
- K'ICHE'.** *Mayan* language spoken in what is now Guatemala, closely related to KAQCHIKEL and TZ'UTUJIL; there are K'iche' words in the Kaqchikel–SPANISH wordlist of D. de Vico, made no later than 1555.
- KIKONGO.** *Bantu* language, now spoken in the Democratic Republic of the Congo, the Republic of the Congo, and Angola; the first extant dictionary is the *Vocabularium Latinum, Hispanicum, et Congense* (1648).
- KIMBUNDU.** *Bantu* language, spoken by the Ambundu people of northern Angola; the first dictionary is that of B. da Canicatti (1804).
- KIMWANI.** *Bantu* language spoken in Mozambique; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856).
- KIPOKOMO (Kipfokomo, Pokomo).** *Bantu* language spoken in Kenya; it was documented in J. Krapf's *Vocabulary of Six East-African Languages* (1850).
- KIRIRI (Kariri).** Language isolate formerly spoken in Brazil; a dictionary of 1696 or earlier by L. Mamiani is now lost.
- KONKANI.** New *Indo-Aryan* language spoken in south-western India; there are records from before the sixteenth century, and a dictionary with PORTUGUESE was compiled by D. Ribeiro in 1626.
- KOREAN.** Language spoken in the Korean peninsula, where the high status of CHINESE as a written language delayed the beginnings of vernacular lexicography until the fifteenth century (though Chinese dictionaries were printed in Korea from the tenth century onwards); bilingual dictionaries with western European languages began to be compiled in the nineteenth century, and the first monolingual dictionary, Sim Uilin's *Botonghaggyo Joseon-eo Sajeon*, was published in 1925.
- KOYUKON.** *Athabaskan* language with a few speakers in Alaska; a major dictionary founded on work of the late nineteenth and early twentieth centuries was published in 2000.
- KRIOL.** ENGLISH-lexifier creole spoken in several parts of Australia; the first dictionary, by J. Sandefur and J. Sandefur, was published in 1979.
- KUKATJA.** *Pama-Nyungan* language spoken in Western Australia; there are lexicographical texts from the second half of the twentieth century.
- Kurdish.** Group of *Iranian* languages spoken in several countries in western Asia; the literary tradition begins around the fifteenth century AD; there is an ITALIAN–Kurdish dictionary by M. Garzoni (1787).
- KUYANI.** *Pama-Nyungan* language formerly spoken in South Australia; a wordlist was included in J. Reuther's comparative work of the early twentieth century.

LADIN. *Romance* language variety spoken by a small population in the north of Italy; Kramer, *Etymologisches Wörterbuch des Dolomitenladinischen* (1988–98), is a scholarly dictionary.

LAKOTA. *Siouan* language spoken in the United States and Canada; there have been several dictionaries in the late twentieth and twenty-first centuries.

LATE EGYPTIAN. Variety of *Egyptian* used as the idiom of *belles lettres* and documentary texts in the later second and early first millennium BC (and in vernacular spoken use at least in the early part of this period), while **CLASSICAL EGYPTIAN** remained in use for religious purposes, calling for wordlists glossing Classical Egyptian in Late Egyptian.

LATIN. *Italic* language originating on the Italian peninsula; widely spoken and written in the territories of the Roman republic and then in the Roman empire, particularly its western part, in antiquity (after which spoken varieties sometimes called Popular Latin or Vulgar Latin developed into the *Romance* languages); very widely used in worship, scholarship, literature, and record-keeping in western Europe throughout and after the Middle Ages; in declining use in these functions thereafter. There has been a strong, complex lexicographical tradition from antiquity onwards, and Latin appeared in innumerable wordlists, monolingual and with all the written vernaculars, in early modern Europe.

LATVIAN. *Baltic* language spoken principally in Latvia. The first texts are from the sixteenth century, and the first dictionary was G. Mancelius' *Latvian–GERMAN Lettus* (1638); K. Mühlenbach and J. Endzelīns, *Lettisch–deutsches Wörterbuch* (4 vols., 1923–32, with two-vol. supplement by Endzelīns and E. Hauzenberga, 1934–46), is the most important modern dictionary.

LAW FRENCH. See **ANGLO-NORMAN**.

LITHUANIAN. *Baltic* language spoken in Lithuania; formerly one of the languages of the Polish-Lithuanian Commonwealth and of the Duchy of Prussia. The first texts are from the sixteenth century, and the first (**POLISH–LATIN–Lithuanian**) dictionary, by K. Sirvydas, is perhaps earlier than 1620; the most comprehensive modern dictionary is the monolingual *Lietuvių kalbos žodynas* (20 vols., 1941–2002).

LOKONO. *Arawakan* language spoken in northern coastal areas of South America; a **SPANISH–Lokono** wordlist dated 1765 is transcribed as part of the *Colección Mutis*.

LULE. *Lule-Vilela* language formerly spoken in what is now northern Argentina; the first printed wordlist was by A. Machoni de Cerdeña (1732).

Lule-Vilela. Small group of extinct South American languages including **LULE** and **TONOCOTÉ**.

LURITJA. *Pama-Nyungan* language, spoken in the Northern Territory of Australia and in Western Australia (see also **PINTUPI / LURITJA**); there are twentieth- and twenty-first-century dictionaries.

LUXEMBURGISH (Letzeburgisch). West Germanic language with official status beside FRENCH and GERMAN in Luxembourg; sometimes regarded as a dialect of German. The first printed dictionary was by J.-F. Gangler (1847); a five-volume *Luxemburger Wörterbuch* was published 1950–77.

MACEDONIAN. South Slavic language spoken in the (Former Yugoslav) Republic of Macedonia; the vernacular on which OLD CHURCH SLAVIC was based has features in common with present-day BULGARIAN and Macedonian. The four-language dictionary of Daniil of Moscopole (1794) presents the Macedonian vernacular of the Bitola region as Bulgarian; a modern dictionary is the *Rečnik na makedonskiot jazik* (3 vols., 1961–6).

MAHICAN (Mohican). Algonquian language formerly spoken in what is now the north-eastern United States; a dictionary was compiled by J. Schmick in the mid eighteenth century.

MALAY. Malayo-Polynesian language very widely spoken by native speakers and as a *lingua franca* in Malaysia, Indonesia (where it is known as Bahasa Indonesia or Indonesian), and elsewhere; inscriptions in a variety of the Grantha script go back to the seventh century AD, and the Devanagari script was used for official purposes from the seventh to the fourteenth centuries, succeeded by Arabic script. A wordlist in Chinese characters probably dates from the fifteenth century; an ITALIAN–Malay wordlist was compiled by A. Pigafetta around 1521, and a DUTCH–Malay dictionary by F. de Houtman was printed in 1603, these being in the Roman alphabet, which is now widely used in Malaysia and Indonesia. There are a number of substantial modern dictionaries.

MALAYALAM. Dravidian language spoken in southern India, and written in a script which evolved from the Grantha script; a literary tradition had begun by the fourteenth century AD, and there were a number of missionary wordlists; a major historical dictionary was begun in 1965.

Malayo-Polynesian. Language family forming the great majority of the Austronesian family, including Hawaiian, Malay, and the Philippine languages BICOL, CEBUANO, HILIGAYNON, IBANAG, ILOCANO, PAMPANGO, TAGALOG, and Visayan.

MANCHU. Tungusic language related to JURCHEN, and now spoken, as a second language, by a minority of ethnic Manchus in China; the Chinese imperial family from 1644 onwards were Manchu-speakers, and Manchu–CHINESE and Chinese–Manchu dictionaries were printed from 1683 onwards, with a monolingual dictionary in 1708.

MANDARIN (*pǔtōnghuà* 普通話 ‘common language’). Variety of CHINESE used as the standard language of the People’s Republic of China (founded in 1949); it is the major language of modern Chinese lexicography. Not to be confused with the MANDARIN TOPOLECTS or with either of two historical varieties: Mandarin (*guānhuà* 官話 ‘official language’), the *lingua franca* of Chinese government officials before the twentieth century (first records in the fifteenth century, first

- dictionaries in the sixteenth century), and Mandarin (*guóyǔ* 國語 ‘officials’ language), the standard language of the Republic of China (founded 1912) (first standard dictionaries in the early twentieth century).
- MANDARIN TOPOLECTS** (*guānhuà fāngyán* 官話方言). Group of topolects of CHINESE spoken in north-eastern, north-western, and south-western China; first lexicographic sources in the eleventh century.
- MANGALA**. *Pama-Nyungan* language formerly spoken in Western Australia; a wordlist was made in the twentieth century by K. McKelson.
- MAPUDUNGUN (Mapuche)**. Mapudungu language spoken in parts of Chile and Argentina; the first printed wordlist was by L. de Valdivia (1606).
- MARATHI**. New *Indo-Aryan* language of western India. Numerous bilingual English dictionaries prepared in the nineteenth century; seven-volume encyclopedic dictionary 1932–8.
- MARTU WANGKA**. *Pama-Nyungan* language spoken in Western Australia (see also YULPARIJA); there is a dictionary by J. Marsh (1992).
- MASSACHUSETT (Natick)**. *Algonquian* language formerly spoken in what is now the north-eastern United States, and now revived after a period of dormancy; a vocabulary was made by J. Cotton before 1756, and a nineteenth-century dictionary was compiled by J. Trumbull on the basis of seventeenth-century texts.
- MATLATZINCA**. *Oto-Manguean* language with a few speakers in Mexico; a wordlist dated 1557 was written in the margins of a copy of A. de Molina’s NAHUATL dictionary by A. de Castro.
- MAYA (Yucatec Maya)**. *Mayan* language spoken in Mexico and Belize; the SPANISH–Maya dictionaries attributed to A. de Solana and G. González de Nájera date from about 1580.
- Mayan**. Family of languages spoken in Mesoamerica and central America, including KAQCHIKEL, K’ICHE’, MAYA, POQOMCHI’, TZELTAL, TZOTZIL, and TZ’UTUJIL.
- MAYO**. *Uto-Aztecan* language spoken in Mexico; a dictionary was published by the Summer Institute of Linguistics in 1962.
- MAZAHUA**. *Oto-Manguean* language spoken in Mexico; the first printed wordlist is that of D. de Nájera Yanguas (1637).
- MBARRUMBATHAMA (Lamu-Lamu)**. *Pama-Nyungan* language with a few speakers in Queensland; there is a dictionary app.
- MIAMI-ILLINOIS**. *Algonquian* language formerly spoken in the United States and currently undergoing some revitalization; early dictionaries were made by Jesuit missionaries, the first being that of P.-F. Pinet (1696–1700).
- MILLCAYAC**. *Huarpean* language formerly spoken in what is now Argentina; it is documented in a grammar and wordlist of 1607 by L. de Valdivia.

- MĪN** (MĪNYŪ 閩語) (southern MĪN is internationally also known as Hokkien, and in Taiwan as TÁIYŪ or Hoklo). Variety of CHINESE with two major subgroups (Northern and Southern), spoken in most parts of the Chinese provinces of FÚJIÀN and Hǎinán, in the island of Taiwan, and among overseas Chinese in south-east Asia; first handwritten dictionaries in the seventeenth century; printed lexicography started in the nineteenth century.
- Mixe.** Group of Mixe-Zoquean languages spoken in Mexico; a wordlist was published by A. de Quintana in 1733.
- Mixtec.** Group of Oto-Manguean languages spoken in Mexico; the first printed dictionary is that of F. de Alvarado (1593).
- MOHAWK.** *Iroquoian* language spoken in parts of Canada and the United States; a dictionary was compiled by J. Bruyas in the seventeenth or early eighteenth century.
- MONGOLIAN.** Mongolic language now spoken in Mongolia and the Inner Mongolia Autonomous Region of China, but formerly also spoken across the vast Mongol empire; written from the thirteenth century, first in an adaptation of the script developed for UYGHUR, and then in the so-called 'Phags-pa script, influenced from the script developed for TIBETAN; there are Mongolian–ARMENIAN and ARABIC–Mongolian wordlists of the thirteenth century, and a CHINESE–Mongolian wordlist of the fourteenth.
- MONO.** *Uto-Aztecan* language with a few speakers in California, currently undergoing some revitalization; a dictionary of Western Mono by R. Bethel and others was published in 1984.
- MONTAGNAIS.** *Algonquian* language spoken in eastern Canada; missionary lexicography began at the end of the seventeenth century.
- MORROBOLAM (Umbuygamu).** *Pama-Nyungan* language formerly spoken in Queensland; there is a dictionary app.
- MOXO.** *Arawakan* language spoken in Bolivia; the first printed wordlist was part of the anonymous *Arte y Vocabulario de la lengua Morocosi* of 1699.
- MURUWARI.** *Pama-Nyungan* language formerly spoken in Queensland and New South Wales; there is a dictionary by L. Oates (1992).
- Muskogean.** Family of languages spoken in and beyond the south-eastern United States, including CHOCTAW and MUSKOGEE.
- MUSKOGEE (Creek).** *Muskogean* language spoken in Oklahoma and Florida; a dictionary by R. Loughridge and D. Hodge was published in 1890.
- MUTSUN.** *Utian* language formerly spoken in California, and currently undergoing some revitalization; a dictionary by F. Arroyo de la Cuesta was published in 1862.
- Na-Dene (Athabaskan-Eyak-Tlingit).** Family of languages spoken in North America, including the *Athabaskan* languages and TLINGIT.

- NAHUATL.** Language spoken in central Mexico; the variety spoken in Tenochtitlan had prestige status in the Aztec empire. An anonymous SPANISH–LATIN–Nahuatl dictionary of c. 1545 remained in manuscript during the colonial period, as did the wordlist in the Nahuatl grammar of A. de Olmos (1547), so that the first published dictionary was that of A. de Molina (1555).
- NARRAGANSETT.** *Algonquian* language formerly spoken in what is now Rhode Island; R. Williams' *Key into the Language of America* (1643) is a phrasebook.
- NAVAJO.** *Athabaskan* language spoken in the south-western United States; substantial missionary dictionaries were produced at the beginning of the twentieth century. Monumental dictionaries were published in 1987 and 1992.
- NDAU (or Cindau).** *Shona* language spoken in Mozambique and Zimbabwe; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856).
- NDEBELE.** *Bantu* language spoken in Zimbabwe and Botswana; the first dictionary is that of W. Elliott (1897).
- NGAMINI.** *Pama-Nyungan* language formerly spoken in South Australia; a wordlist was included in J. G. Reuther's comparative work of the early twentieth century.
- NGARRINDJERI.** *Pama-Nyungan* language formerly spoken in South Australia; a dictionary of the variety *RAMINDJERI* was published by H. Meyer in 1843.
- NGUMBARL.** *Nyulnyulan* language formerly spoken in Western Australia; represented briefly in H. Nekes and E. Worms' *Australian Languages* (1953).
- Niger-Congo.** Major language family or phylum, occupying a greater area of the African continent than any other (cf. *Afroasiatic* and *Nilo-Saharan*), its major branch being the *Bantu* languages; it also includes *BAMBARA*, *IGBO*, *WOLOF*, and *YORUBA*. It probably constitutes the largest group of languages in the world, closely followed by Austronesian (for which see *Malayo-Polynesian*).
- Nilo-Saharan.** Major language family or phylum of the northern half of Africa (cf. *Afroasiatic* and *Niger-Congo*), including the *Nubian* languages.
- NIMANBURRU.** *Nyulnyulan* language formerly spoken in Western Australia; a wordlist is included in H. Nekes and E. Worms' *Australian Languages* (1953).
- NOONGAR (Nyungar).** *Pama-Nyungan* language with a few speakers in Western Australia (*WADJUK* is a variety); early dictionaries were compiled by G. Grey (1839) and G. Moore (1842).
- NORN.** Extinct North *Germanic* language formerly spoken on the Orkney and Shetland archipelagos of the British Isles. The most significant wordlist is by J. Jakobsen (1908–21).
- NORWEGIAN.** North *Germanic* language with official status in Norway (where until 1885 *DANISH* was the sole official language); there are now two varieties, *Bokmål* (formerly *Riksmål*), which grew directly out of the Danish written tradition, and *Nynorsk* (formerly *Landsmål*), based on Norwegian dialects. The first printed

dictionary was by C. Jenssøn (1646); Nynorsk was codified by I. Aasen in dictionaries of 1850 and 1873.

Nubian. Group of Nilo-Saharan languages spoken in what is now southern Egypt and Sudan; Christian texts in Old Nubian of the eighth to fifteenth centuries are extant, and an ITALIAN–Nubian wordlist based on the spoken Kenzi and Nobiin varieties was made by A. Carradori around 1635.

NUNGGUBUYU. *Gunwinjguan* language with a few speakers in the Northern Territory of Australia, currently undergoing some revitalization; there are two dictionaries of the late twentieth century.

NYANGUMARTA. *Pama-Nyungan* language spoken in Western Australia; there is a dictionary by K. McKelson (1989).

NYIKA (Chinyika, Kinyika). *Bantu* language spoken mainly in Malawi and Zambia; it was documented in J. Krapf's *Vocabulary of Six East-African Languages* (1850).

NYIKINA. *Nyulnyulan* language with a few speakers in Western Australia; a wordlist is included in H. Nekes and E. Worms' *Australian Languages* (1953).

NYULNYUL. *Nyulnyulan* language formerly spoken in Western Australia; wordlists with FRENCH and ENGLISH were made from the end of the nineteenth century onwards.

Nyulnyulan. Family of languages formerly spoken in Western Australia, but now with very few speakers in total, including BARDI, JABIRRJABIRR, JAWI, JUKUN, NGUMBARL, NIMANBURRU, NYIKINA, NYULNYUL, and YAWURU.

NYUNGWE (Cinyungwe). *Bantu* language spoken in Mozambique; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856), and dictionaries by V. Courtois were completed in 1889 and published in 1899–1900.

OCCITAN. *Romance* language spoken in southern France. A medieval literary tradition is well documented in scholarly dictionaries of the nineteenth century onwards; the first free-standing printed dictionary was S. Pellas' *Dictionnaire provençal et françois* (1723), and there have been important subsequent dictionaries such as the *Trésor dou Félibrige* of F. Mistral (1879–86), but a modern scholarly dictionary of modern Occitan is lacking.

OJIBWA (Ojibwe). *Algonquian* language spoken in Canada and the United States; nineteenth-century missionary work has been followed by twentieth- and twenty-first-century dictionaries.

OLD AND MIDDLE EGYPTIAN. Variety of *Egyptian* spoken and written in the third (Old Egyptian) and the first half of the second millennium BC (Middle Egyptian). The vestiges of OLD EGYPTIAN are rather sparse and stem mostly from tomb inscriptions. MIDDLE EGYPTIAN as the literary language of the Middle Kingdom (c. 2119–1794 BC) was used later as CLASSICAL EGYPTIAN for official and religious literature exclusively. While OLD AND MIDDLE EGYPTIAN form the older stratum of the idiom, LATE EGYPTIAN, DEMOTIC, and COPTIC comprise its younger development.

- OLD PRUSSIAN.** *Baltic* language formerly spoken in East Prussia (now north-eastern Poland and the Kaliningrad Oblast of Russia); there are manuscript wordlists as early as c. 1400.
- ONONDAGA.** *Iroquoian* language spoken in Ontario and New York state; the first missionary dictionary is of the seventeenth century.
- O'ODHAM (Tepehuan, Papago).** *Uto-Aztecan* language spoken in Arizona and northern Mexico; early wordlists are said to have been made by Jesuit missionaries, and there has been twentieth-century lexicography with **ENGLISH**.
- OPATA.** *Uto-Aztecan* language formerly spoken in northern Mexico; an extensive dictionary was made in 1702.
- OROMO.** *Afroasiatic* language widely spoken in Ethiopia, Kenya, and Somalia; a variety was documented under the name of Galla in a dedicated dictionary of 1842 and a polyglot dictionary of 1850, both by J. Krapf.
- OTOMACO.** Language formerly spoken along the Orinoco River in what is now Venezuela; an eighteenth?-century Otomaco-**SPANISH** wordlist is transcribed as part of the *Colección Mutis*.
- Oto-Manguean.** Family of languages spoken in Mexico (and formerly spoken elsewhere in central America), including **MATLATZINCA**, **MAZAHUA**, *Mixtec*, **OTOMI**, and **ZAPOTEC**.
- Otomi.** Group of *Oto-Manguean* languages spoken in Mexico; glosses in Otomi were added to copies of printed dictionaries of **NAHUATL** and **TARASCAN** in the second half of the sixteenth century.
- PALI.** Middle *Indo-Aryan* literary language variety, in which the canon of Theravāda Buddhism is preserved; the first lexicon was compiled in the thirteenth century, but a tradition of lexicological writing had begun before 500 AD.
- Pama-Nyungan.** Family of languages formerly spoken in much of Australia, and with a number of surviving members; the family includes **ARABANA**, **ARRERENTE**, **AWABAKAL**, **BARNGARLA**, **DIYARI**, **DYIRBAL**, **GAMILARAAY**, **GIDABAL**, **GUGU-YALANJI**, **GUUGU YIMIDHIRR**, **IYURA**, **KARAJARRI**, **KAURNA**, **KUKATJA**, **KUYANI**, **LURITJA**, **MANGALA**, **MARTU WANGKA**, **MBARRUMBATHAMA**, **NGAMINI**, **NGARRINDJERI**, **NOONGAR**, **NYANGUMARTA**, **PINTUPI/LURITJA**, **PITJANTJATJARA**, **RIMANGGUDINHMA**, **THIRRARI**, **WADJUK**, **WALMAJARRI**, **WANGKANGURRU**, **WARLPIRI**, **WESTERN DESERT LANGUAGE**, **WIK MUNKAN**, **WIRADJURI**, **YANDRUWANDHA**, **YAWARRAWARRKA**, **YORTA YORTA**, **YULPARIJA**, and **YUWAALARAAY**.
- PAMPANGO (Kapampangan).** *Malayo-Polynesian* language spoken in the Philippines; the first wordlist was by D. de Bergaño (1732).
- PASHTO.** *Iranian* language widely spoken in Afghanistan and Pakistan; there are nineteenth- and twentieth-century dictionaries with **ENGLISH**, and major dictionaries of the later twentieth century with **RUSSIAN**.

PASSAMAQUODDY. *Algonquian* language with a few speakers in Maine and New Brunswick, currently undergoing some revitalization; an important dictionary by P. LeSourd was published in 1984.

PERSIAN (Farsi). *Iranian* language first attested in inscriptions of the sixth and fifth centuries BC, and widely used as a *lingua franca* and language of high culture in successive empires in south-western Asia and the north of the Indian subcontinent from antiquity to the nineteenth century; now has official status in Afghanistan (where it is called *darī*), Iran, and Tajikistan. There are extant Middle Persian wordlists, the Persian–*AVESTAN* ‘*Frahang ī oīm*’ and the Persian–*ARAMAIC* ‘*Frahang ī pahlawīg*’. The lexicography of New Persian, the variety of the language that has been in use for the last millennium (formed in part by an influx of Arabic words into Middle Persian) begins with the *Lughat-i furs* of Asadī Tūsī (c. 1066; the lost *Risālah* of Sughdī was earlier); a tradition of Persian lexicography flourished in India, the earliest extant work being the *Farhang-i Qavvās* of Fakhr-al-Dīn Mubārakshāh Qavvās Ghaznawī (late thirteenth or early fourteenth century); there was also a learned tradition of Persian lexicography in western Europe, the first substantial dictionary being that of J. Golius, published in 1669.

PIMA BAJO. *Uto-Aztecan* language spoken in northern Mexico; an early wordlist probably made by Jesuit missionaries was published in the twentieth century.

PINTUPI/LURITJA. *Pama-Nyungan* language spoken in the Northern Territory of Australia and in Western Australia; there are twentieth-century dictionaries.

PITJANTJATJARA. *Pama-Nyungan* language spoken in central Australia; wordlists were made by J. Love in the 1940s.

POLABIAN. West *Slavic* language formerly spoken near the River Elbe in what is now Germany; the westernmost of the Slavic languages; wordlists were made as it was becoming extinct in the seventeenth and eighteenth centuries.

POLISH. West *Slavic* language, formerly widely used in the Commonwealth of Poland-Lithuania, and now with official status in Poland. The lexicographical tradition begins in the fifteenth century with wordlists of Polish and *LATIN*; the first monolingual dictionary was S. Linde’s *Słownik języka polskiego* (6 vols., 1807–14); an important modern dictionary is W. Doroszewski, *Słownik języka polskiego* (II vols., 1958–69), and there are also major historical dictionaries.

POQOMCHI’. *Mayan* language spoken in Guatemala; a manuscript wordlist with *SPANISH* and some *LATIN* (c. 1608) is attributed to D. de Zúñiga Marroquín.

PORTUGUESE. *Romance* language spoken in Portugal, Brazil, and other countries formerly part of the Portuguese empire; significant dictionaries with *LATIN* were produced from 1551 onwards, followed by major monolingual dictionaries in Portugal from the eighteenth century onwards and in Brazil more recently, and the language was widely used in missionary lexicography in South America, parts of Africa, and East Asia.

PRAKRIT. Group of Middle *Indo-Aryan* languages, several of which were used in canonical and other writings of Jainism; the first lexicon was compiled in the tenth century, but a tradition of lexicological writing had begun before 500 AD.

QUECHUA. Quechuan language widely spoken in and beyond the Andes; the first printed dictionary was by D. de Santo Tomás (1560); regional varieties were also documented, for instance by J. de Velasco.

RAMINDJERI. Variety of NGARRINDJERI; a dictionary by H. Meyer was published in 1843.

RIKANGGUDINHMA. *Pama-Nyungan* language formerly spoken in Queensland; there is a dictionary app.

Romance. Branch of the *Indo-European* language family, comprising those languages descended from LATIN; it can be divided into Gallo-Romance (including FRANCOPROVENÇAL, FRENCH, GASCON, and OCCITAN), Italo-Romance (including FRIULIAN, ITALIAN, LADIN, and ROMANSH), Ibero-Romance (including CATALAN, GALICIAN, PORTUGUESE, and SPANISH), ROMANIAN, and SARDIC.

ROMANI. New *Indo-Aryan* language, varieties of which are spoken in several parts of Europe by the Roma; wordlists were made from the early sixteenth century onwards, and a scholarly dictionary was published as the second volume of A. Pott's *Die Zigeuner in Europa und Asien* (1844–5).

ROMANIAN. *Romance* language spoken in Romania and Moldova. A manuscript lexicographical tradition began in the sixteenth century, but there was no substantial printed dictionary until the Romanian–LATIN–HUNGARIAN–GERMAN *Lesicon românescu–lătinescu–ungurescu–nemțescu* of 1825; there is a nineteen-volume academy dictionary, and a strong tradition of etymological lexicography.

ROMANSH. *Romance* language spoken in the Swiss canton of Graubünden. There were manuscript wordlists in the seventeenth century, and the first printed wordlist, giving equivalents for ITALIAN headwords in the Sursilvan and Surmiran dialects of Romansh, appeared as part of the *Fundamenti principali della lingua retica, o griggiuna* of F. da Sale (1729).

RONGA (or Shironga). *Bantu* language spoken in Mozambique; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856), and free-standing dictionaries were published in the twentieth century.

RUSSIAN. East *Slavic* language spoken in the lands of the Russian Federation and elsewhere, for instance by many Jews of Russian origin in Israel. The vernacular language of certain medieval East Slavic wordlists is better called RUTHENIAN than Russian, but there were bilingual wordlists of Russian with GERMAN from the sixteenth century onwards and with ENGLISH from the late sixteenth century; major monolingual printed dictionaries, notably the *Slovar' Akademii rossijskoj* published by the Imperial Russian Academy, appeared in the eighteenth century, since when there has been a rich lexicographical tradition.

- RUTHENIAN.** East *Slavic* language variety used in writings of the medieval and early modern periods, and based on early spoken varieties of Belorussian and UKRAINIAN; L. Zyzanij's *Leksis* (1596) was a printed wordlist translating CHURCH SLAVIC words into Ruthenian.
- SALINAN.** Language of uncertain genetic affinities, formerly spoken in California; a dictionary was compiled before 1808 by B. Sitjar.
- SANSKRIT.** Old *Indo-Aryan* language, extant in a rich corpus of texts composed from c. 1500 BC onwards; a learned lexicographical tradition, directed in particular at the exegesis of sacred texts, had begun by c. 500 BC, and continues to the present day.
- SARDIC (Sardinian).** *Romance* language spoken in Sardinia, with a written tradition from the eleventh century onwards; Wagner, *Dizionario etimologico sardo* (²2008), and Pittau, *Nuovo vocabolario della lingua sarda* (2014), are modern scholarly dictionaries.
- SCOTS.** *Germanic* language variety spoken in parts of Scotland and Ulster; sometimes regarded as a dialect of ENGLISH, to which it is closely related. A. Duncan published a LATIN–Scots wordlist in 1595 and T. Ruddiman published a Middle Scots–English wordlist in 1710; the first scholarly dictionary was that of J. Jamieson (1808, supplement 1825).
- Semitic.** Branch of the *Afroasiatic* language family spoken in western Asia and in north and north-eastern Africa, including AKKADIAN, the *Ancient South Arabian* languages; ARABIC, *Aramaic*, GE'EZ, HEBREW, and UGARITIC.
- SENA (or Cisená).** *Bantu* language widely spoken in Mozambique; the first wordlist was published in W. Bleek's *Languages of Mosambique* (1856).
- SERBIAN.** South *Slavic* language spoken in the Balkans, and now the official language of Serbia; it is first recorded in medieval texts, in which there is no hard distinction between CROATIAN and Serbian. The first free-standing printed dictionary of Serbian was that of T. Abramović (1791); the most ambitious modern dictionary is the *Rečnik srpskohrvatskog književnog i narodnog jezika* (1959–; 19 vols. to date).
- SERI.** Language isolate with a few speakers in northern Mexico; there have been dictionaries with SPANISH in the twentieth and twenty-first centuries (the latter also with ENGLISH).
- Shona.** Group of *Bantu* languages spoken in Zimbabwe, Mozambique, and Botswana, including NDAU, standard SHONA, and ZEZURU.
- SHONA (standard Shona).** *Shona* language spoken by a large population in Zimbabwe and Mozambique, codified in the first half of the twentieth century; the first monolingual dictionary was H. Chimhundú's *Duramazwi ReChiShona* (1996).
- SINHALESE.** New *Indo-Aryan* language spoken in Sri Lanka, with a long tradition of literacy; a bidirectional dictionary with DUTCH was compiled before 1705 by S. Cat, and there are substantial nineteenth- and twentieth-century dictionaries.

Sino-Tibetan. Language family including BODO, DEURI, GARO, KARBI, TANGUT, TIBETAN, TIWA, and the CHINESE language varieties.

SIONA. Language spoken in Colombia and Ecuador; Siona-SPANISH wordlists, one of them dated 1788, are transcribed as part of the *Colección Mutis*.

Siouan. Family of languages spoken in the United States and Canada, including DAKOTA and LAKOTA.

Slavic (Slavonic). Branch of the *Indo-European* language family, divided into East, South, and West Slavic groups: East Slavic includes Belorussian, RUSSIAN, and UKRAINIAN; South Slavic includes BOSNIAN, BULGARIAN, CROATIAN, MACEDONIAN, SERBIAN, and SLOVENE; West Slavic includes CZECH, POLABIAN, POLISH, SLOVAK, and SORBIAN. See also CHURCH SLAVIC and RUTHENIAN.

SLOVAK. West Slavic language spoken in Slovakia, which emerged late as a literary language due to the long-standing dominance of CZECH. The earliest dictionary containing Slovak material is the LATIN-HUNGARIAN-Slovak *Verborum in institutione grammatica contentorum in ungaricum et slavonicum translatio* (1648); the first authoritative Slovak dictionary is that of Št. Peciar (6 vols., 1959–68).

SLOVENE. South Slavic language spoken in Slovenia; the first free-standing printed dictionary was H. Megiser's GERMAN-LATIN-Slovene-ITALIAN *Dictionarium quatuor linguarum* of 1592; a major modern dictionary is the *Slovar slovenskega knjižnega jezika* (5 vols., 1970–91).

SORBIAN. West Slavic language spoken in what is now Germany, a long way to the east of the area where POLABIAN was formerly spoken; divided into Upper Sorbian and Lower Sorbian. The first printed dictionary was by J. Swētlík (1721); more recent are the Upper Sorbian-GERMAN dictionary of F. Jakubaš (1954) and the Lower Sorbian dictionary of A. Muka (1911–28).

SPANISH. *Romance* language, now spoken in Spain, in former Spanish territories (especially in South and central America including Mexico), and in the United States. Records of the written language begin with Aragonese glosses from the tenth or eleventh century; the modern lexicographical tradition begins in the fifteenth century with the work of Elio Antonio de Nebrija, and the *Diccionario de la lengua castellana* of the Real Acadēmia Española (1726–39) began a tradition of academy dictionaries of Castilian, by then the most prestigious variety, followed by an increasingly diverse modern tradition; a number of New World varieties have been the subject of dictionaries. Spanish was very widely used in missionary lexicography with languages of South America, Mesoamerica, parts of North America, and East Asia.

SRANAN TONGO. ENGLISH-lexifier creole spoken in Surinam and by a diasporic population in the Netherlands. The first dictionary is by C. Schumann (1780s); there have recently been printed and online dictionaries with DUTCH.

ŠTOKAVIAN. Variety of CROATIAN.

SUMERIAN. Language isolate formerly spoken in southern Mesopotamia (modern Iraq), attested from c. 3200 BC and replaced in normal use by **AKKADIAN** from c. 2000 BC (it was used for learned purposes and for temple rituals thereafter); the lexicographical record, which begins in the late fourth millennium BC, being among the earliest attestations of the language, is the most ancient in the world.

SWAHILI (Kiswahili). *Bantu* language (with lexical influence from **ARABIC**), widely spoken as a mother tongue and as a second language in Tanzania; also spoken in Kenya, Mozambique, and Somalia, and very widely used as a *lingua franca* in east Africa. There are written texts from the eighteenth century, in Arabic script; the first substantial wordlists are a polyglot of 1850 and a dedicated dictionary completed in 1860 and published in 1882, both by J. Krapf, and both in the Roman alphabet in which the language continues to be written.

SWEDISH. North *Germanic* language spoken in Sweden and by a minority in Finland, with a literary tradition from the thirteenth century onwards. The lexicographical tradition begins in the fifteenth century; a major scholarly dictionary, the *Ordbok över Svenska språget utgiven av Svenska Akademien* or *Svenska Akademiens Ordbok*, has been in progress since 1893.

SYRIAC. Variety of *Aramaic* used in a wide variety of writings from Mesopotamia in the first millennium AD, and as a liturgical language by some Christian communities in the Middle East and southern India; learned dictionaries were made by early modern western Europeans.

TAGALOG. *Malayo-Polynesian* language, more widely spoken in the Philippines than any other. There was a literate tradition before European contact, using a syllabic alphabet ultimately of Indian origin; the first dictionary, by P. de San Buenaventura (1613), is, however, of Christian missionary origin, and in Roman script.

TAJIK. *Iranian* language with official status in Tajikistan, sometimes regarded as a regional variety of **PERSIAN**; there are twentieth-century dictionaries.

TAMIL. *Dravidian* language widely spoken in southern India, Sri Lanka, Singapore, and Malaysia, and in a wider diaspora (notably in South Africa), and written with its own syllabic alphabet. Literacy began in the first millennium BC, and there is a great literary tradition, the language of which may be called Classical Tamil; two other varieties, (modern) Formal Tamil and Spoken Colloquial Tamil, are in a diglossic relationship. The earliest lexicographical work is a chapter in the poetic manual *Tolkāppiyam* (first half of the first millennium AD), and there was a strong monolingual tradition from the eighth or ninth century onwards (including the multi-volume *Tamil Lexicon*, 1924–36), supplemented from the seventeenth century by the work of missionary lexicographers.

TANGUT. *Sino-Tibetan* language formerly spoken in the Xi Xia or Tangut empire; a script was reportedly created for it in 1036 (before then, classical **CHINESE** had probably been the main written language of the Tanguts).

- TARAHUMARA.** *Uto-Aztecan* language spoken in Mexico; an early wordlist was published in 1809.
- TARASCAN (Purépecha).** Tarascan language spoken in Mexico; the first printed dictionary was by M. Gilberti (1559).
- TATAR.** Kipchak *Turkic* language spoken in Tatarstan; the first published dictionary was compiled by S. Khal'fin (1775). The word *Tatar* was formerly used as a generic term for what is called 'Turkic' today.
- TELUGU.** *Dravidian* language spoken in several southern Indian states, with a long literary tradition (the first inscriptions date from the mid first millennium AD); numerous early vocabularies in verse (seventeenth century and later) were modelled on the Sanskrit *Amarakośa*; seven-volume comprehensive monolingual dictionary 1936–72; missionary wordlists are the earliest records of the colloquial language.
- Tepehuan.** Pair of *Uto-Aztecan* languages (Northern Tepehuan and Southern Tepehuan) spoken in northern Mexico; missionary lexicography may have begun before 1616, and a dictionary of Northern Tepehuan was published in 1743.
- THIRRARI.** *Pama-Nyungan* language formerly spoken in South Australia; a wordlist was included in J. Reuther's comparative work of the early twentieth century.
- TIBETAN.** *Sino-Tibetan* language spoken in and beyond the present Tibet Autonomous Region of China. The language was written (in a script of north Indian origin) from the seventh century, and the first SANSKRIT–Tibetan dictionary, the *Mahāvryutpatti* (*Bye brag tu rtogs par byed pa*), was completed in the early ninth century; there were later traditions of lexicography with CHINESE and MONGOLIAN.
- TIWA.** *Sino-Tibetan* language spoken in north-eastern India; there is a dictionary with ASSAMESE and ENGLISH (2004).
- TLINGIT.** *Na-Dene* language spoken in the north-west of North America; linguistically sophisticated dictionaries have been produced since the 1970s.
- TONOCOTÉ.** *Lule-Vilela* language formerly spoken in what is now northern Argentina; the first printed wordlist was by A. Machoni de Cerdeña (1732).
- TUPI (Tupinambá, língua geral, língua brasílica, língua brasileira).** *Tupi-Guarani* language formerly used as a *lingua franca* in what is now Brazil; an anonymous PORTUGUESE–Tupi *Vocabulário na língua brasílica* of 1621–2 was succeeded by other wordlists deriving from it.
- Tupi-Guarani.** Language family of South America including GUARANI and TUPI.
- TURKIC.** Language family including CHAGATAY, the Kipchak languages (CUMAN, TATAR), the Oguz languages (TURKISH, TURKMEN), UYGHUR, and YAKUT.
- TURKISH.** Oguz *Turkic* language spoken in Turkey, Cyprus, and parts of the Balkans, and in diasporic use in Germany and elsewhere in western Europe; the Turkic written tradition goes back to the eleventh century, with a major

lexicographical text, Maḥmūd al-Kāšġarī's *Dīwān Luġāt at-Turk*, among the earliest documents, and the first Anatolian Turkish documents are from the thirteenth century.

TURKMEN. Oguz *Turkic* language spoken in Turkmenistan and neighbouring countries; the printed dictionaries are of the twentieth century.

TUSCAN. Variety of **ITALIAN**, with high literary prestige from the age of Dante, Petrarch, and Boccaccio (late thirteenth to late fourteenth centuries) onwards, and consequent high lexicographical importance.

TZELTAL. *Mayan* language spoken in Mexico; the first dictionary, by D. de Ara, was compiled around 1560 and transmitted in manuscript.

TZOTZIL. *Mayan* language spoken in Mexico; a **SPANISH**–Tzotzil dictionary was compiled in the late sixteenth or early seventeenth century.

TZ'UTUJIL. *Mayan* language spoken in Guatemala, closely related to **KAQCHIKEL** and **K'ICHE'**; there are Tz'utujil words in the Kaqchikel–**SPANISH** wordlist of D. de Vico, made no later than 1555.

UGARITIC. *Semitic* language formerly spoken in what is now Syria, and attested in documents from the fourteenth to the twelfth century BC; occasional **SUMERIAN**–**AKKADIAN** lexicographical texts have an additional translation column in Ugaritic.

UKRAINIAN. East *Slavic* language spoken in Ukraine; **CHURCH SLAVIC** was the earliest written language used in the area (see also **RUTHENIAN**), and Ukrainian asserted its distinctness from **RUSSIAN** only in the late eighteenth and nineteenth centuries. The pioneering Ukrainian–Russian dictionary of B. Hrinčenko (4 vols., 1907–9) was followed by *Slovnýk ukrajins'koji movy* (II vols., 1970–80).

UMBUNDU. *Bantu* language, spoken by the Ovibumdu people of the central plateau of Angola and neighbouring areas; the first dictionary is that of W. Sanders and W. Fay (1885).

UMPITHAMU (Umbindhamu). *Pama-Nyungan* language formerly spoken in Queensland; there is a dictionary app.

UNGARINYIN (Ngarinjin). *Worrorran* language with a few speakers in Western Australia; a dictionary by H. Coate and A. Elkin was published in 1974.

URDU. See **HINDI-URDU**.

Uto-Aztecan. Family of languages spoken in the western United States and Mexico, including **EUDEVE**, **HOPi**, **MAYO**, **MONO**, **O'ODHAM**, **OPATA**, **PIMA BAJO**, **TARAHUMARA**, *Tepehuan*, and **YAQUI**.

UYGHUR. *Turkic* language spoken in the Xinjiang region of China and in Kazakhstan, Kyrgyzstan, and Uzbekistan; literacy (and a script of Aramaic origin, subsequently used for writing **MONGOLIAN** and **MANCHU**, but after the seventeenth century abandoned by the Uyghurs themselves in favour of a version of the Arabic script) developed in the first millennium AD; some of the *Turkic*

forms in Maḥmūd al-Kāšgarī's *Dīwān Luġāt at-Turk* can be identified with Uyghur, and it appears in a number of later dictionaries; for instance, an eighteenth-century dictionary with Manchu headwords has equivalents in CHINESE, Mongolian, TIBETAN, and Uyghur (in Arabic script).

VIETNAMESE (Annamese). Austroasiatic language spoken in Vietnam and by a widely scattered diasporic population, formerly written in modified Chinese characters (Chữ Nôm; classical CHINESE was used for formal writing), used in a native tradition of lexicography from 1641 or earlier; a modified Roman alphabet, Quốc Ngữ, which is now the normal writing system, was perfected by A. de Rhodes, who also compiled the first missionary dictionary (1651); a Chinese–Vietnamese glossary using Chinese characters probably dates from the beginning of the seventeenth century.

Visayan (Bisaya). Group of *Malayo-Polynesian* languages spoken in the Philippines, including HILIGAYNON and forming a continuum with the *Bicol* languages; the first extant wordlist was by A. de Métrida (1637).

WADJUK. Variety of NOONGAR; a wordlist with ITALIAN is in the *Memorie storiche* of R. Salvado (1851).

WALMAJARRI. *Pama-Nyungan* language spoken in Western Australia (see also JUWALINY); there is a dictionary by E. Richards and J. Hudson (1990).

WANGKANGURRU. *Pama-Nyungan* language formerly spoken in the Northern Territory of Australia; comparative wordlists with DIYARI were made around 1900.

WARAO. Language isolate spoken in the Orinoco Delta region of Venezuela; a SPANISH–Warao wordlist of 1789 or earlier is transcribed as part of the *Colección Mutis*.

WARLPURI. *Pama-Nyungan* language spoken in the Northern Territory of Australia and in Western Australia; there is a dictionary by L. Reece (1979).

WELSH. *Celtic* language spoken in Wales; the earliest inscription may be from the eighth century, and there is a continuous literary tradition from the ninth century onwards. Early glosses were followed by monolingual wordlists from the fifteenth century onwards; the first printed dictionary, by W. Salesbury, was of 1547.

WESTERN APACHE. *Athabaskan* language spoken in Arizona; the first wordlist was made by A. Whipple in 1861×1863.

WESTERN DESERT LANGUAGE. *Pama-Nyungan* language spoken in and near Western Australia; there is an unpublished dictionary by E. Ten Raa and S. Woenne (1970–3).

WIK MUNKAN (Wik Mungkan). *Pama-Nyungan* language spoken in Queensland; there is a dictionary by C. Kilham et al. (1986).

WIRADJURI. *Pama-Nyungan* language with a few speakers in New South Wales, currently undergoing some revitalization; there were several missionary wordlists in the nineteenth century, and there is now a dictionary app.

- WOLOF.** *Niger-Congo* language spoken in Senegal, Gambia, and Mauritania; the first dictionary is that of J. Dard (1825).
- WORRORRA.** *Worrorran* language with a few speakers in Western Australia; wordlists were made by J. Love in the twentieth century.
- Worrorran.** Group of languages spoken in northern Western Australia, including GUNIN / KWINI, UNGARINYIN, and WORRORRA.
- WÚ (WÚYŮ 吳語).** Variety of CHINESE spoken in the city of Shànghǎi and surrounding areas; no significant lexicography before the nineteenth century.
- XEBERO (Jebero).** Cahuapanan language spoken in part of Peru; anonymous eighteenth-century wordlists with SPANISH and QUECHUA have been edited as *Eighteenth Century Xebera* (2016) and *Vocabulario de la lengua xebera* (2016).
- XHOSA (or Isixhosa).** *Bantu* language widely spoken in South Africa and Lesotho; the first substantial dictionary was that of W. Davis (1872).
- YAKUT.** *Turkic* language of Siberia, touched on in a glossary in N. Witsen's encyclopedic *Noord en Oost Tartarye* of 1692 and treated in detail in E. Piekarski's dictionary of 1907–30.
- YANDRUWANDHA.** *Pama-Nyungan* language formerly spoken in South Australia and Queensland (see also YAWARRAWARRKA); a wordlist was included in J. Reuther's comparative work of the early twentieth century.
- YAO (also called Chiyao or Ajawa).** *Bantu* language spoken in Malawi, Mozambique, Tanzania, and Zambia; it was documented in J. Krapf's *Vocabulary of Six East-African Languages* (1850).
- YAQUI (Yoeme).** *Uto-Aztecan* language spoken in Arizona and northern Mexico; at least one early wordlist was made by missionaries.
- YASSIC.** *Iranian* language formerly spoken in Hungary; there is a fifteenth-century wordlist.
- YAWARRAWARRKA.** Variety of YANDRUWANDHA; a wordlist was included in J. Reuther's comparative work of the early twentieth century.
- YAWURU.** *Nyulnyulan* language formerly spoken in Western Australia; a wordlist is included in H. Nekes and E. Worms' *Australian Languages* (1953).
- YIDDISH.** West *Germanic* language with borrowings from HEBREW and the *Slavic* languages, spoken in many Jewish communities in central and eastern Europe until the Holocaust, and still used by more scattered Jewish communities in Europe, Israel, and North America. There was a medieval glossarial tradition; the first printed dictionary was Rabbi Anshel's Hebrew–Yiddish *Mirkevet ha-Mishneh* (1534×1535).
- YORTA YORTA.** *Pama-Nyungan* language with a few speakers in New South Wales, currently undergoing some revitalization; there is a dictionary app.
- YORUBA.** *Niger-Congo* language spoken mostly in Nigeria and Benin; the first dictionary is that of S. Crowther (1843).

YUÈ (Yuèyǔ 粵語) (**internationally also known as Cantonese**). Variety of CHINESE spoken in the Chinese province of Guǎngdōng (Canton), with official status in the cities of Hong Kong and Macao, and widely spoken in overseas Chinese communities throughout the world; no major dictionary before the nineteenth century.

YULPARIJA. Variety of MARTU WANGKA; a wordlist was made in the twentieth century by K. McKelson.

YUP'IK. *Eskimo-Aleut* language spoken in Alaska; there are twentieth- and twenty-first-century dictionaries.

YUWAALARAAY. Variety of GAMILARAAY; there is a dictionary app.

ZAPOTEC. Group of closely related *Oto-Manguean* languages spoken in what is now Mexico; the first printed dictionary is that of J. de Córdova (1578).

ZEZURU. *Shona* language spoken in Zimbabwe, documented in A. Hartmann's *English–Mashona Dictionary* (1894), the first free-standing dictionary of any Shona variety.

ZULU (Isizulu). *Bantu* language, widely spoken as a mother tongue and as a second language in South Africa, and also in Lesotho, Malawi, Mozambique, and Swaziland; the first dictionaries were those of J. Perrin (1855).

Appendix 2 *The Lexicographers*

The following is a list of the principal lexicographers named in this volume. It gives the following information, where possible.

- (1) Name. Many of the persons listed have been known by more than one name, or by different forms of a name: the form used in this volume has been given in first place, and other forms have been given very selectively. Alphabetization takes no account of diacritics or digraphs; the prefix *al-* in Arabic names is ignored. Names in small capitals within entries are cross-references.
- (2) Date and place of birth and death; where these are unknown, then some indication of region and period (*from* indicates a significant place of residence when birthplace is not known). The sign \times indicates ranges of possible dates: '1409 \times 1419' means 'a date in the range 1409 to 1419' whereas '1409–19' means 'the eleven-year period beginning 1409 and ending 1419'. Centuries are AD unless otherwise indicated; double dates in the form 1077/1666 are AH/AD. Where possible, an attempt has been made to accompany dates of birth and death with the name of a settlement and then (except in the case of capital or self-governing cities), the country or sovereign territory (or county, in the case of England, or state, in the case of the United States of America) in which the settlement was located at the stated date and, if different and non-obvious, the country in which it is located now.
- (3) Principal occupation other than 'lexicographer', if known.
- (4) Principal contribution to lexicography, very briefly stated.

AASEN, IVAR (Åsen, Norway 1813 – Christiania [now Oslo] 1896). Philologist and dramatist, codifier of Nynorsk; achieved this partly by the compilation of influential Norwegian–Danish dictionaries.

‘ABD-AL-LATĪF (ABDÜLLATIF) IBNĪ MELEK (late 13th cent.). Compiler of an Arabic–Turkish wordlist in verse.

- ‘ABD-AL-RASHĪD (b. Thatta, Sindh [now in Pakistan], late 16th or early 17th cent.; d. 1077/1666). Compiler of a Persian dictionary.
- ABDEL-NOUR, JABBOUR (from Lebanon; 1913–1991). Professor; compiler, with S. IDRIS, of a French–Arabic dictionary.
- ABRAMOVIĆ, TEODOR (from Ruma, Habsburg empire [now in Serbia]; 18th cent.). Proofreader; compiler of the first free-standing printed dictionary of Serbian.
- ABŪ L-BARAKAT IBN KABAR (from Egypt; died 1324). Coptic scholar and encyclopedist; compiler of a classified Coptic–Arabic vocabulary.
- ABŪ JA‘FAR AL-BAYHAQĪ (BŪ JA‘FARAK) (Bayhaq, Khorasan [now in Iran] 470/1077 – Nishapur, Khorasan [now in Iran] 544/1150). Imam and teacher; compiler of an Arabic–Persian dictionary.
- ABŪ MIṢĤAL AL-A‘RĀBĪ (d. 231/845). Compiler of a miscellany of strange Arabic usage (*ġarīb*).
- ABŪ L-ṬAYYIB AL-LUĠAWĪ (d. 351/962). Compiler of an alphabetical wordlist documenting consonantal substitution in Arabic.
- ABŪ ‘UBAYD (Harāt, Abbasid Caliphate [now in Afghanistan] 154/770 – Mecca 224/838). Grammarian, Qur’ānic scholar, and lawyer; compiler of the earliest extant multi-thematic Arabic lexicon, and of a proverb collection.
- ABŪ ZAYD AL-ANṢĀRĪ (d. Baghdad? 214×215/830×831). Grammarian; compiler of several single-theme Arabic wordlists.
- ADDY, SIDNEY OLDALL (Norton, Derbyshire 1848 – London 1933). Solicitor and antiquary; compiler of a dictionary of Sheffield regionalisms.
- ADELUNG, JOHANN CHRISTOPH (Spanketow, Pomerania [now in Germany] 1732 – Dresden, Kingdom of Saxony [now in Germany] 1806). Librarian and philologist; compiler of a major monolingual German dictionary.
- ADĪB NAṬANZĪ (from Naṭanz near Isfahan, Seljuk empire [now in Iran]; d. c. 497/1103). Compiler of an Arabic dictionary.
- AGGAVAṢSA (from Myanmar; 12th cent.). Author of the Pali grammatical work *Saddanīti*, which includes lexical material.
- AITKEN, ADAM JACK (Edinburgh 1921 – Edinburgh 1998). University lecturer; editor, succeeding W. CRAIGIE, of the *Dictionary of the Older Scottish Tongue*.
- AJAYAPĀLA (from Bengal [now in India]; late 11th or early 12th cent.). Compiler of a homonymic Sanskrit dictionary.
- AKTUŇ, HULKĪ (b. Istanbul 1949). Writer and poet; compiler of a Turkish slang dictionary.
- ALCALÁ, PEDRO DE (from Spain; mid 15th–early 16th cent.). Friar; compiler of the Spanish–Arabic *Vocabulista arauigo*, which was both the first printed dictionary of Arabic and the first substantial printed dictionary of two European vernaculars.

- ALCALAY, REUBEN (Jerusalem 1907 – Jerusalem 1976).** Translator and editor; compiler of *The Complete Hebrew–English Dictionary*.
- ALCEDO, ANTONIO DE (Quito, Spanish empire [now in Ecuador], 1735 – Madrid 1812).** Soldier, geographer, and historian; compiler of an important early wordlist of Americanisms in Spanish.
- ALDHelm (d. Douling, Kingdom of Wessex [now in Somerset] 707×710).** Bishop, poet in Latin, and scholar; involved in compilation of a Latin wordlist with Old English interpretamenta.
- ALLSOPP, STANLEY REGINALD RICHARD (Georgetown, British Guiana [now Guyana] 1929 – Barbados 2009).** Professor of English, editor of the *Dictionary of Caribbean English Usage*.
- ALONSO, JUAN (later 16th cent.).** Franciscan missionary in what is now Guatemala; compiler of a Spanish–Kaqchikel wordlist.
- ALVARADO, FRANCISCO DE (from the Viceroyalty of New Spain, Spanish empire [now Mexico]; fl. 1574–1593).** Dominican missionary in what is now Mexico; compiler of the first dictionary of Mixtec.
- ALVES, ALBINO (JOÃO ALBINO ALVES MANSO) (Serapicos, Bragança, Portugal 1908 – Nova Lisboa [now Huambo], Angola 1956).** Spiritan missionary in Angola; compiler of a Kimbundu–Portuguese dictionary.
- AMARASIMHA (India 4th×8th cent.).** Compiler of the first Sanskrit lexicon extant in its entirety, titled *Nāmaṅgānuśāsana*, and popularly known as *Amarakośa*.
- AMENEMOPE (from Egypt; 12th cent. BC).** Scholar and temple scribe, ‘scribe of sacred books in the house of life’; compiler of a monolingual ancient Egyptian ‘onomasticon’ (encyclopedic glossary organized according to thematic classes). He is the earliest lexicographer in the world whose name has come down to us.
- ANGE DE SAINT JOSEPH (JOSEPH DE LA BROSSE) (Toulouse, France 1636 – Perpignan, France 1697).** Carmelite missionary and pharmacologist; compiler of a grammar with an Italian–Latin–French–Persian wordlist.
- ANSHEL (RABBI ANSHEL) (from central Europe; 15th or early 16th cent.).** Compiler of the first printed dictionary of Yiddish.
- APION (b. in El Kharga oasis, Roman empire [now in Egypt]; fl. 1st cent. AD).** Teacher; compiler of an alphabetically arranged Homeric glossary.
- APOLLONIUS SOPHISTA (from the Roman empire, 2nd cent. AD).** Compiler of a monolingual alphabetical lexicon of the Greek of Homer.
- ARA, DOMINGO DE (later 16th cent.).** Dominican missionary in what is now Mexico; compiler of the first wordlist of Tzeltal.
- ARAGONA, ALONSO DE (Naples, Spanish empire [now in Italy] 1585 – Asunción, Spanish empire [now in Paraguay] 1629).** Jesuit missionary in South America; compiler of a large, and apparently lost, bidirectional dictionary of Spanish and Guarani.

- ARENAS, PEDRO DE** (**Viceroyalty of New Spain, Spanish empire [now Mexico], early 17th cent.**). Merchant; compiler of a small dictionary of Nahuatl and Spanish.
- ARGENTI, FILIPPO** (**from Florence, Tuscany [now in Italy]; 16th cent.**). Diplomat in Constantinople; compiler of two Italian–Turkish wordlists.
- ARISTOPHANES OF BYZANTIUM** (**b. 257 BC; fl. Alexandria, Egypt; d. 180 BC?**). Librarian and literary scholar; the founding father of ancient lexicography.
- ARROYO DE LA CUESTA, FELIPE** (**Cubo de Bureba, Spain 1780 – Mission Santa Inés, Mexico [now in Solvang, California] 1840**). Franciscan missionary in California; compiler of a dictionary of Mutsun.
- ARTS, TRESSY** (**b. Cuijk, Netherlands 1974**). Editor of the Arabic–English and English–Arabic *Oxford Arabic Dictionary*.
- ĀRZŪ, SIRĀJ AL-DĪN ‘ALĪ KHĀN** (**Akbarābād [now Agra] or Gwalior, Mughal empire [now in India] 1687 – Lucknow [now in India] 1756**). Compiler of a Persian dictionary, and of a lexicon of vernacular words that had entered the Persianate cultural sphere.
- ASADĪ TŪSĪ** (**from Tus in eastern Persia; c. 1000 – c. 1073**). Poet and copyist; compiler of the earliest extant Persian wordlist.
- AL-‘ASKARĪ** (**d. after 395/1005**). Compiler of a wordlist distinguishing supposedly synonymous Arabic words.
- AL-AṢMA‘Ī** (**d. 213×216/828×831**). Philologist; compiler of numerous single-theme Arabic wordlists.
- ASSUMPÇAM (ASSUMPÇAO), MANOEL DA** (**from Évora, Portugal; first half of the 18th cent.**). Augustinian missionary in India; compiler of a Bengali grammar and bidirectional wordlist of Bengali and Portuguese.
- ATHANASIUS OF QUS** (**from Egypt; fl. 2nd half of the 14th cent.**). Bishop, grammarian, and interpreter; compiler of Coptic–Arabic glossaries and author of a Coptic grammar in Arabic.
- AUBERY, JOSEPH** (**Gisors, France 1673 – Saint-François-de-Sales mission, New France [now Saint-François-du-Lac, Canada] 1756**). Jesuit missionary in what is now Canada; compiler of an Abenaki dictionary.
- AULETE, FRANCISCO JÚLIO DE CALDAS** (**Lisbon 1823 – Lisbon 1878**). Teacher and politician; compiler of a Portuguese dictionary.
- AVIS, WALTER SPENCER** (**Toronto, Canada 1919 – Kingston, Canada 1979**). Professor of English; editor of the *Dictionary of Canadianisms on Historical Principles*.
- AVNEYON (né BERGSTEIN), EITAN** (**b. Siedlce, Poland 1930**). Compiler of a Hebrew dictionary.
- AZAR, MOSHE** (**Brussels 1934 – Haifa, Israel 2012**). Professor of the Hebrew language; editor of the Hebrew dictionary *Milon Even-Shoshan* (‘Even-Shoshan dictionary’), on the basis of the work of A. EVEN-SHOSHAN.

- AL-AZHARĪ (Harāt, Abbasid Caliphate [now in Afghanistan] 282/895 – Harāt 370/980).** Compiler of a phonetically arranged Arabic lexicon in fifteen volumes.
- BAALBAKI, MOUNIR (1918 – Beirut 1999).** Translator and publisher; compiler of Arabic–English and English–Arabic dictionaries.
- BAALBAKI, RAMZI MOUNIR (b. Beirut 1951).** Professor of Arabic (and contributor to this volume); compiler of Arabic dictionaries in the tradition founded by his father M. BAALBAKI.
- BAALBAKI, RUHI (b. Beirut 1947).** Compiler of Arabic dictionaries in the tradition founded by his father M. BAALBAKI.
- BADAWI, EL-SAID (El-Nakhas, Egypt 1929 – Cairo 2014).** Professor of Arabic; compiler, with M. HINDS, of a dictionary of Egyptian Arabic.
- BADR-AL-DĪN IBRĀHĪM (from Persianate India; late 14th or early 15th cent.).** Compiler of a Persian dictionary.
- BAHĀR, ṬEK CHAND (from Delhi, Mughal empire [now in India]; 18th cent.).** Compiler of a Persian dictionary.
- BAILEY, NATHAN (baptized London 1691; d. London 1742).** Teacher; compiler of an important predecessor of S. JOHNSON'S *Dictionary*.
- BALBI, GIOVANNI (b. in Genoa [now in Italy]; d. Genoa 1298?).** Dominican friar and theologian; compiler of the monolingual alphabetically ordered Latin *Catholicon*.
- BALDINGER, KURT (Binningen, Switzerland 1919 – Heidelberg, Germany 2007).** Philologist; founding editor of the *Dictionnaire étymologique de l'ancien français*, and compiler of a dictionary of Old Gascon.
- BALDWIN, CALEB COOK (Bloomfield, New Jersey 1820 – East Orange, New Jersey 1911).** Presbyterian missionary in China; compiler, with R. MACLAY, of *An Alphabetic Dictionary of the Chinese Language in the Foochow Dialect*.
- AL-BANDANĪĠĪ (d. 284/897).** Compiler of the earliest extant Arabic lexicon to follow the rhyme system of arrangement.
- BANKS, JOSEPH (London 1743 – Spring Grove, Middlesex 1820).** Naturalist; he and J. COOK compiled wordlists in Guugu Yimidhirr in 1770.
- BARAGA, FREDERIC (Knežja Vas, Habsburg empire [now in Slovenia] 1797 – Marquette, Michigan 1868).** Missionary (Leopoldine Society) in North America and titular bishop of Amyzonias; compiler of an influential dictionary of Ojibwa.
- BARBIER DE MEYNARD, CHARLES ADRIEN CASIMIR (Marseille, France 1826 – Paris 1908).** Historian and Orientalist; compiler of a Turkish–French dictionary.
- BARDHI, FRANG (FRANCISCUS BLANCUS) (Kallmeti, Ottoman empire [now in Albania] 1606 – Rome 1643).** Bishop of Sapa and Sarda; compiler of the first dictionary of Albanian.

- BARLEMENT, NOEL VAN (DE BERLAIMONT)** (perhaps from Velaines, Duchy of Burgundy [now in Belgium]; d. Antwerp, Habsburg Netherlands [now in Belgium] 1531). Schoolmaster; compiler of a popular polyglot dictionary.
- BARTLETT, JOHN RUSSELL** (Providence, Rhode Island 1805 – Providence 1886). Politician and historian; compiler of *Dictionary of Americanisms*.
- BARUA, JADURAM DEKA** (from Assam; 1801–1869). Writer and reformer; his system of orthography was used in the Assamese dictionary of M. BRONSON.
- BARUWĀ, HEMCHANDRA** (from Assam; 1835–1896). Scholar and civil servant; compiler of a major dictionary of Assamese.
- BARZIZZA, GASPARINO** (Bergamo? c. 1360 – Milan 1431). Teacher and humanist; compiler of a Latin–Bergamask wordlist.
- BASALENQUE, DIEGO** (Salamanca, Spain 1577 – Charo, Spanish empire [now in Mexico] 1651). Augustinian missionary in what is now Mexico; compiler of a grammar and dictionary of Matlatzinca.
- BASILIO, TOMÁS** (Palermo, Spanish empire [now in Italy] 1580 – Sinaloa, Spanish empire [now in Mexico] 1654). Jesuit missionary in what is now Mexico; a grammar and vocabulary of Yaqui have been attributed to him and to J. DE VELASCO.
- AL-BATALYAWŚĪ** (Badajoz, al-Andalus [now in Spain] 444/1052 – Valencia, al-Andalus [now in Spain] 521/1127). Grammarian and philosopher; compiler of a collection of groups of three words whose meanings vary according to their vowels.
- BATES, DAISY MAY** (Tipperary, Ireland 1859 – Prospect [now part of Adelaide], Australia 1951). Welfare worker among Australian Aborigines and anthropologist; she made an important early collection of wordlists of languages of Western Australia.
- BATHE, WILLIAM** (Drumcondra, Ireland 1564 – Madrid? 1614). Jesuit; compiler of the Latin–Spanish *Janua linguarum*, a forerunner to the work of COMENIUS.
- BATTAGLIA, SALVATORE** (Catania, Italy 1904 – Naples, Italy 1971). Literary scholar; founding editor of the *Grande dizionario della lingua italiana*.
- BEDERSI, ABRAHAM BEN ISAAC** (from Perpignan, Kingdom of Aragon [now in France]; c. 1230–c. 1300). Poet; compiler of a Hebrew synonym dictionary.
- BÉLISLE, LOUIS-ALEXANDRE** (St-Éloi de Témiscouata, Canada 1902 – Quebec City, Canada 1985). Publisher and translator; compiler of the normative *Dictionnaire général de la langue française au Canada*.
- BELLINI, BERNARDO** (Griante, Austrian empire [now in Italy] 1792 – Turin, Italy 1876). Poet and printer; compiler, with N. TOMMASEO, of the seven-volume *Dizionario della lingua italiana*.
- BELOSTENEC, IVAN** (Varaždin, Kingdom of Hungary [now in Croatia] c. 1594 – Lepoglava, Kingdom of Hungary [now in Croatia] 1674). Paulist monk and

- prior, homilist, and poet; compiler of a dictionary containing Kajkavian, Štokavian, and Čakavian material.
- BENEŠOVSKÝ, MATOUŠ (PHILONOMUS) (from Benešov, Bohemia [now in the Czech Republic]; c. 1550–1595?).** Humanist; author of a Czech grammar and the first monolingual Czech dictionary.
- BENTLEY, WILLIAM HOLMAN (Sudbury, Suffolk 1855 – Bristol, England 1905).** Baptist missionary; compiler, with the assistance of NLEMVO, of a dictionary of Kikongo.
- BEN-YEHUDA (né PERLMAN), ELIEZER (Lužki [now in Belarus] 1858 – Jerusalem 1922).** Hebrew language revivalist; compiler of *Milon halashon haivrit hayeshana vehakhadasha* ('A complete dictionary of Ancient and Modern Hebrew').
- BERGAÑO, DIEGO DE (1690–1747).** Augustinian missionary in the Philippines; compiler of the first dictionary of Pampango.
- BERNOLÁK, ANTON (Slanica, Kingdom of Hungary [now an inundated village in Slovakia] 1762 – Nové Zámky, Kingdom of Hungary [now in Slovakia] 1813).** Catholic priest, who attempted to codify the Slovak language; author of grammatical treatises and a Slovak/Czech–Latin–German–Hungarian dictionary.
- BERTONIO, LUDOVICO (Rocca Contrada, Papal States [now Arcevia, Italy] 1552 – Lima 1625).** Jesuit missionary in South America; compiler of a dictionary of Aymara.
- BERYNDÁ, PAMVA (Čajkovyči [now in Ukraine] 1555×1560 – Kiev 1632).** Poet and translator; compiler of a Church Slavic–Ruthenian dictionary.
- BESCHI, CONSTANZO GIUSEPPE (VĪRAMĀMUṆĪVAR) (Castiglione delle Stiviere [now in Italy] 1680 – Kerala [now in India] 1747).** Jesuit missionary in India, poet in Tamil, and grammarian; compiler of a monolingual Tamil dictionary (and of other dictionaries).
- BESKROVNYJ, VASILIJ MATVEEVIČ (from Russia; 1908–1978).** Indologist; compiler of major dictionaries of Urdu and Hindi with Russian.
- BIANCHI, THOMAS XAVIER (Paris 1783 – Paris 1864).** Orientalist and secretary-interpreter to the king of France; coauthor with J. KIEFFER of a Turkish–French dictionary.
- BIBBESWORTH, WALTER DE (from Essex; fl. 1219–1270).** Landowner and Anglo-Norman poet; his *Tretiz* (1240×1250?) is a rhyming guide to learning French.
- BILODID, IVAN KOSTJANTYNOVYČ (IVAN KONSTANTINOVYČ BELODED) (Uspenska, Russian empire [now in Ukraine] 1906 – Kiev 1981).** Linguist; editor of the Ukrainian Academy dictionary *Slovyk ukrains'koji movy*.
- BIRABAN (JOHN MCGILL, WE-POHNG) (from New South Wales; fl. 1819–1842).** Awabakal leader; he taught L. THRELKELD about the Awabakal language.

- BISCHOF, JOSEPH (Dülken, German empire 1878 – George, South Africa 1958).** Pallottine missionary in Australia; compiler of an English–N̄yulnyul wordlist on the basis of earlier French–N̄yulnyul work by Trappist monks.
- BLEEK, WILHELM HEINRICH IMMANUEL (Berlin 1827 – Mowbray, Cape Colony [now in South Africa] 1875).** Comparative linguist; compiler, on the basis of manuscript wordlists by W. PETERS, of *The Languages of Mosambique*, which includes the first wordlists of Ronga and Shona.
- BLEMMYDES, NIKEPHOROS (Constantinople? [now Istanbul, Turkey] 1197 – Ephesus, Byzantine empire [now in Turkey] c. 1269).** Monk and teacher; possibly the compiler, or one of the compilers, of the thirteenth-century Greek lexicon formerly attributed to Joannes Zonaras.
- BLOCH, OSCAR (Le Thillot, France 1877 – Paris 1937).** Philologist; co-editor with W. VON WARTBURG of a French etymological dictionary.
- BLUTEAU, RAFAEL (London 1638 – Lisbon 1734).** Theatine monk; compiler of a ten-volume dictionary of Portuguese.
- BÖHTLINGK, OTTO VON (St Petersburg, Russia 1815 – Leipzig, Germany 1904).** Philologist; compiler, with R. VON ROTH, of a seven-volume *Sanskrit-Wörterbuch*.
- BONELLI, LUIGI (Brescia, Italy 1865 – Naples, Italy 1947).** Orientalist; compiler of a Turkish–Italian dictionary.
- BONET, JEAN PIERRE JOSEPH (Bages, France 1844 – Paris 1907).** Interpreter and translator; compiler of a Vietnamese–French dictionary.
- BOUCHER, JONATHAN (Blencogo, Cumberland 1738 – Epsom, Surrey 1804).** Clergyman in North America and England; compiler of *A Glossary of Archaic and Provincial Words*.
- BOURKE, JOHN GREGORY (Philadelphia 1843 – Philadelphia 1896).** Soldier and anthropologist; compiler of a Western Apache wordlist.
- BOWREY, THOMAS (d. in or near Wapping, Middlesex [now in London] 1703).** Merchant; compiler of the first dictionary of English and Malay.
- BRADLEY, HENRY (Manchester, Lancashire 1845 – Oxford, Oxfordshire 1923).** Philologist; one of the four editors (with J. MURRAY, W. CRAIGIE, and C. ONIONS) of the first edition of the *Oxford English Dictionary*.
- BRENDER, FRANZ (Basel, Switzerland 1894 – Kaunas, Lithuania 1938).** Linguist; co-editor of a major Lithuanian dictionary started by M. NIEDERMANN and A. SENN; see also A. SALYS.
- BRETON, RAYMOND (Beaune, France 1609 – Caen, France 1679).** Dominican missionary in the Caribbean; compiler of a dictionary of an Arawakan language variety.
- BRIGHT, WILLIAM OLIVER (Oxnard, California 1828 – Louisville, Colorado 2006).** Professor of linguistics; editor of *Native American Placenames of the United States*.

- BRITO, GULIELMUS (GUILLAUME LE BRETON)** (from Brittany? and France; fl. 1249). Franciscan friar and theologian; compiler of a monolingual Latin biblical dictionary.
- BRODOWSKI, JACOB (JOKŪBAS BRODOVSKIS)** (Goldap, Polish-Lithuanian Commonwealth [now in Poland] c. 1692 – Trempen, Kingdom of Prussia [now Novostroevo, Kaliningrad Oblast, Russia] 1744). Precentor; compiler of a wide-ranging unpublished bidirectional dictionary of Lithuanian and German.
- BROLLO, BASILIO** (Gemona del Friuli, Republic of Venice [now in Italy] 1648 – Xīān, China 1704). Franciscan missionary in China; compiler of an unpublished 'Dictionarium Sino-Latinum', which was adapted and published by C.-L.-J. DE GUIGNES.
- BRONSON, MILES** (Norway, New York 1812 – Eaton Rapids, Michigan 1883). Baptist missionary in Assam and compiler of dictionaries of Assamese, with the assistance of JADURAM DEKA BARUA.
- BROWN, CHARLES PHILIP** (Calcutta, Bengal Presidency [now Kolkata, India] 1798 – London 1884). Colonial administrator and Indologist; compiler of Telugu dictionaries.
- BRUYAS, JACQUES** (Lyon, France 1635 – Caughnawaga mission, New France [now Kahnawake Mohawk Territory, Canada], 1712). Jesuit missionary in what is now Canada; compiler of a dictionary of Mohawk.
- BRYANT, ALFRED THOMAS (FATHER DAVID)** (London 1865 – Cambridge, Cambridgeshire 1953). Catholic priest and ethnologist; compiler of a Zulu dictionary.
- BUARQUE DE HOLANDA FERREIRA, AURÉLIO** (Passo de Camaragibe, Brazil 1910 – Rio de Janeiro 1989). Philologist, translator, and writer; compiler of a major dictionary of Portuguese.
- BUDÉ, GUILLAUME** (Paris 1467 – Paris 1540). Humanist; author of *Commentarii linguae graecae*, an important source for subsequent Greek-Latin dictionaries.
- BUECHEL, EUGENE** (Schleid, German empire 1874 – O'Neill, Nebraska 1954). Jesuit missionary in the United States; compiler of a Lakota dictionary.
- BŪGA, KAZIMIERAS (Pažiegė, Russian empire [now in Lithuania] 1879 – Königsberg, Germany [now Kaliningrad, Russia] 1924)**. Linguist of great importance for the study of his native tongue; founding editor of a major monolingual Lithuanian dictionary.
- BUGARÍN, JOSÉ** (Spain 1606 – Manila 1676). Dominican missionary in the Philippines; compiler of a wordlist of Ibanag.
- BURCHFIELD, ROBERT WILLIAM** (Wanganui, New Zealand 1923 – Abingdon, Oxfordshire 2004). Editor of the four-volume *OED Supplement*.

- BURNELL, ARTHUR COKE (St Briavels, Gloucestershire 1840 – West Stratton, Hampshire 1882).** Sanskrit scholar; compiler (with H. YULE) of *Hobson-Jobson*, a dictionary of English as used in India.
- BUSBECQ, OGIER GHISELIN DE (Comines, Habsburg Netherlands [now on the French–Belgian border] 1522 – Saint-Germain, near Rouen, France 1592).** Diplomat and botanist; made the first and only wordlist of Crimean Gothic.
- AL-BUSTĀNĪ, BUTRUS (Dibbiye, Ottoman empire [now in Lebanon] 1819 – Beirut 1883).** Encyclopedist and public intellectual; compiler of a monolingual Arabic dictionary.
- BUXTORF, JOHANNES, THE ELDER (Kamen, Duchy of Cleves [now in Germany] 1564 – Basel, Swiss Confederacy 1629).** Hebraist, father of J. BUXTORF THE YOUNGER; compiler of dictionaries of Biblical Hebrew and Aramaic.
- BUXTORF, JOHANNES, THE YOUNGER (Basel, Swiss Confederacy 1599 – Basel 1664).** Hebraist, son of J. BUXTORF THE ELDER; prepared his father's *Lexicon chaldaicum, talmudicum, et rabbinicum* for publication.
- BYDGOSZCZY, BARTŁOMIEJ Z (Bydgoszcz, Poland c. 1480 – Poznań, Poland 1548).** Bernardine friar; compiler of two Polish wordlists based on Latin sources.
- BYINGTON, CYRUS (Stockbridge, Massachusetts 1793 – Belpre, Ohio 1868).** Missionary (American Board of Commissioners for Foreign Missions) in the United States; compiler of an influential Choctaw dictionary.
- CÀI BIÀN 蔡卞 (from Xianyou, China; 1048–1117).** Son-in-law of WÁNG ĀNSHÍ and official; compiler of a dictionary of terms of botanical and zoological interest in the classical literary text *Shījīng*.
- CALEPINO, AMBROGIO (province of Bergamo [now in Italy] c. 1435 – Bergamo? 1509×1510).** Augustinian friar and humanist; his *Dictionarium* was the basis for a tradition of Latin and polyglot dictionaries extending to the eighteenth century.
- CALLIMACHUS OF CYRENE (fl. Alexandria, Egypt, 3rd cent. BC).** Poet and scholar; compiler of a lost wordlist dealing with regional differentiation within Greek.
- CAMPBELL, WILLIAM (Glasgow, Scotland 1841 – Bournemouth, Sussex 1921).** Missionary to Taiwan and historian; compiler of *Dictionary of the Amoy Vernacular*.
- CANDREA, IOAN AUREL (Bucharest 1872 – Paris 1950).** Philologist and folklorist; compiler with O. DENSUIANU of a Romanian etymological dictionary.
- CANDY, GEORGE (from England; b. 1804; fl. 1854).** Soldier and missionary; compiler, with his twin brother T. CANDY and with J. MOLESWORTH, of a dictionary of Marathi.
- CANDY, THOMAS (England 1804 – Mahabaleshwar, India 1877).** Soldier, educator, and translator; compiler, with his twin brother G. CANDY and with J. MOLESWORTH, of dictionaries of Marathi.

- CANICATTI, BERNARDO MARIA DA (BERNARDUS MARIA A CANECATTIM; family name CASSARO) (Canicatti, Kingdom of Naples [now in Italy] 1746×1751 – Lisbon 1834).** Capuchin missionary; compiler of a dictionary of Kimbundu and a wordlist of a mixture of Kikongo and Kimbundu.
- CANTEMIR, DIMITRIE (Silișteni, Ottoman empire [now in Romania] 1673 – Dmitrovka [now Dmitrovsk], Russia 1723).** Voivode of Moldavia and polymath; compiler of a glossary of the words of foreign origin which he used in his Romanian-language novel *Istoria ieroglifica*.
- CAPELL, ARTHUR (Newtown, Australia 1902 – Gordon, Australia 1986).** Linguist and anthropologist; he collected wordlists in many languages of Arnhem Land, the Kimberley, and other regions of Australia.
- CARDOSO, JERÓNIMO (Lamego, Portugal c. 1508 – Lisbon? c. 1569).** Teacher; compiler of important dictionaries of Latin and Portuguese, from one of which was derived the wordlist of the Portuguese–Chinese dictionary of M. RICCI and M. RUGGIERI.
- CAREY, WILLIAM (Paulerspury, Northamptonshire 1761 – Serampore, Danish empire [now in India] 1834).** Baptist missionary in India, Bible translator, and botanist; collaborator with V. SARMĀ on a Marathi dictionary, and compiler of grammars and dictionaries of other Indian languages.
- CARRADORI, ARCANGELO (Pistoia, Tuscany [now in Italy] late 16th cent.? – Pistoia 1652).** Franciscan missionary in Egypt; compiler of Italian–Turkish and Italian–Nubian dictionaries, which remained in manuscript until the twentieth century.
- CASSIDY, FREDERIC GOMES (Kingston, Jamaica 1907 – Madison, Wisconsin 2000).** Professor of English; chief editor (succeeded by J. HALL) of the *Dictionary of American Regional English* and (with R. LE PAGE) the *Dictionary of Jamaican English*.
- CASTELL, EDMUND (baptized East Hatley, Cambridgeshire 1606; d. Higham Gobion, Bedfordshire 1686).** Orientalist; compiler of *Lexicon Heptaglotton* (1669).
- CASTRO, ANDRÉS DE (from Burgos, Spain; went to Mexico 1542; d. Toluca, Spanish empire [now in Mexico] 1577.** Franciscan missionary in what is now Mexico; the first wordlist of Matlatzinca was written by him in the margins of a copy of the Nahuatl dictionary of A. DE MOLINA.
- CAT (KAT, CATH), SIMON (from Zaandam, Netherlands; d. Ceylon [now Sri Lanka] 1704).** Preacher and Bible translator; compiler of a bidirectional dictionary of Dutch and Sinhalese.
- CAWDREY, ROBERT (b. 1537×1538; d. in or after 1604).** Minister of religion; his *Table Alphabeticall* is the first free-standing non-specialized monolingual English dictionary.

- CHANGUION, ANTOINE NICHOLAS ERNEST** (**The Hague, Netherlands 1803 – Lörrach, German empire 1881**). Educator; compiler of the first Afrikaans wordlist.
- CHAO, EDUARDO** (**Ribadavia, Spain 1822 – Madrid 1887**). Scientist, historian, and politician; compiler of an encyclopedic dictionary of Spanish.
- CHAO, YUEN REN** (**ZHÀO YUÁN RÈN 趙元任**) (**Tiānjīn, China 1892 – Cambridge, Massachusetts 1982**). Linguist, co-editor of the monolingual Chinese dictionary *Guóyǔ cídiǎn* ('Dictionary of the national language').
- CHÉN CHÚN 陳淳** (**from Longxi, China; 1159–1253**). Compiler of the lexicological work *Běixī zìyì*, which offers in-depth philosophical definitions of twenty-six key terms in Neo-Confucian metaphysics.
- CHÉN Dì 陳第** (**from Lianjiang, China; 1541–1617**). Compiler of a dictionary of ancient Chinese character pronunciations.
- CHÉN PÉNGNIÁN 陳彭年** (**from Nancheng, China; 961–1017**). Court scholar; chief editor of the *Guǎngyùn*, a recension of the rhyme dictionary *Qièyùn* (compiled by Lǚ FAYÁN), and of *Dàguǎng yìhuì Yùpiān*, a recension of the *Yùpiān* (compiled by Gǔ YEWÁNG).
- CHÉN SHIYUÁN 陳士元** (**from Yingcheng, China; 1516–1597**). Scholar and official; compiler of a dictionary of vernacular character forms in ancient texts, an encyclopedia of vernacular usages in classical works and more recent texts, and a dictionary of foreign-language terms in various historical works.
- CHIRINO, PEDRO** (**Osuna, Spain 1557 – Manila 1635**). Jesuit missionary in the Philippines; compiler of the first extant missionary dictionary of Hokkien from Manila.
- CHOEROBOSCUS, GEORGE** (**from Constantinople [now Istanbul, Turkey]; second half of the 8th cent.?**). Grammarian; compiler of *Epimerismi in Psalmos*, on the vocabulary of the Greek translation of the Psalms, and perhaps of a similar work on the Iliad.
- CHOLINUS, PETRUS** (**PETER KOLIN**) (**Zug, Swiss Confederacy c. 1508 – Zürich, Swiss Confederacy 1542**). Biblical scholar; compiler, with J. FRISIUS, of a Latin–German dictionary.
- CHOS SKYONG BZANG PO** (**from Tibet; 1441–1528**). Translator of Sanskrit lexicographical works into Tibetan.
- CHOUKA, YAACOV** (**b. Cairo 1936**). Scholar in the digital humanities; compiler of a Hebrew dictionary.
- CIGALE, MATEJ** (**Lome, Austrian empire [now in Slovenia] 1819 – Vienna 1889**). Jurist and editor of a newspaper; editor of a German–Slovenian dictionary drawing on the materials of V. VODNIK.
- CIHAC, ALEXANDRU** (**Iași, Ottoman empire [now in Romania] 1825 – Mainz, German empire 1887**). Philologist; compiler of the first etymological dictionary of Romanian.

- CIORANESCU, ALEJANDRO (ALEXANDRU CIORĂNESCU) (Moroeni, Romania 1911 – Santa Cruz de Tenerife, Spain 1999).** Historian, bibliographer, writer, and linguist; compiler of a Romanian etymological dictionary.
- CIUDAD REAL, ANTONIO DE (Ciudad Real, Spain 1551 – Mérida, Spanish empire [now in Mexico] 1617).** Franciscan missionary in what is now Mexico; compiler of a large Maya–Spanish dictionary.
- CLARETUS DE SOLENCIA (BARTOLOMĚJ Z CHLUMCE) (Chlumec nad Cidlinou, Bohemia [now in the Czech Republic] c. 1320 – Prague? 1370×1379).** Scholar and poet; author of three Latin–Old Czech lexicographical works in verse.
- CLODIUS, JOHANN CHRISTIAN (Großenhain, Saxony [now in Germany] 1676 – Leipzig, Saxony [now in Germany] 1745).** Professor of Arabic; compiled a Latin–Turkish–German dictionary.
- COATE, HOWARD H. J. (England 1910 – Australia 2002).** Itinerant missionary in Australia; compiler of wordlists in several Australian languages, including a dictionary of Ungarinyin with A. ELKIN.
- COLEBROOKE, HENRY THOMAS (London 1765 – London 1837).** East India Company administrator, Sanskritist, and polymath: he translated the *Amarakośa* of AMARASIṂHA into English.
- COLENZO, JOHN WILLIAM (St Austell, Cornwall 1814 – Bishopstowe, Colony of Natal [now in South Africa] 1883).** Anglican bishop of Natal and controversial theologian; compiler of an important Zulu dictionary.
- COLLADO, DIEGO (DIDACUS COLLADUS) (Miajadas, Spain c. 1587 – at sea between the northern Philippines and Manila, 1641?).** Dominican missionary to the Philippines and Japan; compiled a Latin–Spanish–Japanese dictionary.
- COMENIUS, JOHANNES AMOS (JAN AMOS KOMENSKÝ) (Moravia, Austrian empire [now Slovakia] 1592 – Amsterdam 1670).** Theologian and writer on education; compiler of an illustrated pedagogical Latin dictionary.
- CONSTANTIN, ROBERT (Caen, France c. 1530 – Montauban, France 1605).** Physician and humanist; editor of a Greek–Latin dictionary.
- COOK, JAMES (Marton in Cleveland, Yorkshire 1728 – Kealakekua Bay, Hawaii 1779).** Explorer; he and J. BANKS compiled wordlists in Guugu Yimidhirr in 1770.
- COOPER, THOMAS (Oxford, Oxfordshire c. 1517 – Winchester, Hampshire 1594).** Teacher and bishop; compiler of a Latin–English dictionary.
- CÓRDOVA, JUAN DE (d. Antequera, Spain 1595).** Dominican missionary in what is now Mexico; compiler of the first printed dictionary of Zapotec.
- CORMAC MAC CUILENNÁIN (d. Belach Mugna, Leinster [now in Ireland] 908).** King of Munster and bishop; traditionally identified as compiler of the monolingual Irish Cormac’s Glossary.

- CORNEILLE, THOMAS (Rouen, France 1625 – Les Andelys, France 1709).** Dramatist and academician; compiler of the *Dictionnaire des arts et des sciences* issued in 1694 to accompany the *Dictionnaire de l'Académie française*.
- COROMINES I VIGNEAUX, JOAN (JUAN COROMINAS) (Barcelona, Spain 1905 – Pineda de Mar, Spain 1997).** Philologist; compiler of major etymological dictionaries of Spanish and Catalan.
- CORRIENTE CÓRDOBA, FEDERICO (b. Granada, Spain 1940).** Professor of Arabic and Islamic studies; compiler of Arabic–Spanish and Spanish–Arabic dictionaries.
- CORTÉS Y ZEDEÑO, JERÓNIMO TOMÁS DE AQUINO (b. Tlajomulco de Zúñiga, Spanish empire [now in Mexico] 1724; d. late 18th cent.?).** Priest; compiler of a grammar and vocabulary of Nahuatl in the variety of Jalisco.
- COTGRAVE, RANDLE (probably from Cheshire; fl. 1587–1612).** Compiler of a major French–English dictionary.
- COTTON, JOSIAH (Plymouth, Massachusetts 1679 – Plymouth 1756).** Magistrate and lay missionary; compiler of a posthumously published wordlist of Massachusett or Natick.
- COURTOIS, VICTOR-JOSEPH (Livron, Drôme, France 1846 – São José do Mongue, Inhambane, Mozambique 1894).** Jesuit missionary in Mozambique; compiler of the first free-standing dictionaries of Nyungwe.
- COVARRUBIAS HOROZCO, SEBASTIÁN DE (Toledo, Spain 1539 – Cuenca, Spain 1613).** Compiler of a major Spanish dictionary.
- COWAN, J MILTON [sic: J was a single-letter first name, not an abbreviation] (Salt Lake City, Utah 1907 – Northampton, Massachusetts 1993).** Professor of modern languages; translator of the English–Arabic version of the German–Arabic dictionary of H. WEHR.
- CRAIGIE, WILLIAM ALEXANDER (Dundee, Scotland 1867 – Watlington, Oxfordshire 1957).** Philologist; one of the four editors (with J. MURRAY, H. BRADLEY, and C. ONIONS) of the first edition of the *Oxford English Dictionary*; founding editor (succeeded by A. AITKEN) of the *Dictionary of the Scottish Tongue*; and editor (with J. HULBERT, and with M. MATHEWS and A. READ as assistants) of the *Dictionary of American English*.
- CRASTONI (CRASTONE), GIOVANNI (Castel San Giovanni, near Piacenza, Duchy of Milan [now in Italy] 1410×1420 – Milan? after 1497).** Compiler of the first printed Greek–Latin dictionary.
- CROOKE, WILLIAM (Cork, Ireland 1848 – Cheltenham, Gloucestershire 1923).** Orientalist and folklorist; prepared the second edition of A. BURNELL and H. YULE's *Hobson-Jobson*.
- CROWTHER, SAMUEL AJAYI (Osogun, Oyo empire [now in Nigeria] c. 1807 – Lagos [now in Nigeria] 1891).** Bishop of western Africa and translator; compiler

of the first Yoruba dictionary and, with assistance from J. SCHÖN, the first substantial dictionary of Igbo.

CÙA, PAULUS (HUÏNH/HUỖNH TỈNH CÙA) (Bà Rịa, Vietnam 1834 – Saigon [now Ho Chi Minh City], Vietnam 1908). Confucian scholar and translator; compiler of a pioneering monolingual Vietnamese dictionary.

CUERVO, RUFINO JOSÉ (Bogotá 1844 – Paris 1911). Linguist; author of a study of Spanish as spoken in Colombia, and compiler of an unfinished Spanish dictionary.

CŪLA MOGGALLĀNA (from Sri Lanka; late 13th cent.). Compiler of a Pali lexicon modelled on the *Amarakośa* of AMARASĪMHA.

CUOQ, JEAN-ANDRÉ (Le Puy, France 1821 – Lac-des-Deux-Montagnes, Canada 1898). Sulpician missionary in Canada; compiler of dictionaries of Algonquin and Mohawk.

CURR, EDWARD MICKLETHWAITE (Hobart Town, Van Diemen's Land [now in Australia] 1820 – St Kilda, Victoria 1889). Squatter and author; he gathered and published some 300 wordlists of Australian languages.

CYRIL OF ALEXANDRIA (Mahalla, Eastern Roman empire [now in Egypt] 378 – Alexandria, Eastern Roman empire [now in Egypt] 444). Patriarch of Alexandria and theologian; the Greek *Lexicon of Cyril* is sometimes ascribed to him.

DÀI TÔNG 戴侗 (from Yongjia, China; 13th cent.). Scholar and official; author of the dictionary of script *Liùshū gǔ*, which made an advance on *Shuōwén jiězì* by ordering characters into nine thematic categories.

DAL', VLADIMIR IVANOVICH (Luganskij Zavod, Russian empire [now Luhans'k, Ukraine] 1801 – Moscow 1872). Physician and writer; compiled a Russian dictionary in four volumes based on extensive fieldwork.

D'ALBERTI DI VILLANUOVA, FRANCESCO (Nice, Duchy of Savoy [now in France] 1737 – Lucca [now in Italy] 1801). Encyclopedist and translator; compiler of an Italian–French dictionary.

DANIIL OF MOSCOPOLE (DANIEL MOSCOPOLITES) (from Moscopole, Ottoman empire [now Voskopolë, Albania]; 1754–1825). Teacher; compiler of a quadrilingual dictionary in which the major language groups of the Balkans were represented.

ĐÀO DUY ANH 陶維英 (Thanh Hóa, Vietnam 1904 – Hanoi 1988). Historian and intellectual revolutionary; compiler of Chinese–Vietnamese and French–Vietnamese dictionaries.

DARD, JEAN (Maconge, France 1789 – St Louis, Senegal 1833). Educator; compiler of the first Wolof dictionary.

DARMESTER, ARSÈNE (Château-Salins, France 1846 –Paris 1888). Philologist; co-editor with A. HATZFELD (and succeeded by A. THOMAS) of a major French dictionary.

- DĀS, ŚYĀM SUNDAR (b. Varanasi, India 1875; d. 1945).** Educator, and champion of the Hindi language; compiler of a major Hindi dictionary.
- DASYPODIUS, PETRUS (PETER HASENFRATZ?) (in or near Frauenfeld, Swiss Confederacy c. 1490 – Strasburg, Holy Roman empire [now Strasbourg, France] 1555).** Schoolmaster; compiler of the first humanistic Latin–German dictionary (1535).
- DAVE, NARMADA ŚANKARA (Surat, India 1833 – Bombay [now Mumbai, India] 1886).** Founder of modern Gujarati literature; compiler of a Gujarati dictionary.
- DAVELUY, MARIE-NICOLAS-ANTOINE (Amiens, France 1818 – Galmaemot, Korea 1866).** Missionary (Missions étrangères de Paris) in Korea and titular bishop of Akka; a Latin–Korean dictionary is attributed to him or to M. PETITNICOLAS.
- DAVIS, WILLIAM JAFFERD (JEFFERD, JEFFORD) (Salisbury, Wiltshire 1810 – Grahamstown, Cape Colony [now in South Africa] 1883).** Wesleyan Methodist missionary in South Africa; compiler of the first substantial Xhosa dictionary.
- DAWES, WILLIAM (Portsmouth, Hampshire? 1762 – Antigua 1836).** Astronomer and colonial administrator; he made early wordlists of Australian languages.
- DE FRANCIS, JOHN (Bridgeport, Connecticut 1911 – Honolulu, Hawaii 2009).** Professor of Chinese; compiler of a Chinese–English dictionary.
- DELLA BELLA, ARDELIO (Foggia, Spanish empire [now in Italy] 1654 – Split, Republic of Venice [now in Croatia] 1737).** Jesuit priest; compiler of an Italian–Latin–Croatian dictionary.
- DE MAURO, TULLIO (Torre Annunziata, Italy 1932 – Rome 2017).** Linguist and politician; compiler of an eight-volume dictionary of Italian.
- DENSUȘIANU, OVID (Făgăraș, Habsburg empire [now in Romania] 1873 – Bucharest 1938).** Philologist, folklorist, and poet; compiler, with I. CANDREA, of a Romanian etymological dictionary.
- DE VRIES, MATTHIAS (Haarlem, Netherlands 1820 – Leiden, Netherlands 1892).** Philologist; founding editor of the huge monolingual Dutch dictionary *Woordenboek der Nederlandsche Taal*.
- DHANAPĀLA (from the Paramara kingdom [now in India]; 10th cent.).** Man of letters, writing in Sanskrit and Prakrit; compiler of the first extant Prakrit dictionary and of a lost Sanskrit dictionary.
- DHARAṆIDĀSA (from eastern India; early 12th cent.).** Compiler of a homonymic dictionary of Sanskrit.
- DÍAZ, FRANCISCO (b. 1606; d. Fujian?, China 1646).** Dominican missionary in China; compiler of a Chinese–Spanish dictionary transmitted in several manuscripts.

- DIEZ, FRIEDRICH CHRISTIAN** (Giessen, Landgraviate of Hesse-Darmstadt [now in Germany] 1794 – Bonn, German empire 1876). Pioneering Romance philologist; compiler of an etymological dictionary of the Romance languages.
- DIHKHUDĀ, ‘ALĪ AKBAR** (Tehran c. 1879 – Tehran 1956). Bureaucrat, editor, critic, and politician; founding compiler of the major modern dictionary of Persian.
- DIHLAVĪ, SAYYID AḤMAD** (Delhi, India 1846 – Delhi 1918). Teacher, and writer on Indian history and culture; compiler of a major Urdu dictionary.
- AL-DĪNAWĀRĪ** (from Dinawar, Abbasid Caliphate [now in Iran]; d. 281×290/894×903). Historian, mathematician, and philologist; compiler of a lexicon of Arabic plant names, part of which is in alphabetical order.
- DĪNG SHENGSHÙ 丁声树** (Dèngzhōu, China 1909 – Běijīng 1989). Editor-in-chief of the first published version of the monolingual Chinese dictionary *Xiàndài Hànyǔ cídiǎn*.
- DIAGENIANUS** (from Heraclea Pontica, Roman empire [now Karadeniz Ereğli, Turkey]; early 2nd cent. AD). Compiler of an abridgement of the Greek lexicon of PAMPHILUS, which was a source for the work of HESYCHIUS OF ALEXANDRIA and of PHOTIUS.
- DIOSKOROS OF APHRODITOPOLIS** (from Aphroditopolis, Byzantine empire [now Kom Ishqau, Egypt]; d. c. 585 AD). Advocate, notary, and curator of the nearby monastery founded by his father; compiler of Greek–Coptic glossaries.
- DIXON, ROBERT MALCOLM WARD** (b. Gloucester 1939). Professor of linguistics; editor (with W. RAMSON and M. Thomas) of *Australian Aboriginal Words in English*.
- DOBSON, EDWARD SCOTT** (Blyth, Northumberland 1918 – Gozo, Malta 1986). Art teacher and critic, and popularizer of Geordie culture; compiler of *The Geordie Dictionary*.
- DOKE, CLEMENT MARTYN** (Bristol, England 1893 – East London, South Africa 1980). Missionary (Baptist Missionary Society) and historian of the study of the Bantu languages; compiler, with B. VILAKAZI, of a significant Zulu–English dictionary.
- DOMINGUEZ, RAMÓN JOAQUÍN** (Verín, Spain 1811 – Madrid 1848). Teacher of French; compiler of a Spanish dictionary.
- DONIACH, NAKDIMON SHABBETHAY** (London 1907 – Oxford, Oxfordshire 1994). Intelligence officer, civil servant, and teacher; compiler of bilingual dictionaries of Russian, Arabic, and Hebrew with English.
- DOROSZEWSKI, WITOLD** (Moscow 1899 – Warsaw 1976). Linguist; editor of the large monolingual Polish dictionary *Słownik języka polskiego*.
- DOUGLAS, CARSTAIRS** (Kilbarchan, Scotland 1830 – Xiàmén, China 1877). Presbyterian missionary in China; compiler of a Chinese–English dictionary.

- DOUGLAS, WILFRID HENRY** (Belfast, Ireland [now in Northern Ireland] 1917 – Mandurah?, Australia 2004). Missionary (United Aborigines Mission) in Australia; compiler of wordlists of Bardi and Western Desert Language.
- DOZY, REINHART PIETER ANNE** (Leiden, Netherlands 1820 – Leiden 1883). Historian; compiler of an Arabic–French *Supplément aux dictionnaires arabes*.
- DU CANGE, CHARLES DU FRESNE** (Amiens, France 1610 – Paris 1688). Medievalist; compiler of the first large-scale scholarly dictionaries of medieval Latin and Byzantine Greek.
- DUDEN, KONRAD ALEXANDER FRIEDRICH** (Lackhausen, Prussian Rhineland [now in Germany] 1829 – Sonnenberg, Germany 1911). Teacher, and eponym of a major brand of German language reference books; compiler of the first official orthographic dictionary in the unified German empire after 1871.
- DUNCAN, ANDREW** (from Scotland; fl. 1571–1626). Teacher and minister of religion; compiler of the first printed wordlist of Scots.
- DUNGLISON, ROBLEY** (Keswick, Cumberland 1798 – Philadelphia, Pennsylvania 1869). Physician; compiler of a medical dictionary and an early glossary of Americanisms.
- ĐURĐEVIĆ, BARTOL (BARTOLOMEJ)** (Mala Mlaka?, Croatia c. 1506 – Rome? c. 1566). Author of several works in Latin inspired by a spell of captivity among the Turks; compiler of a Latin–Croatian wordlist.
- ECKART, ANSELM** (Mainz [now in Germany] 1721 – Polotsk, Russian empire [now in Belarus] 1809). Jesuit missionary in Brazil; compiler of a Tupi dictionary.
- EDKINS, JOSEPH** (Nailsworth, Gloucestershire 1823 – Shànghǎi, China 1905). Missionary (London Missionary Society) in China and philologist; compiler of *A Vocabulary of the Shanghai Dialect*.
- EGEDE, POUL HANSEN** (Kabelvåg, Kingdom of Denmark [now in Norway] 1708 – Copenhagen 1789). Missionary and translator; compiler of a Greenlandic dictionary.
- ELGER, GEORG** (Wolmar, Polish-Lithuanian Commonwealth [now Valmiera, Latvia] 1585 – Dünaburg, Polish-Lithuanian Commonwealth [now Daugavpils, Latvia] 1672). Jesuit priest; compiler of a Polish–Latin–Latvian dictionary.
- ELIAS, ELIAS ANTOON** (from Cairo; d. 1952). Publisher; compiler of widely used English–Arabic and Arabic–English dictionaries.
- ELKIN, ADOLPHUS PETER** (Maitland, New South Wales 1891 – Sydney, Australia 1979). Anglican clergyman and anthropologist; compiler of a dictionary of Ungarinyin with H. COATE.
- ELLIOTT, WILLIAM ALLAN** (Cheltenham, Gloucestershire 1851 – England? 1932). Missionary (London Missionary Society) in what is now Zimbabwe; compiler of a dictionary of Ndebele and Shona.

- ELYOT, THOMAS** (southern England c. 1490 – Carleton, Cambridgeshire 1546). Humanist and diplomat; his English adaptation of A. CALEPINO's *Dictionarium* was the first post-medieval Latin–English dictionary, and the first English book to have *Dictionary* as its title.
- ENDZELĪNS, JĀNIS** (Kauguri, Russian empire [now in Latvia] 1893 – Koknese, Soviet Union [now in Latvia] 1961). Linguist; co-editor of the iconic Latvian–German dictionary started by K. MÜHLENBACH, and co-editor with E. HAUZENBERGA of its supplement.
- ES'AD EFENDI, MEHMET** (Istanbul 1685 – Istanbul 1753). Ottoman religious, poet, and composer; compiler of the *Lehçetü'l-luğât*.
- ESTIENNE, HENRI** (Paris 1531 – Lyon, France 1598). Printer and humanist, son of R. ESTIENNE; compiler of the Greek–Latin *Thesaurus Graecae linguae*, the most ambitious European dictionary of the sixteenth century.
- ESTIENNE, ROBERT** (Paris c. 1503? – Geneva [now in Switzerland] 1559). Printer and humanist, father of H. ESTIENNE; compiler of the large and learned Latin dictionary *Linguae Latinae thesaurus*, which was influential directly and through its bilingual derivatives.
- EVEN-SHOSHAN (né ROSENSTEIN), AVRAHAM** (Minsk 1906 – Jerusalem 1984). Educator; compiler of two of the leading Hebrew dictionaries of the later twentieth century, his work being continued by M. AZAR.
- FABER (SCHMIDT), BASILIUS** (Sorau, kingdom of Bohemia [now Żary, Poland] c. 1520 – Erfurt, Holy Roman empire [now in Germany] 1575×1576). Compiler of a widely disseminated Latin dictionary for the use of students.
- FABRA I POCH, POMPEU** (Gràcia, Spain 1868 – Prades, France 1948). Engineer and linguist; compiler of an important Catalan dictionary.
- FABRICIUS, OTTO** (Rudkøbing, Denmark 1744 – Copenhagen 1822). Lutheran missionary in Greenland and naturalist; compiler of a Greenlandic dictionary.
- FABVRE, BONAVENTURE** (Troyes, France 1655 – Québec, New France [now Quebec City, Canada] 1700). Jesuit missionary in what is now Canada; compiler of an early dictionary of Montagnais.
- FAKHR-AL-DĪN MUBĀRAKSHĀH QAWWĀS GHAZNAWĪ** (from Delhi; d. after 743/1342). Compiler of the oldest extant Persian dictionary from what is now India.
- FAKHR-AL-DĪN NAKHJAWĀNĪ** (14th cent.). Compiler of the oldest Persian–Turkish glossary.
- FALLON, SAMUEL WILLIAM** (Calcutta [now Kolkata, India] 1817 – London 1880). Educator; compiler of major dictionaries of English and Hindustani.
- FĀNG YǐZHì 方以智** (from Tongcheng, China; 1611–1671). Prominent literary and intellectual figure; compiler of the encyclopedic dictionary *Tōngyǎ*.

- AL-FĀRĀBĪ, IŠĤĀQ BIN IBRĀHĪM (from Fārāb [now in Kazakhstan]; d. 350/961?). Teacher; compiler of a morphologically based Arabic lexicon in which words are arranged by their final letters.
- FARMER, JOHN STEPHEN (Bedford, Bedfordshire 1854 – West Bergholt, Essex 1916). Miscellaneous writer and literary scholar; compiler of dictionaries of English slang and of Americanisms.
- FAVORINO, GUARINO (GUARINO OF FAVERA, PHAVORINUS, VARINUS CAMERS) (b. Pievefavera, near Camerino [now in Italy] 1445×1450?; d. 1537). Benedictine monk, humanist, and bishop of Nocera: compiler of an important Greek lexicon.
- FAY, WILLIAM EDWARDS (Louisville, Kentucky 1855 – Cleveland, Ohio 1907). Missionary (American Board of Commissioners for Foreign Missions) in Angola; compiler, with W. SANDERS, of the first Umbundu dictionary.
- FEBRÈS, ANDRÉS DE (Manresa, Spain 1731 – Genoa [now in Italy] 1790). Jesuit missionary in South America; compiler of Mapudungun wordlists.
- FENNE (FONNE?), TÖNNIES (from Lübeck? fl. 1586–1627). Commercial assistant; author of a Low German–Russian manual.
- FENTON, JAMES (b. Drumdarragh, Antrim 1931). Poet in Ulster Scots; compiler of the Ulster Scots dictionary *The Hamely Tongue*.
- FÉRAUD, JEAN-FRANÇOIS (Marseille, France 1725 – Marseille 1807). Jesuit priest; compiler of a major dictionary of French.
- FERNÁNDEZ DE PALENCIA, ALFONSO (province of Soria, Castile [now in Spain] 1423 – Seville 1492). Historian and translator; compiler of the first printed Latin–Spanish dictionary.
- FERRAGUTO, PIETRO (Messina, Sicily c. 1580 – Naples, Spanish empire [now in Italy] 1656). Jesuit; compiler of an Italian–Turkish dictionary.
- FIGUEIREDO, CÂNDIDO DE (Lobão da Beira, Portugal 1846 – Lisbon 1925). Writer and public servant; compiler of a large Portuguese dictionary.
- FIGUEREDO, JUAN DE (Huancavelica, Spanish empire [now in Peru], 1646 – Lima 1723). Jesuit missionary in Peru; revised the Aymara grammar and wordlists of D. DE TORRES RUBIO, adding a wordlist of the Chinchaysuyo variety of Quechua.
- FIGUEROA, JERÓNIMO DE (Mexico City? 1604 – Mexico City 1683). Jesuit missionary in what is now Mexico; he is said to have compiled a ‘Vocabulario copioso de la lengua tepehuana y tarahumara’.
- AL-FĪRŪZĀBĀDĪ (Kāzarūn, Mongol empire [now in Iran] 729/1329 – Zabīd, Yemen 817/1415). Polymath; compiler of the Arabic lexicon *al-Qāmūs al-Muḥīṭ*, which influenced Arabic and Persian lexicography into the nineteenth century (cf. the work of G. FREYTAG); its name became synonymous with ‘dictionary’.

- FISCHER, AUGUST WILHELM HERMANN GUSTAV (Halle (Saale), Kingdom of Prussia [now in Germany] 1865 – Leipzig, Germany 1949).** Professor of Oriental philology; he left unpublished materials for a very large historical dictionary of Arabic.
- FLIERL, JOHANN (Fürnried, Kingdom of Bavaria [now in Germany] 1858 – Neuendettelsau, Germany 1947).** Lutheran missionary in Australia and New Guinea; compiler of a comparative Diyari–Wangkangurru wordlist.
- FLORIO, JOHN (London 1553 – Fulham, Middlesex [now part of London] 1625).** Translator and teacher; compiler of an Italian–English dictionary.
- FORTE, JUAN (Tarrassa, Spain 1575 – El Zape, Spanish empire [now in Mexico] 1616).** Jesuit missionary in what is now Mexico; he is said to have compiled a wordlist of Northern Tepehuan.
- FRANCARD, MICHEL (b. Bastogne, Belgium 1951).** Linguist; compiler of a dictionary of Belgicisms in French.
- FRANCOSINI, LORENZO (b. in Florence or Castelfiorentino, Tuscany [now in Italy]; d. after 1645?).** Compiler of a bidirectional Italian and Spanish dictionary.
- FREYTAG, GEORG WILHELM (Lüneburg, Electorate of Hanover [now in Germany] 1788 – Bonn, Kingdom of Prussia [now in Germany] 1861).** Professor of Oriental languages; compiler of an Arabic–Latin dictionary based on the *Qāmūs* of AL-FĪRŪZĀBĀDĪ.
- FRISCH, JOHANN LEONHARD (Sulzbach [now in Germany] 1666 – Berlin 1743).** Educator, philologist, and entomologist; compiler of a German dictionary rich in technical vocabulary.
- FRISIUS (FRIES), JOHANNES (Greifensee, Swiss Confederacy 1505 – Zürich, Swiss Confederacy 1565).** Teacher, translator, and musicologist; compiler, with P. CHOLINUS, of a Latin–German dictionary.
- FURETIÈRE, ANTOINE (Paris 1619 – Paris 1688).** Man of letters and academician; compiler of the *Dictionnaire universel* (1690), a large monolingual French dictionary perceived as a rival to the *Dictionnaire de l'Académie française*.
- FYCHAN, SIMWNT (SIMON VAUGHAN) (d. north Wales 1606).** Poet; compiler of a monolingual Welsh wordlist which is a late example of the bardic tradition.
- GALE, JAMES SCARTH (Alma, Canada 1863 – Bath, Somerset 1937).** Protestant missionary in Korea and Bible translator; compiler of a Korean–English dictionary.
- GANGLER, JEAN-FRANÇOIS (Luxemburg 1788 – Luxemburg 1856).** Teacher, poet, and police superintendent; compiler of the first printed dictionary of Luxemburgish.
- AL-ĠARBĀDAQĀNĪ (fl. in the second half of the 4th/10th cent.).** Philologist; author of *al-Rawḥa*, the most extensive lexicon on *ḏ* and *z* in the Arabic tradition.

- GARZONI, MAURIZIO (from Turin; 1734–1804).** Dominican missionary in the region of Mosul in the Ottoman empire; the Italian–Kurdish wordlist in his *Grammatica e vocabolario della lingua Kurda* is the first of Kurdish and a European language.
- GATSCHET, ALBERT SAMUEL (Beatenberg, Switzerland 1832 – Washington, DC 1907).** Philologist and ethnologist; he published wordlists of several languages of the south-western United States and compiled a manuscript wordlist of Western Apache.
- AL-ĞAWĀLĪQĪ (d. 540/1145).** Compiler of a lexicon of foreign loanwords in Arabic arranged alphabetically.
- AL-ĞAWHARĪ (d. c. 400/1010).** Compiler of the first Arabic lexicon in which the order of the radicals of the root, starting with the last, is the sole criterion of arrangement.
- ĞEMĀL-AL-DĪN IBNĪ MŪHENNĀ (IBN MUHANNĀ) (late 13th or early 14th cent.).** Compiler of the Arabic–Persian, Arabic–Turkic, and Arabic–Mongolian dictionary called *Ībnī Mūhennā Lūgati*.
- GÉNIBREL, JEAN-FRANÇOIS-MARIE (Castres, France 1851 – Saigon [now Ho Chi Minh City], Vietnam 1914).** Missionary (Missions étrangères de Paris) in Vietnam; compiler of Vietnamese–French dictionaries.
- GEROV, NAJDEN (NAJDEN GEROV XADŽIDOBREVIČ) (Koprivštica, Ottoman empire [now in Bulgaria] 1823 – Plovdiv, Bulgaria 1900).** Writer and linguist; compiler of the first comprehensive Bulgarian dictionary.
- GHĀLIB, MĪRZĀ ASADULLĀH KHĀN (Agra, Mughal empire [now in India] 1797 – Delhi 1869).** Poet in Persian and Urdu; author of a critique of the seventeenth-century Persian dictionary *Burhān-i Qāṭi*.
- GHARIEB, GHARIEB MOHAMED (b. Ismailia, Egypt 1940).** Compiler, with G. KRAHL, of an Arabic–German dictionary.
- GHĀZĪ AL-DĪN ḤAYDAR SHĀH (from Awadh [now in India]; 1769–1827).** Sovereign of Awadh; alleged author of the Persian dictionary *Haft Qulzum* ('Seven seas').
- GILBERTI, MATURINO (Poitiers, France 1507×1508 – Tzintzuntzan, Spanish empire [now in Mexico] 1585).** Franciscan missionary in what is now Mexico; compiler of the first printed dictionary of Tarascan.
- GILCHRIST, JOHN BORTHWICK (Edinburgh 1759 – Paris 1841).** Teacher; compiler of a dictionary of Hindustani and English.
- GILES, HERBERT ALLEN (Oxford, Oxfordshire 1845 – Cambridge, Cambridgeshire 1935).** Sinologist; compiler of a Chinese–English dictionary.
- GILG, ADÁN (b. Römerstadt, Habsburg empire [now Rýmařov, Czech Republic] 1653; d. in Mexico? in or after 1710).** Jesuit missionary in what is

now Mexico; a 'Vocabulario de las lenguas eudeve, pima, y seris' has been attributed to him.

GIORDA, JOSEPH (province of Piedmont, Kingdom of Sardinia [now in Italy] 1823 – Helena, Montana 1882). Jesuit missionary in the United States; compiler of a Kalispel dictionary.

GODEFROY, FRÉDÉRIC-EUGÈNE (Paris 1826 – Lestelle-Bétharram, France, 1897). Literary scholar; compiler of a ten-volume dictionary of medieval French.

GOLIUS, JACOBUS (JACOB GOOL) (The Hague, Netherlands 1596 – Leiden, Netherlands 1667). Orientalist and mathematician; compiler of a Persian dictionary printed after his death by E. CASTELL.

GONÇALVES, JACOME (Goa [now in India] 1676 – Bolawatta, Ceylon [now Sri Lanka] 1742). Oratorian missionary in Ceylon and author of Christian texts in Sinhalese and Tamil; compiler of a Portuguese–Tamil–Sinhalese dictionary.

GONZÁLEZ DE NÁJERA, GASPAR (fl. Yucatán, Spanish empire [now in Mexico] 1580×1603). Franciscan missionary in what is now Mexico; an early Spanish–Maya wordlist is attributed to him.

GONZÁLEZ HOLGUÍN, DIEGO (Cáceres, Spain 1552 – Mendoza, Spanish empire [now in Argentina] 1618). Jesuit missionary in South America; compiler of a dictionary of Quechua.

GORO D'AREZZO (from Tuscany [now in Italy]; 14th cent.). Grammarian; compiler of a Latin–Tuscan wordlist.

GOVE, PHILIP BABCOCK (Concord, New Hampshire 1902 – Warren, Michigan 1972). Editor of *Webster's Third New International Dictionary of the English Language*.

GRAGS PA RGYAL MTSHAN (from Tibet; 1242–1346). Teacher; translator into Tibetan, with KĪRTICANDRA, of the *Amarakośa* of AMARASIMHA, and the commentary on it by SUBHŪTICANDRA.

GRAVIER, JACQUES (Moulins, France, 1651 – Mobile, New France [now in Alabama] 1708). Jesuit missionary in what is now the United States; compiler of a dictionary of Miami-Illinois.

GREY, GEORGE (Lisbon 1812 – London 1898). Explorer and colonial governor; compiler of an early dictionary of Noongar.

GRIMM, JACOB LUDWIG CARL (Hanau, Hesse-Kassel [now in Germany] 1785 – Berlin 1863). Philologist and folklorist; together with his brother W. GRIMM, founder of the great *Deutsches Wörterbuch*.

GRIMM, WILHELM CARL (Hanau, Hesse-Kassel [now in Germany] 1786 – Berlin 1859). Philologist and folklorist; together with his brother J. GRIMM, founder of the great *Deutsches Wörterbuch*.

GÙ YEWÁNG 顧野王 (from Jiāngsū, China; 519–581). Court scholar, compiler of the seminal dictionary *Yùpiān*; the original text is lost, but the recension by CHÉN PÉNGNIÁN is extant.

- GUÐMUNDUR ANDRÉSSON (from Iceland; d. Copenhagen 1654).** Compiler of the first Icelandic dictionary to include material from contemporary usage.
- GUEVARA, MIGUEL DE (Viceroyalty of New Spain, Spanish empire [now Mexico] 1585 – Michoacán, Spanish empire [now in Mexico] 1646).** Augustinian friar and poet; compiler of an introduction to Matlatzinca which includes lexicographical material.
- GUIGNES, CHRÉTIEN-LOUIS-JOSEPH DE (Paris 1759 – Paris 1845).** Diplomat and travel writer; published an unacknowledged adaptation of a Chinese dictionary by B. BROLLO.
- GÜNTHER, JAMES (JAKOB WILHELM) (Oberschwandorf, Principality of Württemberg [now in Germany] 1806 – Mudgee, New South Wales 1879).** Missionary (Church Mission Society) in Australia; compiler of a posthumously published Wiradjuri wordlist.
- GUR (né GRAZOVSKI), YEHUDA (Minsk 1862 – Tel Aviv 1950).** Essayist and author of textbooks; compiler of multiple Hebrew dictionaries.
- HAACK, FRIEDRICH WILHELM (Crottingen, Kingdom of Prussia [now Kretingalė, Lithuania] 1706 – Pillkallen, Kingdom of Prussia [now Dobrovol'sk, Kaliningrad Oblast, Russia] 1754).** Compiler of a bidirectional dictionary of Lithuanian and German.
- HAAS, MARY ROSAMOND (Richmond, Indiana 1910 – Berkeley, California 1996).** Linguist; many dictionaries of indigenous American languages were produced under her direction at the University of California, Berkeley, and she compiled a Thai–English dictionary.
- HABDELIĆ, JURAJ (Staro Čiče, Kingdom of Hungary [now in Croatia] 1609 – Zagreb 1678).** Jesuit priest and author; compiler of a Kajkavian–Latin dictionary.
- HAENSCH, GÜNTHER (Munich, Germany 1923 – Stadtbergen, Germany 2018).** Linguist; compiler, with R. WERNER, of dictionaries of Americanisms in Spanish.
- HAI GAON (Pumbadita, Abbasid Caliphate [now in Iraq] 939 – Baghdad 1038).** Teacher and spiritual leader; last Gaon of the Pumbadita Jewish Academy in Baghdad; compiler of an anagrammatically arranged Hebrew dictionary.
- HAILE, BERARD (Akron, Ohio 1874 – Santa Fe, New Mexico 1961).** Franciscan missionary to the Navajo people; compiler of Navajo–English and English–Navajo dictionaries.
- HALĀYUDHA (b. before 950; fl. Dhārā, Paramara kingdom [now Dhār, India] after 974).** Compiler of a synonymic Sanskrit dictionary.
- HALBERTSMA, JUSTUS HIDDUS (Grou, Netherlands 1789 – Deventer, Netherlands 1869).** Promoter of the Frisian language; compiler of a pioneering, and unfinished, Frisian dictionary.

- HALE, HORATIO EMMONS** (Newport, New Hampshire 1817 – Clinton, Canada 1896). Philologist; he made wordlists of North American and Australian languages.
- AL-ḤALĪL BIN AḤMAD** (d. 175/791). Grammarian; compiler of the first Arabic lexicon, *Kitāb al-ʿAyn*, which is arranged phonetically and whose introduction lays the foundations of the lexicographical tradition (it is also sometimes ascribed to AL-LAYṬ).
- ḤĀLIŞ, YŪSUF** (19th cent.). Compiler of the French–Turkish verse wordlist *Mitfāh-ı Lisān*.
- HALL, JOAN HOUSTON** (b. Ohio 1946). Chief editor (in succession to F. CASSIDY) of the *Dictionary of American Regional English*.
- HALLIWELL-PHILLIPPS, JAMES ORCHARD** (London 1820 – Hollingbury Copse, Sussex 1889). Antiquary and literary scholar; compiler of *Dictionary of Archaic and Obsolete Words*.
- HALMA, FRANÇOIS** (Langerak, Netherlands 1653 – Leeuwarden, Netherlands 1722). Printer and publisher; compiler of a Dutch–French dictionary which became the basis for the first Dutch–Japanese dictionary.
- HANAGUID, SAMUEL** (Córdoba, al-Andalus [now in Spain] 993 – Granada, al-Andalus [now in Spain] 1055×1056). Poet and statesman; compiler of a lost Hebrew–Arabic biblical dictionary.
- HAN DÀOZHĀO** 韓道昭 (from Zhengding, China; fl. 1211). Compiler of the rhyme dictionary *Sishēng piānhǎi*, an adaptation of the *Guǎngyùn* (see CHÉN PÉNGNIÁN), begun by his father Han Xiaoyan 韓孝彥.
- ḤAN DEHLAWĪ, called DĀRĀWĀL** (b. in Delhi; settled in the principality of Malwa [now in India]; 15th cent.). Scholar; compiler of a Persian dictionary.
- HANDT, JOHANN CHRISTIAN SIMON** (Aken, Kingdom of Prussia 1783 – Geelong, Victoria 1863). Missionary (Church Mission Society) in Australia; compiler, with W. Watson, of an unpublished Wiradjuri wordlist.
- HANXLEDEN, JOHANN ERNST** (ARNOS PADIRI) (Ostercappeln, Bishopric of Osnabrück [now in Germany] 1681 – Pazhuvil, Kerala [now in India] 1732). Jesuit missionary in India, poet in Malayalam, and Sanskritist; he compiled a Malayalam dictionary, and made a transcript of the *Amarakośa* of AMARASĪMHA which would be a basis for the partial publication of the *Amarakośa* by PAULINUS A SANCTO BARTHOLOMAEO.
- ḤAQ, ʿABDUL** (Hapur, India 1870 – Karachi, Pakistan 1961). Champion of the Urdu language, and founding editor of a major historical dictionary of Urdu.
- AL-HARAWĪ, ABŪ ʿUBAYD** (d. 401/1011). Compiler of a dictionary which combines strange usage (*ḡarīb*) in the Qurʾān and the Prophetic tradition of Arabic texts.
- AL-ḤARBĪ** (d. 285/898). Compiler of a voluminous lexicon of rare Arabic words in the *Ḥadīṭ*.

- AL-ḤARĪRĪ** (near **Baṣra**, **Abbasid Caliphate** [now in **Iraq**] **446/1054** – **Baṣra** **516/1122**). Philologist and man of letters; author of a collection of linguistic errors in Arabic made by the elite or people of distinction.
- HARṢAKĪRTI** (fl. **c. 1535–1610**). Jain leader and polymath; compiler of a Sanskrit synonymic lexicon.
- HARTMANN, ANDREW M. (ANDREAS)** (**county of Tyrol, Austria** **1851** – **Southern Rhodesia** [now **Zimbabwe**] **1928**). Jesuit missionary in what is now Zimbabwe; compiler of the first dictionary dedicated to a Shona language.
- HATZFELD, ADOLPHE** (**Paris** **1824** – **Paris** **1900**). Linguist and philosopher; co-editor with **A. DARMESTETER** (and subsequently with **A. THOMAS**) of a major French dictionary.
- HAUZENBERGA, EDĪTE (EDĪTE HAUZENBERGA-ŠTURMA, EDITH HAUSENBERG)** (**Svētciems, Russian empire** [now **Latvia**] **1901** – **Bonn, West Germany** [now **Germany**] **1983**). Linguist; co-editor with **J. ENDZELĪNS**, of the supplement to the Latvian–German dictionary by **K. MÜHLENBACH**.
- HAYYIM, SULAYMAN** (**Tehran** **1886** – **Tehran** **1970**). Translator and playwright, compiler of bilingual dictionaries of Persian with English, French, and Hebrew.
- ḤAYYŪJ, YEHUDA** (**Fez, Fatimid Caliphate** [now in **Morocco**] **mid 10th cent.** – **Córdoba, al-Andalus** [now in **Spain**] **late 10th or early 11th cent.**). Compiler of grammatical dictionaries of Hebrew.
- HEATH, JEFFREY G. (b. Exeter, New Hampshire** **1949**). Linguist; compiler of dictionaries of Nunggubuyu and other Australian and African languages.
- HECKEWELDER, JOHN GOTTLIEB ERNESTUS** (**Bedford, Bedfordshire** **1743** – **Bethlehem, Pennsylvania** **1823**). Moravian missionary in the United States; compiler of a Mahican wordlist.
- HEMACANDRA (Dhandhukā, Gujarat** [now in **India**] **1087×1088** – **Gujarat** **1172**). Jain monk and polymath; compiler of a synonymic Sanskrit dictionary and of a Prakrit dictionary.
- HENISCH, GEORG** (**Bartfeld, Habsburg empire** [now **Bardejov, Slovakia**] **1549** – **Augsburg, Holy Roman empire** [now in **Germany**] **1618**). Physician and teacher; compiler of an elaborate unfinished German dictionary.
- HERBERT, THOMAS** (**York, Yorkshire** **1606** – **York** **1682**). Traveller and government official; published two Persian glossaries.
- HERENNIUS PHILO** (**from Byblos, Roman empire** [now in **Lebanon**]; **c. 70** – **c. 160**). Historian; compiler of a lost Greek synonym dictionary on which many extant Byzantine synonymica were based.
- HESYCHIUS OF ALEXANDRIA (from Alexandria, Eastern Roman empire** [now in **Egypt**]; **5th or 6th cent.**). Compiler of a major alphabetical Greek lexicon, which drew via the abridgement of **DIOGENIANUS** on the lost work of **PAMPHILUS**.

- HINDS, MARTIN (b. Penarth, Wales 1941; d. 1988).** University lecturer in Arabic; compiler, with E. BADAWI, of a dictionary of Egyptian Arabic.
- HIRAETHOG, GRUFFUDD (b. Llangollen, Wales; d. Llangollen 1564).** Poet and herald; compiler of a monolingual Welsh wordlist in the bardic tradition.
- HODGE, DAVID McKELLOP (Choska, Indian Territory [now near Coweta, Oklahoma] 1841 – Tulsa, Oklahoma 1920).** Creek Nation politician; compiler with R. LOUGHRIDGE, of a dictionary of Muskogee.
- HOGARTH, RICHARD (b. Westmorland?; d. London 1718).** Schoolmaster and writer of textbooks; produced an abridged translation of the etymological dictionary of S. SKINNER.
- HOIJER, HARRY (Chicago, Illinois 1904 – Santa Monica, California 1976).** Anthropologist and linguist; compiler of a Navajo dictionary and of an unpublished lexical file on Western Apache.
- HOLDERMANN, JEAN-BAPTISTE DANIEL (Strasbourg, France 1694 – Istanbul 1730).** Jesuit and Orientalist; wrote a Turkish grammar which includes a wordlist.
- HOLM, JOHN (Jackson, Michigan 1943 – Azeitão, Portugal 2015).** Creolist; editor (with A. W. Shilling) of the *Dictionary of Bahamian English*.
- HOOGLAND, JAN (b. Rotterdam, Netherlands 1957).** Professor of Arabic (and contributor to this book); compiler of Arabic–Dutch and Dutch–Arabic dictionaries.
- ḤOSAYN ZAWZANĪ (d. 486/1093).** Compiler of an Arabic–Persian dictionary, and of the first Qurʾān glossary in Persian.
- HOUAISS, ANTÔNIO (Rio de Janeiro, Brazil 1915 – Rio de Janeiro 1999).** Diplomat, translator, and encyclopedist; founding editor of a major dictionary of Portuguese.
- HOUTMAN, FREDERIK DE (Gouda, Habsburg Netherlands 1571 – Alkmaar, Netherlands 1627).** Explorer and astronomer; compiler of the first printed Malay dictionary.
- HRINČENKO, BORYS DMYTROVYČ (Vilʹxovyj Jar, Russian empire [now in Ukraine] 1863 – Ospedaletti, Italy 1910).** Political activist, ethnographer, and writer; compiler of an important Ukrainian–Russian dictionary.
- HUE, GUSTAVE (Bernières-sur-Mer, France 1870 – Hanoi 1946).** Missionary (Missions étrangères de Paris) in Vietnam; compiler of a Vietnamese–Chinese–French dictionary.
- HUGUTIO (UGUCCIONE) (from the Italian peninsula; 12th cent.).** Compiler of the derivationally structured monolingual Latin *Derivationes*; perhaps, but not certainly, identical with the bishop and canonist Hugutio of Pisa (Pisa c. 1140 – Ferrara 1210).
- HULBERT, JAMES ROOT (Eldora, Iowa 1884 – Chicago, Illinois 1969).** Professor of English; editor (with W. CRAIGIE) of the *Dictionary of American English*.

- HUMPHREYS, DAVID** (Derby, Connecticut 1752 – New Haven, Connecticut 1818). Soldier, politician, diplomat, and poet; his play *The Yankey in England* includes an early printed glossary of American English.
- HUNDERTPFUND, ROCHUS** (Bregenz, Austria 1709 – Bregenz 1777). Jesuit missionary in Brazil; possible compiler of a Tupi dictionary.
- HUNTER, JOSEPH** (Sheffield, Yorkshire 1783 – London 1861). Antiquary; compiler of an early dialect dictionary.
- HUNTER, WILLIAM** (Montrose, Scotland 1755 – Java 1812). Surgeon and Orientalist; compiler, with J. Taylor, of a Hindustani dictionary.
- ḤUSAYN INJŪ** (Shiraz, Safavid empire [now in Iran], 16th cent. – Agra, Mughal empire [now in India] 1626). Compiler of a Persian dictionary.
- IBN AL-ANBĀRĪ, ABŪ BAKR** (al-Anbār, Abbasid Caliphate [now in Iraq] 231/885 – Baghdad 328/940). Philologist; compiler of the largest collection of Arabic words which have two contradictory meanings.
- IBN BALʿAM, YEHUDA** (Abu Zakariya Jahja ibn Balʿam) (from Iṣbīliya, al-Andalus [now Seville, Spain]; fl. 3rd quarter of the 10th cent.). Rabbi and exegete; author of works on Hebrew grammar containing wordlists, and of a dictionary of Hebrew homonyms.
- IBN DURAYD** (Baṣra, Abbasid Caliphate [now in Iraq] 223/837 – Baghdad 321/933). Compiler of an Arabic lexicon in which roots are alphabetically arranged in chapters based on the number of their radicals.
- IBN EZRA, MOSHE** (from Granada; c. 1055 – c. 1138). Rabbi, poet, and philosopher; compiler of a Hebrew homonym dictionary in verse.
- IBN FĀRIS** (d. Rayy, Buyid empire [now in Iran] 395/1004). Compiler of a unique Arabic lexicon which distinguishes between the various meanings that a root can indicate.
- IBN ḤĀLAWAYHI** (d. 370/980). Philologist; compiler of a collection of Arabic morphological patterns and phenomena that are only rarely attested.
- IBN JANĀḤ, YONA** (Córdoba, al-Andalus [now in Spain] c. 990 – Zaragoza, al-Andalus [now in Spain] c. 1050). Physician and grammarian; compiler of the great Hebrew–Arabic lexicon *Kitāb al-ʿUṣūl*, which was translated by Y. IBN TIBBON to make the monolingual Hebrew dictionary *Sefer ha-šorašim*.
- IBN MANZŪR** (630/1233 – 711/1311×1312). Compiler of the famous *Lisān al-ʿArab* (20 vols.), which incorporates 5 earlier Arabic lexica and includes 9,273 lemmata.
- IBN AL-QATTĀʿ** (Sicily 433/1041 – Egypt 515/1121). Anthologist, historian, and grammarian; compiler of an Arabic lexicon specializing in verbal patterns.
- IBN QURAYSH, YEHUDA** (b. Tahort, Rustamid Kingdom [now in Algeria]; fl. Fez, Fatimid Caliphate [now in Morocco], first half of the 10th cent.). Compiler of comparative wordlists of Hebrew with Aramaic and Arabic.

- IBN AL-ŠAĠARĪ** (b. in Baghdad; d. Baghdad 542/1148). Philologist and anthologist; compiler of a wordlist of Arabic homonyms.
- IBN SĪDA** (b. in Murcia, al-Andalus [now in Spain]; d. Denia?, al-Andalus [now in Spain] 458/1066). Philologist; compiler of the last major phonetically arranged Arabic lexicon.
- IBN AL-SIKKĪT** (Baghdad c. 186/802 – Baghdad c. 44/858). Philologist; author of *Iṣlāḥ al-manṭiq*, one of whose chapters is the earliest specimen of arranging Arabic roots based on their last radical.
- IBN TIBBON, YEHUDA** (Granada, al-Andalus [now in Spain] 1120 – Marseille, County of Provence [now in France] 1190). Physician and translator; his translation of the Hebrew–Arabic lexicon *Kitāb al-ʿUṣūl* of Y. IBN JANĀH to create the monolingual Hebrew dictionary *Sefer ha-šorašim* made it available to Jews outside the Arabic-speaking world.
- IDRISS, SOUHEIL** (Beirut 1925 – Beirut 2008). Novelist and journalist; compiler, with J. ABDEL-NOUR, of a French–Arabic dictionary.
- AL-IŠFAHĀNĪ, ḤAMZA BIN AL-ḤASAN** (Isfahan, Abbasid Caliphate [now in Iran] c. 280/893 – Isfahan after 351/962). Philologist and historian; compiler of the first Arabic book on proverbs to introduce alphabetical ordering.
- ISHIKAWA MASAMOCCHI** (Edo [now Tokyo] 1753 – Edo 1830). Poet in the satirical *kyōka* genre and *kokugaku* (nativist philology) scholar; compiler of a Japanese literary dictionary.
- ISIDORE** (Cartagena or Hispalis, Kingdom of the Visigoths [now in Spain] c. 560 – Seville 636). Bishop; compiler of the encyclopedic Latin *Etymologiae*.
- JAKOBSEN, JAKOB** (Tórshavn, Faroe Islands 1864 – Copenhagen 1918). Philologist; compiler of pioneering printed dictionaries of Faroese and Norn.
- JAKUBAŠ, FILIP** (Kleinpostwitz/Bójswey, German empire [now Germany] 1895 – Schirgiswalde/Šērachow, East Germany [now Germany] 1966). Teacher; compiler of an Upper Sorbian–German dictionary.
- JAMIESON, JOHN** (Glasgow, Scotland 1759 – Edinburgh 1838). Minister of religion; compiler of the first scholarly dictionary of Scots.
- JENSSEN, CHRISTEN** (d. Askvoll, Norway 1653). Minister of religion; compiler of the first printed dictionary of Norwegian.
- JEROME** (EUSEBIUS HIERONYMUS) (Stridon, Western Roman empire [perhaps now in Croatia or Slovenia] c. 347 – Bethlehem, Eastern Roman empire [now in Palestine] 420). Biblical translator and ascetic, made a wordlist of Hebrew names from the Bible with Latin interpretamenta.
- JETTÉ, JULES** (Montreal, Canada 1864 – Akulurak, Alaska 1927). Jesuit missionary in Canada and Alaska; his Koyukon lexicographical materials were the basis for the dictionary brought to completion by E. JONES and published in 2000.

- JOHN (YUHANNA) SAMANNUDI (from Egypt; fl. 1235–1257).** Bishop, grammarian, and interpreter; compiler of a Coptic–Arabic glossary.
- JOHNSON, FRANCIS (b. 1795×1796; d. Hertford, Hertfordshire 1876).** Compiler of a Persian–English dictionary based on that of J. RICHARDSON.
- JOHNSON, FREDERICK (England? 1889 – Tanganyika? [now in Tanzania] 1937).** Lay missionary (Universities Mission to Central Africa) in east Africa and colonial Education Officer; compiler of the first monolingual Swahili dictionary.
- JOHNSON, SAMUEL (Lichfield, Staffordshire 1709 – London 1784).** Poet and man of letters; compiler of *A Dictionary of the English Language*.
- JONES, ELIZA (b. Cut-off, Alaska 1938).** Teacher and knowledge-keeper; she completed and published a Koyukon dictionary in 2000 on the basis of materials by J. JETTÉ.
- JONES, WILLIAM (London 1746 – Calcutta, Bengal Presidency [now Kolkata, India] 1794).** Orientalist, judge, and pioneer of Indo-European comparative philology; with J. URI, he planned a revision of the *Thesaurus* of F. MENINSKI with added Persian material, and he made an English translation of the *Amarakośa* of AMARASIMHA which was used by H. COLEBROOKE in the preparation of his own translation.
- JUNGMANN, JOSEF (Hudlice, Bohemia [now in the Czech Republic] 1773 – Prague 1847).** Poet and linguist; author of an important Czech–German dictionary.
- JUNIUS, FRANCISCUS (Heidelberg, Electoral Palatinate [now in Germany] 1591 – Windsor, Berkshire 1677).** Classical and Germanic philologist; compiled the first printed dictionary of Biblical Gothic.
- JUNOD, HENRI-ALEXANDRE (Saint-Martin, Switzerland 1863 – Geneva 1934).** Missionary (Mission Suisse Romande) in Mozambique, ethnographer, and naturalist; compiler of a Ronga grammar with a Ronga–Portuguese–French–English wordlist.
- JU SI-KYEONG (Bongsan County, Korea 1876 – Seoul? 1914).** Grammarian; originator of the major Korean dictionary project of the twentieth century.
- JUŠKA, ANTANAS (ANTON VASIL’EVIČ JUŠKEVIČ) (Daujotai, Russian empire [now in Lithuania] 1819 – Kazan’, Russian empire 1880).** Catholic priest and musicologist; compiler of a Lithuanian dictionary based on the spoken language.
- KANKEYAR (from Mōrūr, Koṅku Nāṭu [now in India] 14th cent.?).** Compiler of the Tamil lexicographical text *Uric col nikaṇṭu*.
- KARADŽIĆ, VUK (Tršić, Ottoman empire [now in Serbia] 1787 – Vienna 1864).** Philologist and language reformer; author of an influential Serbian–German–Latin dictionary.
- KARŁOWICZ, JAN ALEKSANDER LUDWIK (Subortowicze, Poland [now Subartonys, Lithuania] 1836 – Warsaw 1903).** Linguist, ethnographer, and

musicologist; co-editor of the large monolingual Polish dictionary *Słownik języka polskiego* and compiler of a Polish dialect dictionary.

KASPI, JOSEF (JOSEPH IBN KASPI; BONAFOS DE L'ARGENTIERA) (Arles, County of Provence [now in France] c. 1280 – Majorca, Kingdom of Aragon [now in Spain] 1345). Logician and philosopher; compiler of a Hebrew dictionary.

KAYĀTARAR (KEYĀTARAR) (from Tēvai, near Rameshvaram [now in India] 15th×16th cent.). Compiler of the Tamil lexicographical text *Kayātaram*.

KELHAM, ROBERT (Billingborough, Lincolnshire 1717 – Edmonton, Middlesex 1808). Antiquary; compiler of an Anglo-Norman dictionary.

KENBŌ HIDETOSHI (Morioka, Japan 1914 – Tokyo 1992). Linguist and obsessive documenter of contemporary usage; actual editor of the dictionaries of contemporary Japanese *Meikai kokugo jiten* and *Sanseidō kokugo jiten*, the figurehead editor being KINDAICHI KYŌSUKE.

KENNEDY, VANS (Pinmore, Scotland 1784 – Bombay [now Mumbai, India] 1846). Soldier; compiler of a pioneering bidirectional dictionary of English and Marathi.

KENRICK, WILLIAM (Hertfordshire? 1729×1730 – London 1779). Miscellaneous writer; compiler of the first English pronouncing dictionary.

KHAL'FIN, SAGIT (from Kazan, Russian empire; 1732–1785). Teacher; compiled the first published dictionary of Tatar.

ḲHURĀSĀNĪ, YŪSUF BIN MUḤAMMAD YŪSUF (from western Afghanistan, probably d. in Agra, Mughal empire [now in India]; fl. early 16th cent.). Court physician; compiler of a medical vocabulary which provides Hindawī (northern Indian) equivalents for Arabic and Persian terms.

KIEFFER, JEAN DANIEL (Strasbourg, France 1767 – Paris 1833). Orientalist and secretary-interpreter of Oriental languages to the French government; coauthor with TH. BIANCHI of a Turkish–French dictionary.

KILIAAN, CORNELIS (VAN KIEL) (Duffel, Habsburg Netherlands [now in Belgium] c. 1530? – Antwerp, Spanish Netherlands [now in Belgium] 1607). Proof-corrector; compiler of a series of Dutch dictionaries.

KIMCHI, DAVID (RADAK) (Narbonne? [now in France] 1160 – Narbonne? 1235). Rabbi, exegete, and grammarian; compiler of a monolingual Hebrew dictionary which circulated in and after the medieval period, and is still sometimes used.

KINDAICHI KYŌSUKE (Morioka, Japan 1882 – Tokyo 1971). Linguist and scholar of Ainu literature; figurehead editor of the dictionaries of contemporary Japanese *Meikai kokugo jiten* and *Sanseidō kokugo jiten*, the actual editor being KENBŌ HIDETOSHI.

KINO (CHINI), EUSEBIO FRANCISCO (Segno, Bishopric of Trento [now in Italy] 1645 – Mission Santa Maria Magdalena, Spanish empire [now Magdalena de Kino, Mexico] 1711). Jesuit missionary in what are now parts of Mexico and the United States; compiler of a wordlist of O'odham.

- KIRCHER, ATHANASIUS** (Geisa, principality of Fulda [now in Germany] 1602 – Rome 1680). Jesuit and polymath; compiler of the first modern dictionary of Coptic; the first Chinese dictionary printed in Europe appeared as an appendix to the French translation of his *China illustrata* in 1670.
- KĪRTICANDRA** (from Tibet; later 13th×earlier 14th cent.). Teacher; translator into Tibetan, with GRAGS PA RGYAL MTSHAN, of the *Amarakośa* of AMARASIMHA and the commentary on it by SUBHŪTICANDRA.
- KIRWIN, WILLIAM JAMES** (Newport, Rhode Island 1925 – St John's, Canada 2016). Professor of English; editor (with G. STORY and J. WIDDOWSON) of the *Dictionary of Newfoundland English*.
- AL-KISĀ'Ī** (north of Baghdad c. 119/737 – near Rayy, Abbasid Caliphate [now in Iran] 189/805). Grammarian and one of the seven canonical readers of the Qur'ān; compiler of the first extant monograph on solecism in Arabic.
- KLAPPENBACH, RUTH** (Niedersteinbach, German empire [now in France] 1911 – Schöneiche, East Germany [now Germany] 1977). Founding editor, with W. STEINITZ, of the first post-World War II German dictionary, *Wörterbuch der deutschen Gegenwartssprache*.
- KLAUSNER, JOSEPH GEDALIAH** (Olkeniki, Russian empire [now in Lithuania] 1874 – Jerusalem 1958). Historian and professor of Hebrew literature; compiler of a Hebrew–Russian–German dictionary (1900, with Y. GUR).
- KLEINSCHMIDT, SAMUEL PETRUS** (Lichtenau, Kingdom of Denmark [now Alluitsoq, Greenland] 1814 – Neu-Herrnhut, Kingdom of Denmark [now part of Nuuk, Greenland] 1886). Moravian missionary in Greenland; compiler of a Greenlandic dictionary.
- KNAANI (né KAUFMANN), YAAKOV** (Hîncești, Russian empire [now in Moldova] 1894 – Jerusalem 1978). Teacher; compiler of *Otsar halashon haivrit litkufoteyha hashonot* ('Thesaurus of the Hebrew language according to its historical periods').
- KOELLE, SIGISMUND WILHELM** (Cleebronn, Württemberg [now in Germany] 1820 – London 1902). Missionary (Church Missionary Society) in Sierra Leone and Constantinople; his *Polyglotta Africana* was a pioneering collection of lexical data from African languages.
- KŌNG ĀNGUÓ 孔安國** (from China; 2nd or 1st cent. BC). Literary scholar; glosses from his commentaries on classical Chinese texts were drawn upon in the important early wordlist *Ĕryǎ*.
- KRAHL, GÜNTHER** (b. 1940; d. Leipzig, Germany 1992). University lecturer in Arabic; compiler, with G. GHARIEB, of an Arabic–German dictionary.
- KRAMER, MATTHIAS** (Cologne, Holy Roman empire [now in Germany] 1640 – Nuremberg, Holy Roman empire [now in Germany] 1729). Teacher of languages; compiler of major bilingual dictionaries of German.

- KRAMVANT, JAGANATHSHASTRI** (from Bombay [now Mumbai, India]; fl. 1829). Scholar and translator; leader of the pandits who produced the first monolingual printed dictionary of Marathi.
- KRAPE, JOHANN LUDWIG** (Derendingen, Württemberg [now in Germany] 1810 – Kornthal, German empire 1881). Missionary (Church Missionary Society) in east Africa; compiler of a dictionary of Oromo, of a *Vocabulary of Six East-African Languages*, including Swahili, and of the first Swahili dictionary.
- KROPF, JOHANN HEINRICH ALBERT** (Potsdam, Kingdom of Prussia [now in Germany] 1822 – Stutterheim, South Africa 1910). Lutheran missionary (Berlin Missionary Society) in southern Africa; compiler of a major Xhosa dictionary.
- KROPFITSCH, LORENZ** (b. Klagenfurt, Austria 1946). Compiler of dictionaries of German and Arabic.
- KṢĪRASVĀMIN** (from Kashmir [now in India or Pakistan]; first half of the 12th cent.). Author of a commentary in Sanskrit on the *Amarakośa* of AMARASIMHA.
- KŪKAI** 空海 (from Japan; 774–835). Buddhist monk; compiler of the Chinese dictionary *Tenrei banshō* (c. 830), which is the first extant dictionary from Japan.
- KUN DGA' RGYAL MTSHAN** (born Palden Dondup; known as Sakya Pandita) (Sakya, Tibet 1182 – Liangzhou, Tibet 1251). Spiritual leader and Buddhist scholar; compiler of an abridged metrical translation of select passages from the *Amarakośa* of AMARASIMHA.
- KURĀ' AL-NAML** (d. 310/922). Compiler of a collection of Arabic homonyms.
- KURMIN, JAN** (Viļkomir, Russian empire [now Ukmergė, Lithuania] 1795 – Kraslava, Russian empire [now Krāslava, Latvia] 1860). Compiler of a Polish–Latin–(High) Latvian dictionary.
- KURCHAT, ALEXANDER THEODOR** (Kretinga, Russian empire [now in Lithuania] 1857 – Kiefersfelden, Germany 1944). Teacher and linguist; editor of a Lithuanian–German dictionary which drew on the work of his uncle F. KURCHAT.
- KURCHAT, FRIEDRICH** (Noragehlen, Kingdom of Prussia [now in the Kaliningrad Oblast, Russia] 1806 – Cranz, German empire [now Zelenogradsk, Kaliningrad Oblast, Russia] 1884). Priest, professor, and linguist; compiler of German–Lithuanian and Lithuanian–German dictionaries, on which his nephew A. KURCHAT would draw.
- KÝ, PÉTRUS TRƯƠNG VĨNH** 張永記 (Vinh Long province, Vietnam 1837 – Saigon [now Ho Chi Minh City], Vietnam 1898). Interpreter, teacher, and writer; compiler of a Vietnamese–French dictionary which combined alphabetization with traditional subject ordering.
- LACOMBE, ALBERT** (Saint-Sulpice, Canada 1827 – Midnapore, Canada 1916). Oblate missionary; he compiled an early dictionary of Cree, and produced an edition of the Ojibwa dictionary of F. BARAGA.

- LAGADEUC, JEHAN (from Brittany; 15th cent.).** Priest; compiler of the Breton–French–Latin *Catholicon en troys langaiges*.
- LAGUNAS, JUAN BAPTISTA DE (from Castile, Spain; fl. 1530–1576).** Franciscan missionary in what is now Mexico; compiler of a Tarascan–Spanish dictionary.
- LANE, EDWARD WILLIAM (Hereford, Herefordshire 1801 – Worthing, Sussex 1876).** Translator from Arabic and writer on Egypt; compiler of an important Arabic–English dictionary based on the *Tāğ al-‘arūs* of AL-ZABĪDĪ, the latter part of which was seen through publication by Lane’s nephew S. LANE-POOLE.
- LANE-POOLE, STANLEY EDWARD (London 1854 – London 1931).** Numismatist and professor of Arabic; supervised the publication of the last three volumes of the Arabic dictionary of his uncle E. LANE.
- LANGE, JACOB (Königsberg, German empire [now Kaliningrad, Russia] 1711 – Riga, Russian empire [now in Latvia] 1777).** Theologian and linguist; compiler of a bidirectional dictionary of Latvian and German.
- LARA, LUIS FERNANDO (b. Mexico City, 1943).** Linguist; editor of an online dictionary of Mexican Spanish.
- LAROUSSE, PIERRE (Toucy, France 1817 – Paris 1875).** Teacher and publisher; founder of a tradition of encyclopedic French dictionaries.
- LARRAMENDI, MANUEL (Andoain, Spain 1690 – Azpeitia, Spain 1766).** Jesuit; compiler of the major Basque dictionary of the eighteenth century.
- LAS CASAS, CRISTÓBAL DE (from Seville, Spain; d. 1576).** Compiled a bidirectional Spanish and Italian dictionary.
- LAUGHLIN, ROBERT MOODY (b. Princeton, New Jersey, 1934).** Anthropologist and botanist; compiler and editor of dictionaries of Tzotzil.
- LAVES, GERHARDT KURT (Chicago, Illinois 1906 – Chicago 1993).** Linguist; he compiled wordlists in six Australian languages.
- AL-LAYṬ BIN AL-MUẒAFFAR (d. 190/805).** Philologist and disciple of AL-ḤALĪL; the *Kitāb al-‘Ayn* is sometimes ascribed to al-Layṭ rather than to al-Ḥalīl.
- LEGRAND DE LA LIRAYE, THÉOPHILE MARIE (Mauves, France 1819 – Saigon [now Ho Chi Minh City], Vietnam 1873).** Missionary (Missions étrangères de Paris) in Vietnam; compiler of the first Vietnamese–French dictionary and of a lexicon of Chinese characters glossed with their Sino-Vietnamese pronunciations.
- LE GUENNEC, GRÉGOIRE (Brittany, France? 1875 – Bimbe, Angola 1960).** Spiritan missionary in Angola; compiler of an Umbundu dictionary completed and published in 1972 by J. VALENTE.
- LĚNG YŪLÓNG 冷玉龙 (b. Rénshòu County, China 1954).** Editor, with Wéi Yīxīn 韦一心, of the huge monolingual Chinese dictionary *Zhōnghuá zìhǎi* (1994).

- LE PAGE, ROBERT BROCK** (London 1920 – Heslington, Yorkshire? 2006). Creolist; editor (with F. CASSIDY) of the *Dictionary of Jamaican English*.
- LE TALLEUR, GUILLAUME (d. Rouen, France 1494)**. Printer; compiler of a Latin–French dictionary.
- LE VER, FIRMIN** (Abbeville, France c. 1370 – Abbeville 1444). Carthusian prior; compiler of a Latin–French dictionary.
- LEVITA, ELIAHU (BAḤUR)** (Neustadt an der Aisch, Margraviate of Ansbach [now in Germany] 1468 – Venice [now in Italy] 1549). Poet and grammarian; compiler of lexica of Aramaic and Hebrew.
- LEVY, EMIL** (Hamburg, German Confederation 1855 – Freiburg im Breisgau, German empire 1917). Philologist; compiler of an eight-volume supplement to the Old Occitan dictionary of F. RAYNOUARD.
- LÍ JǐNXǐ 黎锦熙** (Xiāngtán, China 1890 – Běijīng 1978). Linguist; co-editor of the monolingual Chinese dictionary *Guóyǔ cídiǎn* and contributor to the *Xiàndài Hànyǔ cídiǎn*.
- Lǐ LùXīng 李禄兴** (b. Xīnjí City, China 1964). Editor of an English–Chinese dictionary.
- Lǐ RÓNG 李荣** (Wēnlíng, China 1920 – Běijīng 2002). Co-editor of the monolingual Chinese dictionary *Xiàndài Hànyǔ cídiǎn*.
- Lǐ SHÍ 李實** (from Suining, China; 1598–1674). Compiler of a dictionary of the topolect of Sichuān province.
- Lǐ Sī 李斯** (Hénán, China 284 BC – Shǎnxī, China 208 BC). Prime minister to the first emperor of China; compiler of one of the first Chinese lexicographical works, a manual of characters called *Cāngjié piān*.
- Lǐ Tǎo 李燾** (from Danleng, China; 1115–1184). Historian and official; editor of *Shuōwén jiězì wúyīn yùnpǔ*, a reordered recension of the dictionary of script *Shuōwén jiězì*.
- Lǐ YÁngBīng 李陽冰** (from Qiao commandery, China; fl. 8th cent.). Calligrapher and official; editor of a recension of the dictionary of script *Shuōwén jiězì*.
- LIANG SHIH-CHIU (LIÁNG SHÍQIŪ) 梁實秋** (Běijīng 1903 – Taipei 1987). Compiler of an English–Chinese dictionary.
- LIN YUTANG (LÍN YŮTÁNG) 林語堂** (Píng Hé County, China 1895 – Hong Kong 1976). Compiler of a Chinese–English dictionary.
- LINDE, SAMUEL BOGUMIŁ** (Toruń, Poland 1771 – Warsaw 1847). Educator and librarian; compiler of the first monolingual Polish dictionary.
- LISBOA, MARCOS DE** (from Lisbon; entered the Franciscan order 1582; d. Spain 1628). Franciscan missionary in the Philippines; compiler of the first extant wordlist of Bicol.

- LITTRÉ, ÉMILE MAXIMILIEN PAUL (Paris 1801 – Paris 1881).** Philosopher and man of letters; compiler of the foremost French dictionary of the nineteenth century.
- LIÚ XÌ 劉熙 (from China; c. 200 AD).** Compiler of a topical glossary which makes folk-etymological connections between Chinese words.
- LOYD, WILLIAM (Tilehurst, Berkshire 1627 – Hartlebury, Worcestershire 1717).** Natural philosopher and bishop; compiled an 'Alphabetical dictionary' as an appendix to John Wilkins' *Essay Towards a Real Character and a Philosophical Language* (1668), which gave equivalents for English words in the terms of an artificial language.
- LOBSCHIED, WILHELM (WILLIAM) (Gummersbach, Kingdom of Prussia [now in Germany] 1822 – Youngstown, Ohio 1893).** Missionary (Rheinische Missionsgesellschaft and Chinese Evangelization Society) in China; compiler of a Chinese dictionary.
- LODERECKER, PETER (from Bohemia; d. 1636).** Benedictine monk; editor of a revised and expanded edition of the multilingual dictionary of F. VRANCIĆ.
- LOMBARDO, NATALE (province of Calabria, Spanish empire [now in Italy] 1648 – Puebla, Spanish empire [now in Mexico] 1704).** Jesuit missionary in what is now Mexico; compiler of a dictionary of Opatá.
- LOMMATZSCH, ERHARD (Dresden, German empire 1886 – Frankfurt am Main, West Germany [now Germany] 1975).** Philologist; editor, on the basis of materials by his old teacher A. TOBLER, of an eleven-volume dictionary of Old French.
- LOPADIOTES, ANDREAS (from Constantinople [now Istanbul, Turkey]; fl. c. 1300–1330).** Man of letters and teacher; compiler of the Atticist Greek lexicon now called *Lexicon Vindobonense*.
- LOUGHRIDGE, ROBERT MCGILL (Laurensville, South Carolina 1809 – Waco, Texas 1900).** Presbyterian missionary to the Muskogee people in Oklahoma; compiler, with D. HODGE, of a dictionary of Muskogee.
- LOVE, JAMES ROBERT BEATTIE (Killeter, Ireland [now in Northern Ireland] 1889 – Adelaide, Australia 1947).** Presbyterian missionary in Australia; compiler of Worrorra wordlists.
- LOVELL, CHARLES J. (Fort McKinley, Maine 1907 – Willow Springs, Illinois 1960).** Writer, assistant (to M. MATHEWS) on *A Dictionary of Americanisms*, and original editor of the *Dictionary of Canadianisms on Historical Principles*.
- LÙ DIÀN 陸佃 (from Shanyin, China; 1042–1102).** Student of WÁNG ĀNSHÍ and official; compiler of a dictionary of terms of botanical and zoological interest in the classical literary text *Shījīng*.
- LÙ ĚRKUÍ 陸爾奎 (from Chángzhōu City, China; 1862–1935).** Compiler of the monolingual Chinese dictionary *Cíyuán*.

- LÙ FǎYÁN 陸法言 (from Linzhang, China; fl. 601). Chief compiler of the influential rhyme dictionary *Qièyùn*.
- LÙ GǔSŪN 陆谷孙 (Yúyáo City, China 1940 – Shànghǎi, China 2016). Editor-in-chief of *The English–Chinese Dictionary*.
- Lǚ SHŪXIĀNG 吕叔湘 (Dānyáng, China 1904 – Běijīng 1998). Linguist; co-editor of the monolingual Chinese dictionary *Xiàndài Hànyǔ cídiǎn*.
- LUDOLF, HIOB (Erfurt, Holy Roman empire [now in Germany] 1624 – Frankfurt am Main, Holy Roman empire [now in Germany] 1704). Diplomat and founder of Ethiopic studies in Europe; compiler of a Ge'ez–Latin dictionary.
- LUÓ ZHÚFĒNG 罗竹风 (Píngdù City, China 1911 – Shànghǎi, China 1996). Editor of an eleven-volume monolingual Chinese dictionary.
- MAALER, JOSUA (Zürich, Swiss Confederacy 1529 – Glattfelden, Swiss Confederacy 1599). Minister of religion; compiler, on the basis of the work of P. CHOLINUS and J. FRISIUS, of the first free-standing German–Latin dictionary.
- MACAFEE, CAROLINE (b. Lanark, Scotland 1956). University lecturer; editor of *The Concise Ulster Dictionary*.
- MACGOWAN, JOHN (Belfast, United Kingdom 1835 – London 1922). Missionary (London Missionary Society) in China; compiler of a Chinese dictionary.
- MACHADO, JOSÉ PEDRO (Faro, Portugal 1914 – Lisbon 2005). Linguist; compiler of etymological dictionaries of Portuguese.
- MACHONI DE CERDEÑA, ANTONIO (Cagliari, Kingdom of Sardinia [now in Italy] 1671 – Córdoba, Spain 1753). Jesuit missionary; author of a grammar and wordlist of Lule.
- MACIVER, DONALD (Plymouth, Devon 1852 – Guǎngzhōu, China 1910). Missionary in China; compiler of *An English–Chinese Dictionary in the Vernacular of the Hakka People in the Canton Province*.
- MACLAY, ROBERT SAMUEL (Concord, Pennsylvania 1824 – Los Angeles, California 1907). Methodist missionary in China and Japan; compiler, with C. BALDWIN, of *An Alphabetic Dictionary of the Chinese Language in the Foochow Dialect*.
- MAĆZYSKI, JAN (Gzików, Poland 1527 – Miłkowice, Duchy of Legnica [now in Poland] c. 1587). Humanist; compiler of a Latin–Polish lexicon.
- MADAN, ARTHUR CORNWALLIS (Cam, Gloucestershire 1846 – Stourton, Wiltshire 1917). Missionary (Universities Mission to Central Africa) in east Africa; compiler of English–Swahili and Swahili–English dictionaries.
- MAGNÚS ÓLAFSSON (Eyafjörður region, Iceland 1573 – Laufás, Iceland 1636). Minister of religion; compiler of the first Icelandic dictionary, the Old Icelandic–Latin *Specimen lexicæ runicæ*.

- MAHEŚVARA (from India; early 12th cent.).** Compiler of a homonymic Sanskrit dictionary.
- MAḤMŪD AL-KĀŠĠARĪ (from the Karakhanid Khanate [now in the approximate area of the Kashgar Prefecture, Xinjiang, China]; c. 1008 – c. 1105).** Traveller and collector of information; compiler of the dictionary of Turkic languages *Dīwān Luġāt at-Turk*.
- AL-MĀLAQĪ (d. 702/1302).** Compiler of an alphabetical wordlist of Arabic particles.
- MALARET, AUGUSTO NICOLÁS (Sabana Grande, Spanish empire [now in Puerto Rico] 1878 – San Juan, Puerto Rico 1967).** Lawyer and linguist; compiler of a dictionary of Americanisms in Spanish.
- MALIGE-KLAPPENBACH, HELENE (from Germany; c. 1907–1996).** Contributor to the *Wörterbuch der deutschen Gegenwartssprache* of which her sister R. KLAPPENBACH was a founding editor.
- AL-MALIK AL-AFDAL (from Yemen; reigned 764–778/1363–1377).** Sixth Rasulid king of Yemen; compiler or patron of the Arabic–Turkish–Persian–Mongolian–Greek–Armenian wordlist called the Rasulid Hexaglot.
- MALLINĀTHA (from Āndhra [now in India]; 14th cent.).** Author of a commentary in Sanskrit on the *Amarakośa* of AMARASIMHA.
- MAMIANI, LUIS VINCENCIO (Pesaro, Papal States 1652 – Rome 1730).** Jesuit missionary in Brazil, and author of a Kiriri grammar and catechism; his *Vocabulario kiriri*, a rare example of a missionary dictionary of a Brazilian language other than Tupi, is lost.
- MANCELIUS, GEORG (Grenzhof, Duchy of Courland and Semigallia [now Augstkalne, Latvia] 1593 – Mitau, Duchy of Courland and Semigallia [now Jelgava, Latvia] 1654).** Lutheran theologian; compiler of the first Latvian dictionary.
- MAṆKHA (from Kashmir [now in India or Pakistan]; early 12th cent.?).** Perhaps to be identified with the Kashmiri poet and politician of that name; compiler of the homonymic Sanskrit dictionary called the *Maṅkhakośa* or *Anekārthakośa*.
- MÁO QÍLÍNG 毛奇齡 (from Xiaoshan, China; 1629–1713).** Prominent literary and intellectual figure; compiler of a dictionary of the Cantonese topolect.
- MARBÁN, PEDRO (Lerida [now Lleida], Spain 1647 – Loreto, Spanish empire [now in Bolivia] 1713).** Jesuit missionary; compiler of a grammar and dictionary of Moxo.
- MARTIN, BENJAMIN (baptized Worplesdon, Surrey 1705; d. London 1782).** Popularizer of science; compiler of the English dictionary *Lingua Britannica reformata*.
- MATHEWS, MITFORD M. (Jackson, Alabama 1891 – Chicago, Illinois 1985).** Assistant editor (under W. CRAIGIE) of the *Dictionary of American English* and editor of *A Dictionary of Americanisms*.

- MATHEWS, ROBERT HAMILTON** (Narellan, New South Wales Colony [now part of Sydney, Australia] 1841 – Parramatta, Australia 1918). Surveyor and anthropologist; he made wordlists of numerous languages of south-eastern Australia.
- MATHEWS, ROBERT HENRY** (Flemington, Australia 1877 – Melbourne, Australia 1970). Missionary (China Inland Mission) in China; compiler of a Chinese–English dictionary.
- MATSUI KANJI** (Chōshi, Japan 1863 – Ashio, Japan 1945). Linguist; editor of the Japanese dictionary *Dai-Nihon kokugo jiten*, on which was based the *Nihon Kokugo daijiten* of his grandson MATSUI SHIGEKAZU.
- MATSUI SHIGEKAZU** (b. Tokyo 1926). Linguist; editor of the large historically oriented Japanese dictionary *Nihon Kokugo daijiten*, based on the *Dai-Nihon kokugo jiten* of his grandfather MATSUI KANJI.
- MATSUMURA AKIRA** (Tokyo 1916 – Higashi-Matsuyama, Japan 2001). Historical linguist; editor of the monolingual Japanese dictionaries *Dajirin* and *Daijisen*.
- MATTA, JOAQUIM DIAS CORDEIRO DA** (Cabiri, Angola 1857 – Barra do Cuanza [now Barra do Kwanza], Angola 1894). Poet and journalist; compiler of the first dictionary of Kimbundu by an Angolan.
- MAWER, ALLEN** (London 1879 – Broxbourne, Hertfordshire 1942). University administrator and place-name scholar; compiler (often in collaboration with F. STENTON) of several place-name dictionaries.
- AL-MAYDĀNĪ, ABU’L-FADL** (d. Nishapur, Khorasan [now in Iran] 518/1124). Philologist; compiler of the most famous collection of Arabic proverbs and of an Arabic–Persian wordlist.
- McKELSON, KEVIN** (Moonee Ponds, Australia 1926 – Melbourne, Australia 2011). Pallottine missionary in Australia; compiler of dictionaries of Northern Nyangumarta and other Australian languages.
- MEDHURST, WALTER HENRY** (London 1796 – London 1857). Missionary (London Missionary Society) in China; compiler of Chinese–English dictionaries.
- MEDINĪKARA** (from India; fl. 1200×1275). Compiler of the *Nānārthaśabdakoṣa* or *Medinīkoṣa*, a homonymic lexicon of Sanskrit.
- MEGISER, HIERONYMUS** (Stuttgart, Württemberg [now in Germany] c. 1554×1555 – Linz, Austria 1619). Historian and philologist; compiler of a German–Latin–Slovenian–Italian dictionary and the all-inclusive polyglot *Thesaurus polyglottus vel dictionary multilingue*.
- MEISTERBURG, ANTON** (Bernkastel, Archbishopric of Trier [now in Germany] 1719 – Trier 1799). Jesuit missionary in Brazil; probable compiler of a Tupi dictionary.
- MÉI YĪNGZUÒ 梅膺祚** (from Xuancheng, China; early 17th cent.). Compiler of the innovative and influential dictionary of script *Zihui*.

- Menahem ben Saruq** (Turtūshah, al-Andalus [now Tortosa, Spain] c. 920 – Qurtuba?, al-Andalus [now Córdoba, Spain] c. 970). Court poet; compiler of a pioneering monolingual Hebrew dictionary.
- MENINSKI, FRANÇOIS À MESGNIEN** (FRANCISCUS MESGNIEN LOTHARINGUS) (Totainville, Lorraine [now in France] 1620×1623 – Vienna 1698). Interpreter and diplomat; his *Thesaurus linguarum orientalium Turcicae–Arabicae–Persicae* was the major printed Turkish dictionary of the seventeenth and eighteenth centuries.
- MÉNTRIDA, ALONSO DE** (Méntrida, Spain 1559 – Manila 1637). Augustinian missionary in the Philippines; compiler of the first extant wordlist of Visayan.
- MEYER, HEINRICH AUGUST EDUARD** (Berlin 1813 – Bethanien [now Bethany]?, South Australia 1862). Lutheran missionary in Australia; compiler of a dictionary of Ramindjeri.
- MEYER-LÜBKE, WILHELM** (Dübendorf, Switzerland 1861 – Bonn, Germany 1936). Philologist; compiler of an etymological dictionary of the Romance languages.
- MIELCKE, CHRISTIAN GOTTLIEB** (Georgenburg, Kingdom of Prussia [now Maėvka, Kaliningrad Oblast, Russia] 1733 – Pillkallen, Kingdom of Prussia [now Dobrovol'sk, Kaliningrad Oblast, Russia] 1807). Protestant cantor, philologist, and translator; compiler of a Lithuanian dictionary partly based on the one by P. RUHIG.
- MIKALJA, JAKOV** (GIACOMO MICAGLIA, JACOBUS MICALIA) (Peschici, Kingdom of Naples [now in Italy], 1601 – Loreto, Papal States [now in Italy] 1654). Jesuit priest; compiler of a Croatian–Italian–Latin dictionary.
- MIKLOŠIČ, FRANC** (FRANZ VON MIKLOSICH) (Pichelberg bei Luttenberg, Austrian empire [now Radomerščak pri Ljutomeru, Slovenia], 1831 – Vienna 1891). Philologist; editor of a monumental Church Slavic dictionary (to which contributions were made by M. PLETERŠNIK).
- MINSHEU, JOHN** (b. 1559×1560; d. London 1627). Compiler of a Spanish–English dictionary and of a large polyglot dictionary with English headwords.
- MĪRZĀ KHĀN** (from the Mughal empire; late 17th cent.). Courtier and man of letters; compiler of a wordlist in Persian defining Indic terms.
- MISCHLICH, ADAM** (Nauheim, Hesse-Kassel [now in Germany] 1864 – Frankfurt am Main, Germany 1948). Missionary (Basel Mission) in Togo, and Africanist; compiler of a Hausa dictionary.
- MISTRAL, FRÉDÉRIC** (Maillane, France 1830 – Maillane 1914). Poet; his dictionary *Trésor dou Félibrige* was meant to establish a standard for the Occitan language.
- MOERIS** (from the Roman empire; 3rd cent.?). Compiler of an alphabetically arranged Atticist Greek lexicon.

- MOHAMMAD NAKHJAWĀNĪ (from Persia; 14th cent.).** Compiler of the Persian dictionary *Ṣiḥāḥ al-furs*.
- MOLESWORTH, JAMES THOMAS (Camberwell, Surrey 1795 – Clifton, Bristol 1872).** Soldier and scholar; compiler, with G. CANDY and T. CANDY, of dictionaries of Marathi.
- MOLINA, ALONSO DE (province of Cáceres, Spain c. 1514 – Mexico City 1579×1585).** Franciscan missionary in what is now Mexico; compiler of the first printed dictionary of Nahuatl.
- MOLINER, MARIA (Paniza, Spain 1900 – Madrid 1981).** Librarian; compiler of an important dictionary of Spanish.
- MOLINO, GIOVANNI (YOVHANES ANKURAC'Ī) (b. 1592).** Dragoman and printer of Armenian books; compiler of an Italian–Turkish dictionary.
- MONET, PHILIBERT (Bonneville, Duchy of Savoy [now in France] 1566 – Lyon, France 1643).** Jesuit and teacher; compiler of a large French–Latin dictionary.
- MONTGOMERY, MICHAEL B. (b. Knoxville, Tennessee 1950).** Professor of English and linguistics; editor of *The Dictionary of Smoky Mountain English*.
- MOON SEYOUNG (Seoul 1888 – missing during the Korean War, 1950×1953).** Teacher and lexicographer; compiler of *Joseon-eo Sajeon*.
- MOORE, GEORGE FLETCHER (Donemana, Ireland [now in Northern Ireland] 1798 – London 1886).** Landowner in Australia and public servant; he made a substantial early dictionary of Noongar.
- MORAIS SILVA, ANTÔNIO DE (Rio de Janeiro, Brazil 1755 – Pernambuco, Brazil 1824).** Lawyer and translator; compiler of a dictionary of Portuguese which played a role like that of an academy dictionary in the nineteenth century.
- MORICE, ADRIEN-GABRIEL (Saint-Mars-sur-Colmont, France 1859 – St Boniface, Canada 1938).** Oblate missionary in Canada; compiler of a monumental dictionary of Carrier.
- MOROHASHI TETSUJI (Niigata, Japan 1883 – Tokyo 1982).** Scholar of Chinese classics; editor of the large scholarly sinograph–Japanese dictionary *Dai kanwa jiten*.
- MORRISON, ROBERT (Morpeth, Northumberland 1782 – Guǎngzhōu, China 1842).** Missionary (London Missionary Society) in China; compiler of a Chinese dictionary.
- MORRISON, WILLIAM THOMAS (New York? 1834×1835 – Běijīng 1869).** Missionary in China; compiler of *An Anglo–Chinese Vocabulary of the Ningpo Dialect*.
- MOSER, EDWARD W. (Joliet, Illinois 1924 – El Desemboque, Mexico 1976).** Linguist with the Summer Institute of Linguistics; compiler, with his wife M. MOSER, of a dictionary of Seri.
- MOSER, MARY MARGARET BECK (Lock Haven, Pennsylvania 1924 – Catalina, Arizona? 2013).** Linguist with the Summer Institute of Linguistics; compiler, with her husband E. MOSER, and with S. Marlett, of dictionaries of Seri.

- MOTH, MATTHIAS** (Odense, Denmark 1649 – Copenhagen? 1719). Civil servant; compiler of a huge unpublished dictionary of Danish.
- MTUZE, PETER TSHOBISA** (b. Middelburg, South Africa 1941). Poet and clergyman; formerly editor-in-chief of the *Greater Dictionary of Xhosa*.
- MU'ARRIĞ AL-SADŪSĪ** (d. 193×198/808×813). Compiler of the Arabic proverb collection *al-Amṭāl*.
- AL-MUFADDAL AL-ḌABBĪ** (d. Kufa, Abbasid Caliphate [now in Iraq] 164×170/781×787?). Compiler of the Arabic proverb collection *Amṭāl al-'Arab*.
- MUḤAMED (BOSNEVĪ) HEVĀĪ USKUFĪ** (from Bosnia, Ottoman empire; 17th cent.). Compiler of a Bosnian–Turkish dictionary.
- MUḤAMMAD ḤUSAYN TABRĪZĪ** (known as BURHĀN) (from Hyderabad [now in India]; 17th cent.). Compiler of the influential Persian lexicon *Burhān-i Qāṭi*.
- MÜHLENBACH, KARL (KÄRLIS MĪLENBAHS)** (Kandava, Russian empire [now in Latvia] 1853 – Võru, Russian empire [now in Estonia] 1916). Teacher and linguist; initiator of the Latvian–German dictionary that was finished by J. ENDZELĪNS.
- MU'ĪN, MUḤAMMAD** (Rasht, Iran c. 1914 – Tehran 1971). Director of the foundation set up by 'A. DIHKHUDĀ to manage the Persian dictionary project which Dihkhudā had initiated; he also compiled a six-volume encyclopedic dictionary of Persian, and edited the seventeenth-century *Burhān-i Qāṭi*.
- MUKA, ARNOŠT (ERNST MUCKE)** (Großhänchen, German empire 1854 – Bautzen, Germany 1932). Teacher and linguist; among his writings on the Sorbian language is a comprehensive Lower Sorbian–German dictionary.
- MUKHLIŠ, ĀNAND RĀM** (b. Sodhra, Mughal empire [now in Pakistan]; 1699–1750). Bureaucrat and poet; compiler of a Persian dictionary.
- MÚLJI, KARSANDÁS** (Gujarat [now in India] 1832 – Kathiawar Agency [now in India] 1875). Teacher, journalist, and social reformer; compiler of *A Pocket Dictionary, Gujarati and English*.
- MULLAWILLABURKA** (South Australia c. 1811 – Adelaide, South Australia 1845). Leader among the Kaurna people; he was one of the Kaurna on whose teaching the dictionary of C. TEICHELMANN and C. SCHÜRMANN was based.
- AL-MURĀDĪ** (d. 749/1348). Compiler of a wordlist of Arabic particles arranged in order of their length.
- MURKO, ANTON** (Spodnja Voličina, Austrian empire [now in Slovenia] 1809 – Spodnje Hoče, Austria-Hungary [now in Slovenia] 1871). Priest and philologist; compiler of a bidirectional dictionary of Slovene and German.
- MURMELLIUS, JOANNES** (Roermond, Duchy of Guelders [now in the Netherlands] 1480 – Deventer, Bishopric of Utrecht [now in the Netherlands] 1517). Humanist; compiler of an influential Latin dictionary for schoolchildren.

- MURRAY, JAMES AUGUSTUS HENRY** (**Denholm, Scotland 1837 – Oxford, Oxfordshire 1915**). Philologist; editor of the first published fascicle of the *Oxford English Dictionary*, and (with H. BRADLEY, W. CRAIGIE, and C. ONIONS) of much of the remainder of the dictionary.
- MYMER, FRANCISZEK** (**b. Lwówek Śląski, Poland c. 1500; d. after 1564**). Translator, and poet; editor of a Latin–German–Polish dictionary after a Czech model.
- NÁGERA YANUAS, DIEGO DE** (**Mexico City 1570? – Jocotitlan?, Spanish empire [now in Mexico] 1635**). Priest; compiler of the first grammar and wordlist of Mazahua.
- NAKHAT** (**NIYĀZ ‘ALĪ BEG**) (**from Delhi; d. 1849×1850**). Compiler of an early monolingual Urdu dictionary.
- NASCIMENTO, JOSÉ PEREIRA DO** (**Ceará, Brazil 1861 – Lisbon 1913**). Medical officer in the Portuguese Royal Navy; compiler of Umbundu wordlists.
- NATHAN BEN YEḤIEL** (**Rome 1035 – Rome 1110**). Rabbi; compiler of *‘Arukh*, a dictionary of post-Biblical Hebrew, which influenced a long succession of subsequent dictionaries including work by J. BUXTORF THE ELDER.
- NEBRIJA, ELIO ANTONIO DE** (**AEIUS ANTONIUS NEBRISSENSIS**) (**Lebrija, Castile [now in Spain] 1441 – Alcalá, Spain 1522**). Humanist; compiler of Latin–Spanish and Spanish–Latin dictionaries which were used as models for dictionaries across the Spanish empire, notably those of A. DE MOLINA and his successors in Mesoamerica.
- NEIRA, ALONSO DE** (**Matapozuelos, Spain c. 1635 – Camoa, Mexico 1703**). Jesuit missionary in South America; compiler of a grammar and dictionary of Achagua completed by J. RIVERO.
- NEKES, HERMANN** (**Essen, German empire 1875 – Kew [now part of Melbourne], Australia 1948**). Pallottine missionary in Cameroon and Australia; compiler, with E. WORMS, of the large polyglot lexical compendium *Australian Languages*.
- NESSELMANN, GEORG HEINRICH FERDINAND** (**Fürstenau, West Prussia [now Kmiecin, Poland] 1811 – Königsberg, German empire [now Kaliningrad, Russia] 1881**). Orientalist and philologist; compiler of a Lithuanian–German dictionary.
- NGAG DBANG ‘JIG RTEN DBANG PHYUG GRAGS PA’I RDO RJE** (**from Tibet; 16th cent.**). Compiler of the Tibetan dictionary *Mkhas pa’i rna rgyan*.
- NGÔ THÌ NHẬM** 吳時壬 (**from Đại Việt [now Vietnam]; 1746–1803**). Scholar, high official, and diplomat; the Chinese–Vietnamese rhyming pedagogical dictionary *Tam thiên tự* has been attributed to him.
- NICOT, JEAN** (**Nîmes, France 1530 – Paris 1604**). Diplomat; compiler of a French dictionary with Latin as metalanguage.

- NIEDERMANN, MAX** (Winterthur, Switzerland 1874 – Neuchâtel, Switzerland 1954). Classicist and Indo-Europeanist; one of the original editors (with A. SENN) of the bilingual *Wörterbuch der litauischen Schriftsprache*.
- NILAMBIKAI AMMAIYAR, T.** (1903–1945×1948). Tamil language purist and feminist; the first female lexicographer of Tamil, she compiled a Sanskrit–Tamil dictionary intended to help Tamilians to avoid Sanskrit terms.
- NIŞANYAN, SEVAN** (b. Istanbul 1956). Turkish-Armenian intellectual, essayist, and linguist; compiler of the historical-etymological Turkish dictionary *Sözlerin Soyağacı*.
- NIZOLIUS, MARIUS** (MARIO NIZZOLI) (Boretto, near Brescello, Duchy of Ferrara [now in Italy] 1488 – Brescello? 1566). Compiler of a dictionary of Ciceronian Latin.
- NLEMVO** (MANTANTU DUNDULU) (Padwa, Angola c. 1865 – Ngombe Lutete, Bas-Congo [now in the Democratic Republic of the Congo], 1938). Translator and author; assisted W. BENTLEY in the compilation of a dictionary of Kikongo.
- NOCEDA, JUAN JOSÉ DE** (Seville, Spain 1681 – Manila? 1747). Jesuit missionary in the Philippines; compiler, with P. DE SAN LÚCAR, of a Tagalog dictionary.
- NOGUEIRA, RODRIGO DE SÁ** (Mindelo, Cape Verde 1892 – Lisbon 1979). Linguist and phoneticist; compiler of a Ronga dictionary.
- NONIUS MARCELLUS** (perhaps from Thubursicum, Roman empire [now Khemissa, Algeria]; 4th cent. AD?). Compiler of an encyclopedic Latin dictionary.
- NOWELL, LAURENCE** (b. Whalley, Lancashire 1530; disappeared in continental Europe in or after 1569). Antiquary and cartographer; compiler of the first scholarly wordlist of Old English.
- NXUMALO, OTTY EZROM HOWARD MANDLAKAYISE** (b. Louwsburg, South Africa 1938). Writer and educator; compiler with S. NYEMBEZI of the first monolingual Zulu dictionary.
- NYEMBEZI, (CYRIL LINCOLN) SIBUSISO** (Babanango, South Africa 1919 – Pietermaritzburg, South Africa 2000). Writer and educator; compiler with O. NXUMALO of the first monolingual Zulu dictionary.
- Ó CLÉIRIGH, MICHÉL** (b. Kilbarron, Ireland in or after 1590; d. Leuven, Spanish Netherlands [now in Belgium] 1643?). Franciscan lay brother, scribe, and chronicler; compiler of the first printed monolingual Irish dictionary.
- OLMOS, ANDRÉS DE** (Oña, Spain c. 1485 – Tampico, Spanish empire [now in Mexico] 1571?). Franciscan missionary in what is now Mexico; compiler of the first grammar of Nahuatl, which includes a Nahuatl wordlist.
- Ó MUIRITHE, DIARMAID** (New Ross, Ireland 1936 – Vienna 2014). Professor of Irish studies and newspaper columnist; compiler of *A Dictionary of Anglo-Irish*.

- ONIONS, CHARLES TALBUT** (Birmingham, Warwickshire 1873 – Oxford, Oxfordshire 1965). Philologist: one of the four editors (with J. MURRAY, H. BRADLEY, and W. CRAIGIE) of the first edition of the *Oxford English Dictionary*.
- ORION OF THEBES** (born in Thebes, Eastern Roman empire [now in Egypt]; fl. 5th cent.). Grammarian; compiler of a pioneering Greek etymologicon.
- ŌTA ZENSAI** (Fukuyama, Japan 1759 – Edo [now Tokyo] 1829). Scholar of Chinese classics interested in Japanese vernacular; compiler of a dictionary of non-literary Japanese.
- ŌTSUKI FUMIHIKO** (Edo [now Tokyo] 1847 – Tokyo 1928). Western studies scholar and linguist; compiler of *Genkai*, ‘the first modern Japanese dictionary’.
- ODIN, CÉSAR** (Bassigny, France c. 1560 – Paris 1625). Translator and writer on modern languages; compiled a bidirectional Spanish and French dictionary.
- OŽEGOV, SERGEJ IVANOVIČ** (Kamennoe, Russian empire [now Kuvšinovo, Russia] 1900 – Moscow 1964). Linguist; editor of a monolingual Russian dictionary based on that of D. UŠAKOV.
- PAHL, HERBERT WALTER** (Fort Murray, Cape Colony [now in South Africa] 1907 – Port Elizabeth, South Africa 1990). Educator; chief editor of the first published volume of the *Greater Dictionary of Xhosa*.
- PALSgrave, JOHN** (b. London, end of the 15th cent.; d. 1554). Courtier and teacher; compiler of a guide to the French language for English learners which included the first substantial printed French–English wordlists.
- PAMPHILUS OF ALEXANDRIA** (from the Roman empire; fl. 1st cent. AD). Scholar; compiler of one of the first universal Greek lexica (now lost), which, in its abridgement by **DIOGENIANUS**, was a source for the work of **HESYCHIUS**.
- PĀṆINI** (from northern India; c. 6th cent. BC?). The pre-eminent grammarian in the Sanskrit tradition; his grammatical work is transmitted with two Sanskrit wordlists, of which the interpretamenta are probably by a later hand.
- PAPIAS** (from Lombardy [now in Italy]; 11th cent.). Grammarian; compiler of the monolingual Latin *Elementarium*.
- PASOR, GEORG** (Ellar, county of Sayn [now in Germany] 1570 – Franeker, Netherlands 1637). Teacher; compiler of the first modern dictionary of New Testament Greek.
- PASSOW, FRANZ LUDWIG CARL FRIEDRICH** (Ludwigslust, Mecklenburg-Schwerin [now in Germany] 1786 – Breslau, Prussia [now Wrocław, Poland] 1833). Classicist; author of an important early statement of the principles of historical lexicography, which he put into practice in successive editions of a Greek–German dictionary.

- PAUL THE DEACON (Duchy of Friuli (now in Italy) c. 720 – Monte Cassino [now in Italy] 799).** Historian and poet; made an epitome of the abridgement of the *De verborum significatu* of VERRIUS FLACCUS by POMPEIUS FESTUS.
- PAUL, HERMANN OTTO THEODOR (Magdeburg, Saxony [now in Germany] 1846 – Munich, Germany 1921).** Philologist; compiler of an innovative German dictionary.
- PAULINUS A SANCTO BARTHOLOMAEO (JOHANN PHILIPP WESDIN) (Hof am Leithaberge, Austria 1746 – Rome 1806).** Carmelite missionary in India and Orientalist; editor of a printed edition of the first part of the *Amarakośa* of AMARASIMHA.
- PEACOCK, EDWARD (Hemsworth, Yorkshire 1831 – Kirton-in-Lindsey, Lincolnshire 1915).** Antiquary and contributor to the *Oxford English Dictionary*; compiler of a Lincolnshire dialect wordlist.
- PECIAR, ŠTEFAN (Nedanovce, Austro-Hungarian empire [now in Slovakia] 1912 – Bratislava, Czechoslovakia [now in Slovakia] 1989).** Linguist; editor of the first authoritative dictionary of the Slovak language.
- PELLAS, SAUVEUR ANDRÉ (b. Comps-sur-Artuby, France 1667; d. 1727).** Minim friar; compiler of the first free-standing dictionary of Occitan.
- PELLEPRAT, PIERRE-IGNACE (Bordeaux, France 1606 – Puebla, Mexico 1667).** Jesuit missionary in the Caribbean and South America; compiler of an early wordlist of Carib.
- PEROTTI, NICCOLÒ (Sassoferrato [now in Italy] 1429×1430 – Sassoferrato 1480).** Humanist and archbishop of Siponto; compiler of *Cornucopiae*, a major study of the vocabulary of classical Latin structured as a lexical commentary on the poems of Martial.
- PERRIN, JAMES (Chichester, Sussex 1801 – Durban, Colony of Natal [now in South Africa] 1888).** Missionary and naturalist; compiler of the first significant Zulu dictionary.
- PERRY, WILLIAM (from Edinburgh; fl. 1774–1801).** Teacher and naval surgeon; compiler of one of the first pronouncing dictionaries of English.
- PETERS, WILHELM KARL HARTWIG (Koldenbüttel, Duchy of Schleswig [now in Germany] 1815 – Berlin 1883).** Explorer and naturalist; his manuscript wordlists were the basis for *The Languages of Mosambique* by W. BLEEK.
- PETITNICOLAS, MICHEL ALEXANDRE (Coinches, France 1828 – near Seoul 1866).** Missionary (Missions étrangères de Paris) in Korea; a Latin–Korean dictionary is attributed to him or to M.-N.-A. DAVELUY.
- PETITOT, ÉMILE-FORTUNÉ-STANISLAS-JOSEPH (Grancey-le-Château, France 1838 – Mareuil-lès-Meaux, France 1916).** Oblate missionary in Canada; compiler of dictionaries of Athabaskan languages and of Western Canadian Inuktitut.

- PETROCCHI, POLICARPO** (Castello di Cireglio, Austrian empire [now in Italy] 1852 – Castello di Cireglio 1903). Writer and literary scholar; compiler of dictionaries of Italian.
- PETTER, RODOLPHE CHARLES** (Vevey, Switzerland 1865 – Lame Deer, Montana 1947). Mennonite missionary to the Cheyenne people; compiler of a Cheyenne dictionary.
- PFUHL, CHRISTIAN TRAUOGOTT** (Křesčan Bohuvěř Pful) (Prauschwitz, Kingdom of Saxony [now in Germany] 1825 – Pirna, German empire [now in Germany] 1889). Linguist and teacher; compiler of an Upper Sorbian–German dictionary.
- PHAM DINH HỒ** 範廷琥 (Đan Loan, Đại Việt [now Vietnam] 1768 – Vietnam 1839). Writer and teacher; compiler of a Chinese–Vietnamese dictionary.
- PHILITAS THE COAN** (from Cos; b. c. 340 BC). Poet and scholar; compiler of the *Ataktoi glōssai*, ‘Rare words with no arrangement’, a list, extant only in fragments, of rare words in Ancient Greek, with no explanatory material provided.
- PHILLIPS, EDWARD** (b. 1630; d. in or after 1696). Miscellaneous writer; his *New World of English Words* was the first monolingual English dictionary to be published in folio format.
- PHOTIUS** (from the Byzantine empire; c. 810 – after 893). Patriarch of Constantinople and man of letters; compiler in his youth (c. 830–840) of a Greek lexicon which drew on the work of **DIOGENIANUS** and other sources.
- PHRYNICHUS ‘THE ARAB’** (from Bithynia, Roman empire [now in Turkey]; 2nd cent. AD). Rhetorician; compiler of two prescriptive wordlists of ancient Greek extant only in later abridgements (one of them by **PHOTIUS**).
- PIANZOLA, BERNARDINO** (Domodossola, Duchy of Milan [now in Italy] 1721 – Padua, Austrian empire [now in Italy] 1803). Franciscan friar and missionary; compiler of Turkish wordlists.
- PICHARDO Y TAPIA, ESTEBAN** (Santiago de los Caballeros, Spanish empire [now in the Dominican Republic] 1799 – Havana, Spanish empire [now in Cuba] 1879). Lawyer and geographer; compiler of a dictionary of Cubanisms.
- PIEKARSKI, EDWARD** (EDUARD KARLOVIČ PEKARSKIJ) (Piatrovičy [now in Belarus] 1858 – St Petersburg 1934). Ethnographer; compiler of the first Yakut dictionary.
- PIGNEAU DE BEHAINE, JOSEPH GEORGES PIERRE** (BÁ ĐA LỘC) (Origny-en-Thiérache, France 1741 – Quy Nhơn, Đại Việt [now Vietnam] 1799). Missionary (Missions étrangères de Paris) in Vietnam, titular bishop of Adran, and friend and adviser to the first Nguyễn emperor; compiler of Vietnamese dictionaries, which were eventually edited and published by J. -L. TABERD.

- PINET, PIERRE-FRANÇOIS** (Périgueux, France 1660 – Kaskaskia, New France [now in Illinois] 1702). Jesuit missionary in what is now the United States; compiler of the first dictionary of Miami-Illinois.
- PINNOCK, OSBORN** (from Gloucester, Gloucestershire; fl. c. 1148). Benedictine monk and theologian; compiled the derivationally structured *Panormia* or *Liber derivationum*.
- PLACIDUS** (from the Roman empire or a post-Roman territory; 5th or 6th cent.). Compiler of a Latin glossary used by ISIDORE and others.
- PLATTS, JOHN THOMPSON** (Calcutta, Bengal Presidency [now Kolkata, India] 1830 – London 1904). Inspector of schools in India and lecturer in Persian at Oxford; compiler of *A Dictionary of Urdu, Classical Hindi, and English*.
- PLETERŠNIK, MAKŠ** (Pišce, Austro-Hungarian empire [now in Slovenia] 1840 – Pišce, Kingdom of Serbs, Croats, and Slovenes [now in Slovenia] 1923). Compiler of a Slovene–German dictionary complementing that of M. CIGALE, and contributor to the Church Slavic dictionary of F. MIKLOŠIČ.
- PODOLSKY, BARUCH** (Moscow 1940 – Holon, Israel 2011). Linguist; compiler of Hebrew–Russian dictionaries.
- POLLUX, JULIUS** (from Naucratis, Roman empire [now south-east of Alexandria, Egypt]; fl. 166×176). Scholar and rhetorician; compiler of an important Greek dictionary, the *Onomasticon*, which survives in abridged form.
- POMEY, FRANÇOIS-ANTOINE** (Prenes-les-Fontaines, France 1618 – Lyon, France 1673). Jesuit; compiler of a French–Latin dictionary and of the subject-ordered *Indiculus universalis*.
- POMPEIUS FESTUS, SEXTUS** (from the Roman empire; late 2nd cent.). Abridger of the *De verborum significatu* of VERRIUS FLACCUS.
- POND, SAMUEL** (New Preston, Connecticut 1808 – Shakopee, Minnesota 1891). Presbyterian missionary to the Dakota people; compiler of a Hebrew–Dakota dictionary.
- PONS, JEAN-FRANÇOIS** (diocese of Rodez, France 1698 – Chandernagore, Mughal empire [now Chandannagar, India] 1752). Jesuit missionary in India; he translated the *Amarakośa* of AMARASIMHA into Latin.
- POTIER, PIERRE-PHILIPPE** (Blandain, Austrian Netherlands [now in Belgium] 1708 – Notre-Dame-de-l'Assomption, British North America [now Windsor, Ontario] 1781). Jesuit missionary in what is now Canada; compiler of a major dictionary of Huron, and of a wordlist of the French spoken in North America.
- POTT, AUGUST** (Nettelrede, Hannover [now in Germany] 1802 – Halle (Saale), German empire 1887). Philologist; the second volume of his *Die Zigeuner in Europa und Asien* is a landmark scholarly dictionary of Romani.

- PRATT, TERRY KENNETH (b. Toronto, Canada 1943).** Professor of English; editor of the *Dictionary of Prince Edward Island English* and (with S. Burke) of *Prince Edward Island Sayings*.
- PROENÇA, ANTAÔ DE (Ramela, Portugal 1625 – Ramanathapuram [now in India] 1666).** Jesuit missionary in India; compiler of the first missionary dictionary of Tamil.
- PURUŠOTTAMADEVA (from Bengal [now in India]; first half of the 12th cent.?).** Author of the *Trikāṇḍaśeṣa*, a Sanskrit lexicon compiled with the aim of supplementing the *Amarakośa* of AMARASIMHA.
- PUTSILLO, MIKHAIL PAVLOVICH (from the Russian empire; 1845–1889).** Official and historian; compiler of a Russian–Korean dictionary.
- AL-QĀLĪ (Malazgirt, Abbasid Caliphate [now in Turkey] 288/901 – al-Andalus [now in Spain] 356/967).** Compiler of the first major lexicon in Arab Spain.
- QUINTANA, AUGUSTÍN DE (Antequera, Spanish empire [now in Mexico] c. 1660 – Antequera? 1734).** Dominican missionary in what is now Mexico; author of works in and about Mixe, including a wordlist.
- QUTRUB (d. 206/821).** Grammarian; compiler of the first collection of groups of Arabic triplets, each word of which is distinguishable in meaning by its characteristic vowel.
- RADA CRUZAT, MARTÍN DE (Pamplona, Spain 1533 – at sea between Borneo and the Philippines 1578).** Augustinian missionary in the Philippines; he is said to have compiled a grammar and vocabulary of Hokkien, now lost, and if so, he was perhaps the first European to make a wordlist of any variety of Chinese.
- RADLOFF, FRIEDRICH WILHELM (VASILIJ VASIL'EVICH RADLOV) (Berlin 1837 – Petrograd [now St Petersburg] 1918).** Turcologist and ethnographer; compiler of a comparative Turkic dictionary.
- RALE (RASLES), SÉBASTIEN (Pontarlier, France 1657 – Norridgewock, Abenaki territory [now Madison, Maine] 1724).** Jesuit missionary in North America; compiler of an Abenaki dictionary.
- RAMSON, WILLIAM STANLEY (Lower Hutt, New Zealand 1933 – Sydney, Australia 2011).** Professor of English and university administrator; first editor of the *Australian National Dictionary*.
- RAPHELENGIUS, FRANCISCUS (FRANS VAN RAVELINGEN) (Lannoy, county of Flanders [now in France] 1539 – Leiden, Netherlands 1597).** Printer and Hebraist; compiled the first Arabic–Latin dictionary.
- RASHI (R. SHELOMO ITZHAKI, RABBI SOLOMON BEN ISAAC) (Troyes, France 1040 – Troyes 1105).** Rabbi and commentator on the Bible and the Talmud (the lexical material in his commentaries has been gathered and presented as a dictionary by Yitṣḥaq Avineri); maker of a major group of early Old French glosses in Hebrew characters.

- RASTELL, JOHN** (Coventry, Warwickshire? c. 1475 – London 1536). Lawyer and printer; compiler of a dictionary of legal terms in Law French and English.
- RAVIUS, CHRISTIAN** (Berlin 1613 – Frankfurt an der Oder, Brandenburg [now in Germany] 1677). Orientalist and biblical scholar; compiler of a *Specimen lexicī Arabico–Persico–Latini*.
- RAY, JOHN** (Black Notley, Essex 1627 – Black Notley 1705). Theologian and naturalist; compiler of the first printed dialect wordlist of English.
- RAYNOUARD, FRANÇOIS JUST MARIE** (Brignoles, France 1761 – Passy, France 1836). Dramatist and philologist; compiler of a six-volume dictionary of Old Occitan, to which a major supplement was added by E. LEVY.
- READ, ALLEN WALKER** (Winnebago, Minnesota 1906 – New York, New York 2002). Professor of English; assistant editor (under W. CRAIGIE) of the *Dictionary of American English*.
- REDHOUSE, JAMES WILLIAM** (Surrey? 1811 – London 1892). Orientalist and diplomat; interpreter to the grand vizier and the English government and compiler of a comprehensive *Turkish and English Lexicon*.
- RESTIVO, PABLO** (BLAS PRETOVIO) (Mazzarino, Spanish empire [now in Italy] 1658 – Candelaria, Spanish empire [now in Argentina] 1741). Jesuit missionary in South America; he revised and augmented the Guarani dictionaries of A. RUIZ DE MONTAYA.
- REUCHLIN, JOHANNES** (Pforzheim, Margraviate of Baden [now in Germany] 1455 – Stuttgart, Austrian empire [now in Germany] 1522). Humanist and Hebraist; the first Christian to compile a wordlist of Hebrew, and editor of the first printed text of the Latin–German *Vocabularius breuiloquus*.
- REUTHER, JOHANN GEORG** (Roßtal, Kingdom of Bavaria [now in Germany] 1861 – near Eudunda, South Australia 1914). Lutheran missionary in Australia; compiler of a thirteen-volume linguistic and anthropological study including Pama-Nyungan comparative wordlists and a dictionary of place-names.
- RÉZEAU, PIERRE** (b. Vouvant, France, 1938). Priest and philologist; compiler of a dictionary of French regionalisms.
- RHODES, ALEXANDRE DE** (Avignon, France 1591 – Isfahan, Safavid empire [now in Iran] 1660). Jesuit missionary in Vietnam; compiler of the first missionary dictionary of Vietnamese.
- RICCI, MATTEO** (Macerata, Papal States [now in Italy] 1552 – Běijīng 1610). Jesuit missionary in China; probable compiler with M. RUGGIERI of *Dicionário Português–Chinês*, the first extant dictionary of a variety of Chinese and a European language.
- RICHARDSON, CHARLES** (Hoxton, Middlesex 1775 – Feltham, Middlesex 1865). Teacher; his *New Dictionary of the English Language* is an important predecessor of the *Oxford English Dictionary*.

- RICHARDSON, JOHN** (Edinburgh? 1740×1741 – Calcutta, Bengal Presidency [now Kolkata, India] 1795). Compiler of the first comprehensive Persian–English dictionary.
- RICHELET, PIERRE** (Cheminon, France 1626 – Paris 1698). Compiler of the first fully monolingual dictionary of French.
- RIDEL, FÉLIX-CLAIR** (Chantenay, France 1830 – Vannes, France 1884). Titular bishop of Philippopolis and Vicar Apostolic of Korea; compiler of a Korean–French dictionary.
- RIDLEY, MARK** (Stretham, Cambridgeshire 1560 – London? 1621×1624). Physician to the tsar of Russia and writer on magnetism; compiler of substantial Russian–English wordlists.
- RIDLEY, WILLIAM** (Hartford End, Essex 1819 – Paddington, New South Wales 1878). Itinerant missionary in Australia; compiler of wordlists of Gamilaraay and other Australian languages.
- RIFAT, KILISLI** (Kilis, Ottoman empire [now in Turkey] 1874 – Ankara 1953). Philologist; editor of MAHMŪD AL-KĀŠGARĪ's *Dīwān Luġāt at-Turk*.
- RIGGS, STEPHEN RETURN** (Steubenville, Ohio 1812 – Beloit, Wisconsin 1883). Presbyterian missionary to the Dakota people; compiler of dictionaries of the Dakota language.
- RINALDINI, BENITO** (Brescia, Venetian Republic [now in Italy] 1695 – Guadalupe y Calvo, Spanish empire [now in Mexico] 1764). Jesuit missionary in what is now Mexico; author of a grammar and dictionary of Northern Tepehuan.
- RIVERO, JUAN** (Miraflores de la Sierra, Spain 1681 – Puerto de San Salvador de Casanare, Spanish empire [now in Colombia or Venezuela] 1736). Jesuit missionary in South America; he completed the Achagua grammar and dictionary of A. DE NEIRA.
- RIVOLA, FRANCESCO** (from Milan; d. after 1631). Priest; compiler of the first printed dictionary of Armenian.
- ROBERT, PAUL** (Orléansville, Algeria 1910 – Mougins, France 1980). Publisher and compiler of an important series of French dictionaries.
- ROBOREDO, MANUEL DE** (name in religion FRANCISCO DE SÃO SALVADOR) (b. Kingdom of Kongo [now in Angola and the Democratic Republic of the Congo] a. 1620; d. Ambuíla, Kingdom of Kongo [now in Angola] 1665). Capuchin friar and priest; probable compiler of a Latin–Spanish–Kikongo dictionary, which has also been attributed to the copyist of the extant manuscript, A. WILLEMS.
- ROTH, HEINRICH** (Dillingen, Bishopric of Augsburg [now in Germany] 1620 – Agra, Maratha empire [now in India] 1668). Jesuit missionary in India and author of the earliest extant grammar of Sanskrit by a European; his work

- towards a Sanskrit–Latin dictionary (which was never realized) makes him the first European lexicographer of Sanskrit.
- ROTH, RUDOLF VON** (Stuttgart, Württemberg [now in Germany] 1821 – Tübingen, Germany 1895). Philologist; compiler, with O. VON BÖHTLINGK, of a seven-volume *Sanskrit-Wörterbuch*.
- RUDDIMAN, THOMAS** (Raggel, Scotland 1674 – Edinburgh 1757). Printer and librarian; his glossary to his edition of Gavin Douglas' version of the *Aeneid* is the first printed wordlist of Middle Scots.
- RUFUS OF EPHEBUS** (from the Roman empire, 2nd half of the 1st cent. AD). Physician; compiler of an influential anatomical lexicon.
- RUGGIERI, MICHELE** (Spinazzola, Spanish empire [now in Italy] 1543 – Salerno, Spanish empire [now in Italy] 1607). Jesuit missionary in China; probable compiler with M. RICCI of *Dicionário Português–Chinês*, the first extant dictionary of a variety of Chinese and a European language.
- RUHIG, PHILIPP** (Kattenau, Duchy of Prussia [now Zavety, Kaliningrad Oblast, Russia] 1675 – Walterkehmen [now Ol'xovatka, Kaliningrad Oblast, Russia] 1749). Lutheran pastor and author of religious works; compiler of a bidirectional Lithuanian and German dictionary, which became a basis for the work of C. MIELCKE.
- RUIZ BLANCO, MATÍAS** (Estepa, Spain 1643 – province of Cumana, Spanish empire [now in Venezuela] 1705). Franciscan missionary in South America; compiler of a wordlist of Cumanagoto.
- RUIZ DE MONTOYA, ANTONIO** (Lima 1585 – Lima 1652). Jesuit missionary in South America; compiler of the first dictionaries of Guarani, revised and augmented by P. RESTIVO.
- RUMPHIUS, GEORG EVERHARD** (Wölfersheim, Holy Roman empire [now in Germany] – Ambon, Dutch East Indies [now in Indonesia] 1702). Merchant and naturalist; compiler of an unfinished Malay dictionary.
- SAADIA GAON** (SA'ĪD B. YŪSUF AL-FAYYŪMĪ) (Fayyūm, Abbasid Caliphate [now in Egypt] 882 – Baghdad 942). Rabbi, Gaon of the Jewish academy of Sura (Iraq), philosopher, and exegete; compiler of the Hebrew lexicographical work called the *Egron*.
- AL-ṢAGĀNĪ** (d. 650/1252). Compiler of *al-'Ubāb al-zāḥir*, one of the most organized Arabic lexica in the internal arrangement of its lemmata.
- SAGARD, GABRIEL** (from France; fl. 1614–1632). Recollect missionary in what is now Canada; compiler of a French–Huron wordlist.
- AL-ṢĀḤĪB BIN 'ABBĀD** (d. 385/995). Compiler of the phonetically arranged Arabic lexicon *al-Muḥīṭ fī l-luġa* (II vols.).

- ȘĂINEANU, LĂZAR (LAZARE SAINÉAN) (Ploiești, Ottoman empire [now in Romania] 1859 – Paris 1934).** Philologist and folklorist; compiler of an encyclopedic dictionary of Romanian.
- SAKAEDA TAKEI (Kōchi, Japan 1879 – Ninomiya, Japan 1962).** Linguist; lead editor of the sinograph dictionary *Daijiten*.
- SALAKṢA (from Gujarat [now in India]; fl. 1365).** Compiler of the earliest-known Sanskrit–Persian lexicon.
- SALE, FLAMINIO DA (Sale Marasino, Venetian Republic [now in Italy] 1667 – Obervatz, canton of Graubünden [now in Switzerland] 1733).** Capuchin missionary in Graubünden; his *Fundamenti principali della lingua retica, o griggiona* includes the first printed wordlist of Romansh.
- SALESBURY, WILLIAM (b. Llansannan, Wales a. 1520; d. after 1574).** Humanist and Bible translator; compiler of *Dictionary in Englyshe and Welshe*, the first free-standing printed dictionary of English and another vernacular.
- SALLE DE L'ETANG, SIMON PHILIBERT DE LA (Reims, France c. 1700 – Paris 1765).** Agronomist; compiler, from written sources, of a bidirectional dictionary of French and Carib.
- SALVÁ, VICENTE (Valencia, Spain 1786 – Paris 1849).** Politician, bibliophile, and literary scholar; the best-known Spanish lexicographer of the nineteenth century.
- SALVADO, ROSENDO (Tui, Spain 1814 – Rome 1900).** Benedictine missionary in Australia and titular bishop of Adriana; his *Memorie storiche* include a Wadjuk wordlist.
- SALYS, ANTANAS (Reketė, Russian empire [now in Lithuania] 1902 – Philadelphia, Pennsylvania 1972).** Linguist; co-editor of the *Wörterbuch der litauischen Schriftsprache, Litauisch–Deutsch*, started by M. NIEDERMANN and A. SENN; see also F. BRENDER.
- SAN ANTONIO, FRANCISCO DE (called 'OREJITA') (early 17th cent.).** Franciscan missionary in the Philippines; compiler of a manuscript Tagalog–Spanish wordlist.
- SAN BUENAVENTURA, PEDRO DE (fl. 1594; d. at sea between the Philippines and Mexico 1627).** Franciscan missionary in the Philippines; compiler of the first extant dictionary of Tagalog.
- SANDERS, DANIEL (Grand Duchy of Mecklenburg-Strelitz [now in Germany] 1819 – Strelitz 1897).** Teacher; compiler of a German dictionary in conscious opposition to the principles of the brothers GRIMM (1860–5).
- SANDERS, WILLIAM HENRY (Tillypally, Ceylon [now Tellipallai, Sri Lanka] 1856 – Pittsburgh, Pennsylvania? 1947).** Missionary (American Board of Commissioners for Foreign Missions) in Angola; compiler, with W. FAY, of the first Umbundu dictionary.

- SAN LÚCAR, PEDRO (b. 1706; fl. 1754).** Jesuit missionary in the Philippines; compiler, with J. DE NOCEDA, of a Tagalog dictionary.
- SANTO TOMÁS, DOMINGO DE (Seville, Spain 1499 – La Plata, Spanish empire [now Sucre, Bolivia] 1570).** Dominican missionary in the Spanish Viceroyalty of Peru and bishop of La Plata o Charcas; compiler of the first printed dictionary of Quechua.
- SANZ DE GALDEANO, SERAPHIM (Villatuerta, Spain 1913 – Perth, Australia 2008).** Benedictine missionary in Australia; compiler of a grammar and dictionary of Gunin/Kwini.
- AL-SARAQUSTĪ (from al-Andalus [now in Spain]; d. after 400/1010).** Compiler of an Arabic lexicon devoted to verbal patterns.
- ŚARMĀ, VAIJNĀTH (fl. Calcutta, Bengal Presidency [now Fort William, India] 1804–1810).** Chief Marathi pandit at the College of Fort William; collaborator with W. CAREY on a Marathi dictionary.
- ŚĀŚVATA (from India; 5th×10th cent.).** Compiler of a Sanskrit homonymic lexicon.
- AL-ŠAYBĀNĪ, ABŪ ‘AMR (from Kūfa, Umayyad Caliphate [now in Iraq]; d. Baghdad 206/821).** Compiler of *Kitāb al-Ġim*, the first semasiological Arabic lexicon to adopt alphabetical order.
- SCAPULA, JOHANNES (JEAN ESPULAZ) (from Lausanne; 16th cent.).** Professor of Greek; compiler of a successful abridgement of the *Thesaurus graecae linguae* of H. ESTIENNE.
- SCHELE DE VERE, MAXIMILIAN RUDOLPH (Växjö, Sweden? 1820 – Washington, DC 1898).** Professor of modern languages; compiler of *Americanisms: The English of the New World*.
- SCHMICK, JOHANN JACOB (Königsberg, Kingdom of Prussia [now Kaliningrad, Russia] 1714 – Lititz, Pennsylvania 1778).** Moravian missionary in North America; compiler of a dictionary of Mahican.
- SCHNEIDER, LUCIEN (Paris 1907 – Paris 1978).** Oblate missionary in Arctic Canada; compiler of dictionaries of Eastern Canadian Inuktitut.
- SCHÖN, JAMES FREDERICK (JACOB FRIEDERICH) (Ober Weiler, Baden [now in Germany] 1802 – New Brompton, Kent 1889).** Missionary (Church Missionary Society) in west Africa; compiled the first dictionary of Hausa, and assisted S. CROWTHER in the preparation of the first substantial dictionary of Igbo.
- SCHOONEES, PIETER CORNELIS (Uniondale, Cape Colony [now in South Africa] 1891 – Strand, South Africa 1970).** Editor of the first published volumes of the *Woordeboek van die Afrikaanse taal*.
- SCHOTTELIUS, JUSTUS GEORGIUS (Einbeck, Braunschweig-Wolfenbüttel [now in Germany] 1612 – Wolfenbüttel 1676).** Poet and promoter of the

- German language; his *Ausführliche Arbeit von der teutschen HauptSprache* includes an important German wordlist.
- SCHREGLE, GÖTZ (b. Erlangen, Germany 1923).** Compiler of a German–Arabic dictionary and an incomplete Arabic–German dictionary.
- SCHREVELIUS, CORNELIUS (Haarlem, Netherlands 1608 – Leiden, Netherlands 1661).** Teacher; compiler of a Greek–Latin school dictionary.
- SCHRODERUS, ERICUS JOHANNIS (Uppsala, Sweden 1609 or earlier – Stockholm 1639).** Compiler of the first printed wordlist of Finnish.
- SCHUEREN, GERARD VAN DER (GERT VAN DER SCHUREN) (from Xanten, duchy of Cleves [now in Germany]; d. in or after 1481).** Secretary to the duke of Cleves; compiler of a dictionary of the Germanic language variety of the Duchy of Cleves, which has been seen as the first printed dictionary of Dutch.
- SCHULTZE, BENJAMIN (Sonnenburg, Duchy of Prussia [now Słońsk, Poland] 1689 – Halle (Saale), Kingdom of Prussia [now in Germany] 1760).** Lutheran missionary (Dänisch-Hallesche Mission) in India; compiler of a comparative wordlist of a number of the languages of India.
- SCHUMANN, CHRISTIAN LUDWIG (Pilgerhut, Berbice [now in Guyana] 1749 – Brüdergarten, Tranquebar [now Tharangambadi, India] 1794×1795).** Moravian missionary; compiler of the first dictionaries of Saramaccan and Sranan Tongo.
- SCHÜRMANN, CLAMOR WILHELM (Schledehausen, Kingdom of Hannover [now in Germany] 1815 – Bethanien [now Bethany], South Australia 1893).** Lutheran missionary in Australia; compiler of a dictionary of Barngala and, with C. TEICHELMANN, and taught by MULLAWILLABURKA, of a grammar and dictionary of Kaurna.
- SECO RAYMUNDO, MANUEL (b. Madrid 1928).** Writer on the Spanish language; compiler of the the first synchronic, descriptive dictionary of peninsular Spanish.
- SEDELMAYR, JACOBO (Bishopric of Freising [now in Germany] 1703 – Aldea de Avila, Spain 1779).** Jesuit missionary in what are now parts of the United States and Mexico; he is said to have compiled a lost wordlist of O’odham.
- SEMSEDDIN SAMI (SAMI FRASHËR) (Frashër, Ottoman empire [now in Albania] 1850 – Istanbul, Turkey 1904).** Intellectual, philosopher, and linguist; compiler of the *Kâmûs-ı Türkî*.
- SENISIO, ANGELO (Catania, Sicily 1305 – San Martino delle Scale, Sicily 1386).** Benedictine abbot; compiler of the Latin–Sicilian *Declarus*.
- SENN, ALFRED (Blotzheim, German empire [now in France] 1899 – Ashord, Connecticut 1978).** Indo-Europeanist; one of the original editors (with M. NIEDERMANN) of the bilingual *Wörterbuch der litauischen Schriftsprache*.
- SERVIUS (from Rome; fl. late 4th and early 5th cents.).** Grammarian; author of a Latin commentary on Vergil with influential lexical information.

- SHĀD, MUḤAMMAD PĀDShĀH** (from Vizianagaram, India; fl. 1877–1889). Chief secretary to the maharajah of Vizianagaram; compiler of the *Farhang-i Ānandrāj*, the largest dictionary compiled in the Persianate world before the twentieth century.
- SHAPIRO, FELIKS** (Xoluj, near Babrujsk [now in Belarus] 1879 – Moscow 1960). Teacher; compiler of an important Hebrew–Russian dictionary.
- SHELOMO BEN SAMUEL** (from Djurdjāniyya, Khorasan [now Kunya-Urgench, Turkmenistan]; fl. 1339). Compiler of a Hebrew–Persian dictionary.
- SHERIDAN, THOMAS** (Ireland 1719? – Margate, Kent 1788). Actor; compiler of an influential pronouncing dictionary of English.
- SHĪ YÓU** 史游 (from China; second half of the 1st cent. BC). Compiler of the mnemonic wordlist *Jíjiù piān*.
- SHINMURA IZURU** (Yamaguchi, Japan 1876 – Kyoto, Japan 1967). Historical linguist and etymologist; compiler with his son SHINMURA TAKESHI of the influential Japanese dictionary *Kōjien*.
- SHINMURA TAKESHI** (Tokyo 1905 – Nagoya, Japan 1992). Scholar and translator of French literature; compiler with his father SHINMURA IZURU of the influential Japanese dictionary *Kōjien*.
- SHŌJŪ** 昌住 (from Japan; fl. c. 900). Buddhist monk; compiler of *Shinsen jikyō*, a Chinese dictionary with Japanese glosses, which is the first extant dictionary to include Japanese.
- SHŪ XĪNCHÉNG** 舒新城 (Xùpǔ County, China 1893 – Shànghǎi, China 1960). Compiler of the monolingual Chinese dictionary *Cihǎi*.
- SĪBAWAYHI** (perhaps from Shiraz [now in Iran]; d. Fars?, Abbasid Caliphate [now in Iran] 180/796). Author of *al-Kitāb*, the first and most influential book in the Arabic grammatical tradition.
- SILVA, DUARTE DE** (Portugal c. 1527 – Japan 1564). Jesuit missionary in Japan; said to have been the compiler of a lost dictionary of Japanese, perhaps the earliest by a European.
- SILVY, ANTOINE** (Aix-en-Provence, France 1638 – Québec, New France [now Quebec City, Canada] 1711). Jesuit missionary in what is now Canada; compiler of an early dictionary of Montagnais.
- SIM UILIN** (Seoul 1894 – Busan, Korea 1951). Teacher, grammarian, and lexicographer; compiler of the first monolingual Korean dictionary.
- SINCLAIR, JOHN McHARDY** (Edinburgh 1933 – Florence, Italy 2007). Corpus linguist; editor of the *Collins COBUILD English Language Dictionary*.
- SIRVYDAS, KONSTANTINAS** (CONSTANTINUS SZYRWID) (Sirvydai?, Polish-Lithuanian Commonwealth [now in Lithuania] 1579 – Vilnius, Polish-Lithuanian Commonwealth [now in Lithuania] 1631). Jesuit priest and author; compiler of the first Lithuanian dictionary, a Polish–Latin–Lithuanian lexicon.

- SITJAR, BUENAVENTURA** (Porreres, Spain 1739 – Mission San Antonio de Padua, Mexico [now near Jolon, California] 1808). Franciscan missionary in California; compiler of a dictionary of Salinan.
- SI TU PAṆ CHEN** (from Tibet; 1699×1700–1774). Polymath and artist; reviser of Tibetan translations of Sanskrit lexicographical works, and author of a Tibetan commentary on the *Amarakośa* of AMARASIMHA.
- SKINNER, STEPHEN** (baptized London 1623; d. Lincoln, Lincolnshire 1667). Physician; compiler of the first printed etymological dictionary of English (1671, with Latin as metalanguage; an abridgement with English as metalanguage was published by R. HOGARTH).
- SOLANA, ALONSO DE (LA)** (from Solana, Spain; d. Mérida, Spanish empire [now in Mexico] 1600). Franciscan missionary in Mesoamerica; compiler of Maya dictionaries.
- SOLOMON BEN ABRAHAM** (from Urbino [now in Italy]; fl. c. 1480). Compiler of the Hebrew synonym dictionary *Ohel Mo'ed*.
- SOMNER, WILLIAM** (baptized Canterbury, Kent 1598; d. Canterbury 1669). Antiquary; compiler of the first printed dictionary of Old English.
- ŚRĪDHARASENA** (from India; c. 1150×1261). Compiler of a Sanskrit lexicon with synonymic and homonymic parts, the former now extant only in a Tibetan translation.
- STAHL, HEINRICH** (Reval, Kingdom of Sweden [now Tallinn, Estonia] c. 1599 – Narva, Kingdom of Sweden [now in Estonia] 1657). Minister of religion; compiler of the first printed dictionary of Estonian.
- STEERE, EDWARD** (London 1828 – Mkunazini, Zanzibar [now in Tanzania] 1882). Anglican bishop of central Africa and translator; his *Handbook of the Swahili Language* (1870) includes the first dictionary of the Unguja variety spoken at Zanzibar.
- STEFFEL, MATTHÄUS** (Jihlava, Habsburg empire [now in the Czech Republic] 1734 – Brno, Austrian empire [now in the Czech Republic] 1806). Jesuit missionary in what is now Mexico; compiler of a dictionary of Tarahumara.
- STEINBACH, CHRISTOPH ERNST** (Sammelwitz, Austrian empire [now Żębowice, Poland] 1698 – Breslau, Austrian empire [now Wrocław, Poland] 1741). Physician and philologist; compiler of the first German dictionary to give literary citations.
- STEINGASS, FRANCIS JOSEPH** (Frankfurt am Main, German Confederation 1825 – London 1903). Linguist and teacher; compiler of a Persian–English dictionary based on the work of FRANCIS JOHNSON.
- STEINITZ, WOLFGANG** (Breslau, Germany [now Wrocław, Poland] 1905 – Berlin 1967). Finno-Ugrian philologist; founding editor, with R.

- KLAPPENBACH, of the first post-Second World War German dictionary, *Wörterbuch der deutschen Gegenwartssprache*.
- STENDER, GOTTHARD FRIEDRICH (Lassen, Duchy of Courland and Semigallia [now Laši, Latvia] 1714 – Sonnaxt, Russian empire [now Sunākste, Latvia] 1796). Lutheran pastor, philologist, and poet; compiler of a bidirectional dictionary of Latvian and German.
- STENTON, FRANK MERRY (Upper Norwood, Surrey [now part of London] 1880 – Reading, Berkshire 1967). Historian of medieval England; compiler (often in collaboration with A. MAWER) of several place-name dictionaries.
- STEPHANUS A S.S. PETRO ET PAULO (perhaps from the Venetian Republic; 1692–1766×1767). Carmelite missionary in India; compiler of a Portuguese–High Malayalam–Latin dictionary.
- STEPHANUS OF BYZANTIUM (from Constantinople? [now Istanbul, Turkey]; fl. early 6th cent.). Grammarian; compiler of the *Ethnica*, a Greek geographic lexicon.
- STEUERWALD, KARL (Strasburg, German empire [now Strasbourg, France] 1905 – Neu-Ulm, West Germany [now Germany] 1989). Linguist and Turcologist; compiler of a Turkish–German dictionary.
- STEWART, GEORGE RIPPEY (Sewickley, Pennsylvania 1895 – San Francisco, California 1980). Historian and novelist; compiler of a dictionary of American place-names.
- STIELER, KASPAR (Erfurt, Holy Roman empire [now in Germany] 1632 – Erfurt 1717). Poet and philologist; his *Der Teutschen Sprache Stammbaum und Fortwachs, oder Teutscher Sprachschatz* was, in its time, the most extensive printed dictionary of any Germanic language.
- STIERNHIELM, GEORG (Vika, Sweden 1598 – Stockholm 1672). Civil servant, poet, and philologist; he projected an etymological dictionary of the languages of the world.
- STORY, GEORGE MORLEY (St John's, Newfoundland [now in Canada] 1927 – St John's? 1994). Literary scholar and folklorist; compiler (with W. KIRWIN and J. WIDDOWSON) of the *Dictionary of Newfoundland English*.
- STREHLOW, CARL FRIEDRICH THEODOR (Uckermark, German empire 1871 – Horseshoe Bend, Australia 1922). Lutheran missionary in Australia; father of T. STREHLOW; compiler of an unpublished Western Arrernte–German–Luritja–Diyari dictionary.
- STREHLOW, THEODOR GEORGE HENRY (Hermannsburg, Australia 1908 – Adelaide, Australia 1978). Linguist, and author of *Songs of Central Australia*; son of C. STREHLOW; compiler of a large unpublished dictionary of Western Arrernte.

- SUBHŪTICANDRA (from India; 1060×1172).** Author of a commentary on the *Amarakośa* of AMARASIMHA, which was translated into Tibetan in or before the mid fourteenth century.
- SUETONIUS TRANQUILLUS, GAIUS (from the Roman empire, c. 70 – c. 130).** Biographer and imperial administrator; compiler of lexicographical tracts in Greek.
- SUGHDI (between the 7th and 11th cents.).** Compiler of *Risālah*, the oldest recorded Persian wordlist (no longer extant).
- SURŪRĪ, MUHAMMAD QĀSIM (from Isfahan, Safavid empire [now in Iran]; fl. late 16th and early 17th cents.).** Compiler of the Persian dictionary *Majma' al-Furs*.
- SVABO, JENS CHRISTIAN (Vágar, Faroe Islands 1746 – Tórshavn, Faroe Islands 1824).** Ethnographer; compiler of the first dictionary of Faroese.
- SWĚTLIK (SWÓTLIK), JURIJ HAWŠTYN (GEORGIUS AUGUSTINUS SWÓTLIK) (Wittichenau, Electorate of Saxony [now in Germany] 1650 – Bautzen, Electorate of Saxony [now in Germany] 1729).** Catholic priest and translator; compiler of the first printed Sorbian dictionary.
- SYMEON (from the Byzantine empire; fl. first half of the 12th cent.).** Grammarian; compiler of the Greek *Etymologicum Symeonis* and of an unpublished Greek synonym dictionary.
- TABERD, JEAN-LOUIS (Saint-Étienne, France 1794 – Calcutta [now Kolkata], India 1840).** Missionary (Missions étrangères de Paris) in Vietnam and India, and titular bishop of Isauropolis; editor of the Vietnamese–Latin and Latin–Vietnamese dictionaries of P. PIGNEAU DE BEHAINE.
- TANḤUM YERUSHALMI (d. Egypt, 1291).** Exegete; compiler of the post-Biblical Hebrew dictionary *Al-Muršid al-Kāfi*.
- TANIKAWA KOTOSUGA (Ise, Japan 1709 – Ise 1776).** *Kokugaku* (nativist philology) scholar; compiler of the innovative Japanese dictionary *Wakun no shiori*.
- TAPLIN, GEORGE (Kingston-upon-Thames, Surrey 1831 – Raukkan, South Australia 1879).** Congregationalist missionary in Australia; compiler of an unpublished Ngarrindjeri wordlist and a comparative wordlist of Australian languages.
- TAUSTE, FRANCISCO DE (MIGUEL TORRALBA DE RADA) (Tauste, Spain 1626 – province of Cumana, Spanish empire [now in Venezuela] 1685).** Franciscan missionary in South America; compiler of a wordlist of Cumanagoto.
- TEICHELMANN, CHRISTIAN GOTTLÖB (Dahme, Kingdom of Saxony [now in Germany] 1807 – Stansbury, South Australia 1888).** Lutheran missionary in Australia; compiler, with C. SCHÜRMANN, and taught by MULLAWILLABURKA, of a grammar and dictionary of Kaurna.

- TELEPHUS OF PERGAMUM (from the Roman empire, 2nd cent. AD).** Grammarian; compiler of a Greek lexicon for words denoting things in common use.
- TENCH, WATKIN (Chester, Cheshire 1758×1759 – Devonport, Devonshire 1833).** Soldier, and writer about Australia; he made an early wordlist of Iyura.
- TENG SHOU-HSIN (DÈNG SHǒUXÌN) 鄧守信 (b. Huālián, Taiwan 1940).** Editor of an English–Chinese dictionary.
- TERREROS Y PANDO, ESTEBAN DE (Trucíos, Spain 1707 – Forlì, Papal States [now in Italy] 1782).** Jesuit priest; compiler of a four-volume encyclopedically oriented Spanish dictionary.
- TERUEL, ANTONIO (b. 1604).** Capuchin missionary in the Kingdom of Kongo, and compiler of a lost Latin–Spanish–Italian–Kikongo dictionary.
- Tew, GWILYM (from Glamorgan, Wales; fl. 1460–1480).** Poet; compiled a monolingual Welsh wordlist of vocabulary from a medieval poetic manuscript.
- THEON (from Alexandria, Egypt; 1st cent. BC).** Grammarian; he has been identified as the author of a monolingual Greek glossary of words occurring mostly in comedy.
- THIBAUT, ANDRÉ (b. Sept-Îles, Canada 1963).** Linguist and historian of the French language; compiler of *Dictionnaire suisse romand*.
- THOMAS MAGISTROS (name in religion THEODOULOS) (Thessalonica, Byzantine empire [now in Greece] c. 1280 – Thessalonica c. 1347).** Diplomat and scholar; his *Ecloga* is an important Atticist Greek lexicon.
- THOMAS, ANDRÉ ANTOINE (Saint-Yrieix-la-Montagne, France 1857 – Paris 1935).** Archivist, medievalist, and philologist; compiler, with A. HATZFELD, and in succession to A. DARMESTETER, of a major French dictionary.
- THORNTON, RICHARD HOPWOOD (Didsbury, Lancashire 1845 – Portland, Oregon 1925).** Professor of law; compiler of *An American Glossary*.
- THRELKELD, LANCELOT (London 1788 – Sydney, New South Wales 1859).** Missionary (London Missionary Society) in Australia; compiler, instructed by BIRABAN, of a pioneering grammar and wordlist of Awabakal.
- TIETZE, ANDREAS (Vienna 1914 – Vienna 2003).** Linguist and Turcologist; compiler of a great historical-etymological dictionary of Turkish.
- TIKTIN, HEIMANN HARITON (Breslau, German Confederation [now Wrocław, Poland] 1850 – Berlin 1936).** Linguist; compiler of one of the classic dictionaries of Romanian.
- TINDALE, NORMAN B. (Perth, Western Australia 1900 – Palo Alto, California 1993).** Anthropologist and entomologist; he gathered wordlists in numerous Australian languages.

- TOBLER, ADOLF** (Hirzel, Switzerland 1835 – Berlin 1910). Philologist; his notes were the basis for an eleven-volume dictionary of Old French edited by his former pupil E. LOMMATZSCH.
- TOLKĀPIIYAṆ** (southern India, earlier 1st millennium AD?). This name is given in a preface to the Tamil grammar called the *Tolkāppiyam* as that of its author and eponym (the name has traditionally been analysed as Tol Kāppiyaṇ ‘Kāppiyaṇ the elder’; other explanations have been proposed).
- TOMMASEO, NICCOLÒ** (Šibenik, Habsburg empire [now in Croatia] 1802 – Florence, Italy 1874). Writer and Italian patriot; compiler, with B. BELLINI, of the seven-volume *Dizionario della lingua italiana*.
- TORRES RUBIO, DIEGO DE** (Alcazar de San Juan, Spain 1557 – Chuquisaca, Spanish empire [now in Bolivia] 1638). Jesuit missionary; author of a grammar of Aymara with wordlists.
- TOURS, FRANÇOIS-MARIE DE** (d. 1709). Capuchin missionary in India; compiler of an unpublished French–Hindustani dictionary.
- TRABOT, PEREZ** (b. in France; settled in Italy c. 1395). Teacher; compiler of a Biblical Hebrew–Italian–Arabic dictionary.
- TRÁVNÍČEK, FRANTIŠEK** (Spešov, Austria-Hungary [now in the Czech Republic] 1888 – Brno, Czechoslovakia [now in the Czech Republic] 1961). Linguist; editor with P. VÁŠA of a comprehensive monolingual Czech dictionary.
- TRIGAULT, NICOLAS** (Douai, Spanish Netherlands [now in France] 1577 – Hángzhōu, China 1628). Jesuit missionary in China; compiler of *Xīrú Ěrmùzī*, the first Chinese dictionary into which there was a systematic attempt to integrate Romanized transcriptions of character readings.
- TRỊNH THỊ NGỌC TRÚC** (Thăng Long, Đại Việt [now Vietnam] 1595 – Bút Tháp Temple, Đại Việt [now Vietnam] 1660). Queen of Đại Việt, as consort of king Lê Thần Tông; probable compiler of the most significant dictionary of Vietnamese by a native speaker before the nineteenth century, *Chỉ nam ngọc âm giải nghĩa*.
- TRUMBULL, JAMES HAMMOND** (Stonington, Connecticut 1821 – Hartford, Connecticut 1897). Antiquarian; compiler of a dictionary of Massachusetts, on the basis of seventeenth-century texts.
- TỰ ĐỨC** 嗣德 (Huế, Vietnam 1829 – Huế 1883). Emperor of Vietnam; author of the Chinese–Vietnamese rhyming pedagogical dictionary *Tự Đức thánh chế tự học giải nghĩa ca*.
- UEDA KAZUTOSHI** (Edo [now Tokyo] 1867 – Tokyo 1937). Founder of modern Japanese linguistics/national language studies (*kokugogaku*); figurehead editor of the large monolingual Japanese dictionary *Dai Nihon kokugo jiten* and the sino-graph dictionary *Daijiten*.

- ULMANN, KARL CHRISTIAN (Riga, Russian empire [now in Latvia] 1793 – Walk, Russian empire [now Valga, Estonia] 1871). Professor of theology in Dorpat (now Tartu) and Lutheran bishop in St Petersburg; compiler of dictionaries of Latvian and German.
- UNDERWOOD, HORACE GRANT (WON DU-U) (London 1859 – Atlantic City, New Jersey 1916). Presbyterian missionary in Korea; compiler of the first bidirectional dictionary of Korean and English.
- URBANO, ALONSO (Mondéjar, Spain c.1529 – Tula de Allende, Spanish empire [now in Mexico] c. 1608). Franciscan missionary in what is now Mexico; compiler of a Spanish–Nahuatl–Otomi dictionary founded on the Spanish–Nahuatl dictionary of A. DE MOLINA.
- URI, JOANNES (JÁNOS, JOHN) (Nagykörös, Hungary 1724 – Oxford, Oxfordshire 1796). Cataloguer of Asian manuscripts; with W. JONES, planned a revision of the *Thesaurus* of F. MENINSKI with added Persian material.
- UŠAKOV, DMITRIJ NIKOLAEVIČ (Moscow 1873 – Tashkent, Soviet Union [now in Uzbekistan] 1942). Editor of a monolingual Russian dictionary in four volumes, which formed the basis for the work of S. OŽEGOV.
- VĀCASPATĪ (from India; fl. before 500 AD?). Compiler of the *Śabdārṇava*, a Sanskrit lexicon, now extant only in fragments.
- VALĀ, AḤMAD AL-‘AZĪZ AL-NĀ’ITĪ (from Hyderabad, India; d. 1924). Compiler of the unfinished *Aṣaf al-Lughāt*, the last important Persian dictionary to be compiled in the Indian subcontinent.
- VALDEMĀRS, KRIŠJĀNIS (CHRISTIAN WOLDEMAR) (Ārlava, Russian empire [now in Latvia] 1825 – Moscow 1891). Leading member of the Latvian First Awakening movement; compiler of a Russian–Latvian–German dictionary.
- VALDIVIA, LUIS DE (Granada, Spain 1561 – Valladolid, Spain 1642). Jesuit missionary in what is now Chile; compiler of the first printed wordlists of Mapudungun and of Allentiac and Millcayac.
- VALENTE, JOSÉ FRANCISCO (Unhais da Serra, Covilhã, Portugal 1912 – Lisbon 1993). Spiritan missionary in Angola; completed and published the Umbundu dictionary of G. LE GUENNEC.
- VALENTIANO, PIETRO LUPIS (perhaps from Valencia; fl. republic of Ancona? [now in Italy], in or before 1527). Compiler of Italian–Croatian and Italian–Turkish wordlists.
- VĀMANABHAṬṬA BĀṆA (from southern India; fl. 14th–15th cents.). Writer of prose and poetry; compiler of a Sanskrit synonymic and homonymic dictionary.
- VAN MOL, MARK LEOPOLD CECILE MARIA (b. Wilrijk, Belgium 1952). Professor of Arabic; compiler of Arabic–Dutch and Dutch–Arabic dictionaries.

- VĀRASTAḤ, SIYĀL KOTĪ MAL** (from Lahore [now in Pakistan]; fl. 1766). Compiler of the Persian dictionary *Mustalahāt al-Shu'arā* ('Expressions of the poets').
- VARO, FRANCISCO** (Seville, Spain 1627 – Fuzhou, China 1687). Dominican missionary to China; compiler of Portuguese–Mandarin and Spanish–Mandarin dictionaries.
- VĀŠA, PAVEL** (Čáslav, Austria-Hungary [now in the Czech Republic] 1874 – Brno, Czechoslovakia [now in the Czech Republic] 1954). Linguist; editor with F. TRÁVNÍČEK of a comprehensive monolingual Czech dictionary.
- VEFIK PAŞA, AHMET** (Istanbul 1823 – Istanbul 1891). Ottoman statesman, diplomat, playwright, and translator; compiler of the first monolingual Turkish dictionary.
- VELASCO, JUAN DE** (Riobamba, Spanish empire [now in Ecuador] 1727 – Faenza, Papal States [now in Italy] 1792). Jesuit missionary; compiler of a wordlist of a variety of Quechua spoken in Ecuador.
- VELASCO, JUAN BAUTISTA DE** (Oaxaca, Spanish empire [now in Mexico] 1562 – Mocorito, Spanish empire [now in Mexico] 1613). Jesuit missionary in what is now Mexico; a grammar and vocabulary of Yaqui have been attributed to him and to T. BASILIO.
- VERRIUS FLACCUS, MARCUS** (from the Roman empire; c. 55 BC? – c. 20 AD?). Scholar and teacher; his large and influential *De significatione uerborum* (preserved in an epitome by POMPEIUS FESTUS and in an abridgement of the epitome by PAUL THE DEACON) is the earliest Latin lexicon.
- VICO, DOMINGO DE** (b. Jaén, Spain; d. near Cobán, Spanish empire [now in Guatemala] 1555). Dominican missionary in Mesoamerica; his Kaqchikel–Spanish wordlist, transmitted in manuscript, is the earliest of a Mayan language.
- VIGUIER, PIERRE FRANÇOIS** (Besançon, France 1745 – Paris 1821). Lazarist father and Prefect Apostolic in Istanbul; author of a Turkish grammar and wordlist.
- VIKRAMĀDITYA** (from India; fl. after 500 AD?). Compiler of the *Saṃsārāvarta*, a Sanskrit lexicon, now extant only in fragments.
- VILAKAZI, BENEDICT WALLET** (formerly BAMBATHA KAMSHINI) (Groutville, Natal [now in South Africa] 1906 – Johannesburg, South Africa 1947). Poet, novelist, and teacher; compiler with C. DOKE of a significant Zulu–English dictionary.
- VIŚVANĀTHA** (from Vidarbha, Mughal empire [now in India]; fl. early 17th cent.). Man of letters; compiler of a large synonymic and homonymic dictionary of Sanskrit.
- VIVAR, PEDRO DE** (Logroño, Spain 1731 – northern Philippines 1771). Augustinian missionary in the Philippines; compiler of the first extant dictionary of Ilocano.

- VODNANSKÝ, JAN** (JAN BOSÁK Z VODNAN, JOHANNES AQUENSIS) (Vodňany, Bohemia [now in Czech Republic] c. 1460 – Bechyně?, Bohemia [now in Czech Republic] after 1534). Franciscan monk; compiler of the first Latin–Czech dictionary.
- VODNIK, VALENTIN** (Zgornja Šiška, Habsburg Monarchy [now in Slovenia] 1758 – Ljubljana, Austrian empire [now in Slovenia] 1819). Franciscan friar, teacher, and poet; compiler of a German–Slovene–Latin dictionary in manuscript, drawn on for the German–Slovene dictionary of M. CIGALE.
- VOSSIUS, GERARDUS JOANNES** (Heidelberg, Electoral Palatinate [now in Germany] 1577 – Amsterdam 1649). Humanist; compiler of *Etymologicon linguae Latinae*.
- VRANČIĆ, FAUST** (FAUSTUS VERANTIUS, FAUSTO VERANZIO) (Šibenik, Republic of Venice [now in Croatia] 1551 – Venice 1617). Bishop, polymath, and inventor; compiler of a multilingual dictionary.
- VULLERS, JOHANN AUGUST** (Bonn, French empire [now in Germany] 1803 – Bonn, German empire 1880). Professor of Oriental languages; compiler of *Lexicon Persico–Latinum etymologicum*.
- VYĀḌI** (from India; fl. before 500 AD?). Compiler of the *Utpalinī*, a Sanskrit lexicon, now extant only in fragments, identified by later commentators as a source for the the *Amarakośa* of AMARASIMHA.
- WAHRIG, GERHARD** (Burgstädt, Germany 1923 – Wiesbaden, West Germany [now Germany] 1978). Linguist; compiler of a popular German dictionary in which the use of computer-aided methods was pioneered.
- WALKER, JOHN** (Colney Heath, Middlesex 1732 – London 1807). Teacher of elocution; compiler of the most successful of the eighteenth-century English pronouncing dictionaries.
- WÁNG ĀNSHÍ** 王安石 (from Linchuan, China; 1021–1086). Statesman; compiler of the dictionary of script *Zishuō*, an ambitious attempt to explore the relationships between Chinese characters.
- WÁNG Lì** 王力 (Bóbái County, China 1900 – Běijīng 1986). Linguist; co-editor of the monolingual Chinese dictionary *Xiàndài Hànyǔ cídiǎn*.
- WĀNG Yí** 汪怡 (from Cháohú, China; 1875–1960). Linguist; chief editor of the monolingual Chinese dictionary *Guóyǔ cídiǎn*.
- WÁNG YĪNGDIÀN** 王應電 (from Kunshan, China; fl. 1540). Compiler of *Tóngwén bèikǎo*, a dictionary with a rhyme table.
- WARTBURG, WALTHER VON** (Riedholz, Switzerland 1888 – Basel, Switzerland 1971). Philologist; compiler with O. BLOCH of a French etymological dictionary, and editor of the monumental *Französisches etymologisches Wörterbuch*.

- WEBSTER, NOAH** (West Hartford, Connecticut 1758 – New Haven, Connecticut 1843). Writer on language; compiler of the hugely influential *American Dictionary of the English Language*.
- WEHR, HANS** (Leipzig, Germany 1909 – Münster, West Germany [now Germany] 1981). Professor of Arabic; compiler of a widely used Arabic–German dictionary, translated by J. COWAN to make an Arabic–English version.
- WÈI JIÀNGŌNG** 魏建功 (from Nántōng, China; 1901–1980). Editor-in-chief of the first edition of the important monolingual Chinese dictionary *Xīnhuá zìdiǎn*.
- WÈI JIÀO** 魏校 (from Kunshan, China; fl. 16th cent.). Scholar and official; compiler of the dictionary of script *Liúshū jīngyùn*, which draws moral conclusions from the analysis of Chinese characters.
- WELLIG, ARNOLD GOTTLIEB** (Riga, Russian empire [now in Latvia] 1778 – Riga, Russian empire [now in Latvia] 1864). Pastor; editor of a supplement to the *Lettsches Lexikon* of G. STENDER.
- WEMMERS, JACOBUS** (Antwerp, Spanish Netherlands [now in Belgium] 1598 – Naples, Spanish empire [now in Italy] 1645). Carmelite friar and titular bishop of Memphis in Egypt; compiler of the first Ge'ez–Latin dictionary.
- WERNER, REINHOLD OTTO** (Regen, Germany 1947 – Augsburg, Germany 2015). Linguist; compiler, with G. HAENSCH, of dictionaries of Americanisms in Spanish.
- WHIPPLE, AMIEL WEEKS** (Greenwich, Massachusetts 1818 – Washington, DC 1863). Military engineer and surveyor; his 'Vocabularies of North American Languages' includes the first wordlist of Western Apache.
- WHITWORTH, GEORGE CLIFFORD** (Great Budworth, Cheshire 1846 – Grasse, France 1917). Civil servant in India; compiler of *An Anglo-Indian Dictionary*.
- WIDDOWSON, JOHN DAVID ALLISON** (b. Sheffield, Yorkshire 1935). Folklorist and linguist; editor (with G. STORY and W. KIRWIN) of the *Dictionary of Newfoundland English*.
- WILKINS, CHARLES** (b. Frome, Somerset; baptized 1749; d. London 1836). Editor of a revised edition of the Persian–English dictionary of J. RICHARDSON.
- WILLEMS, ADRIAEN** (later JORIS VAN GHEEL, JORGE DE GELA) (Oevel, Spanish Netherlands [now in Belgium] 1617 – Ngongo Mbata, Kingdom of Kongo [now in the Democratic Republic of the Congo] 1652). Capuchin missionary; copyist and at one time reputed author of the Kikongo dictionary probably compiled by M. DE ROBOREDO.
- WILLIAMS, ROGER** (London c. 1606 – Rhode Island 1683). Colonist and controversialist: his *Key into the Language of America* is a pioneering phrasebook of Narragansett.
- WILLIAMS, SAMUEL WELLS** (Utica, New York 1812 – New Haven, Connecticut 1884). Missionary (American Board of Commissioners for Foreign Missions) in

- China, diplomat, and professor of Chinese; compiler of *A Tonic Dictionary of the Chinese Language in the Canton Dialect*.
- WILLIAMSON, JOHN POAGE (Lac Qui Parle, Minnesota 1835 – Greenwood, South Dakota 1917).** Presbyterian missionary to the Dakota people; compiler of an English–Dakota dictionary.
- WILTENS, CASPAR (Antwerp, Spanish Netherlands [now in Belgium] 1584 – Ambon, Dutch East Indies [now in Indonesia] 1619).** Preacher; compiler of a bidirectional wordlist of Dutch and Malay.
- WITSEN, NICOLAAS (Amsterdam 1641 – Amsterdam 1717).** Statesman and virtuoso; his encyclopedic *Noord en Oost Tartarye* includes the first wordlist of Yakut.
- WORDSWORTH, CHRISTOPHER (London 1848 – Salisbury, Wiltshire 1938).** Clergyman and historian; compiler of the dialect wordlist *Rutland Words*.
- WORMS, ERNEST (ERNST) AILRED (Bochum, German empire 1891 – Sydney, Australia 1963).** Pallotine missionary in Australia; compiler, with H. NEKES, of the large polyglot lexical compendium *Australian Languages*.
- WRIGHT, JOSEPH (Idle, West Yorkshire 1855 – Oxford 1930).** Philologist; editor of the *English Dialect Dictionary*.
- WRIGHT, THOMAS (Tenbury, Worcestershire 1810 – London 1877).** Historian and antiquary; editor of medieval glossaries and compiler of a *Dictionary of Obsolete and Provincial English*.
- WÚ JǐNGRÓNG 吴景荣 (from Píngyáng County, China; 1915–1994).** Editor in chief of the Chinese–English *Hàn–Yīng cídiǎn*.
- WÚ Yŭ 吴械 (from Shuzhou, China; c. 1100–1154).** Compiler of *Yùnbǔ*, the first dictionary to analyse ancient rhymes systematically.
- XÚ Kǎi 徐鍇 (from Guangling, China; 920–974).** Court scholar (and brother of XÚ XUÀN); author of *Shuōwén jiězì xizhuàn*, an important commentary on the dictionary of script *Shuōwén jiězì*.
- XU SHÈN 许慎 (from Hénán, China; fl. c. 100 AD).** Court scholar; compiler of the dictionary of script *Shuōwén jiězì*, one of the seminal works of Chinese lexicography.
- XÚ XUÀN 徐鉉 (from Guangling, China; 916–991).** Court scholar (and brother of XÚ Kǎi); producer of a critical edition of the dictionary of script *Shuōwén jiězì*.
- YĀDAVAPRAKĀSA (b. Tiruppuṭṭuḷi, near Kanchipuram, Chola empire [now in India]; alleged lifespan 1017–1137).** Compiler of the synonymic and homonymic Sanskrit dictionary *Vaijayantī*.
- YAMADA TADAO (Tokyo 1916 – Tokyo 1996).** Historical linguist and dictionary historian; the driving force behind the Japanese dictionary *Shin meikai kokugo jiten*.
- YÁNG XIÓNG 楊雄 (from Sìchuān, China; 53 BC–AD 18).** Compiler of the *Fāngyán*, a wordlist of forms from different varieties of Chinese and from non-Chinese languages spoken in the Chinese empire.

- YĀSKA (from India; fl. c. 500BC?).** Author of the *Nirukta*, which is both a commentary on the oldest Sanskrit lexicographical text (the *Nighaṇṭu*) and a major contribution to Sanskrit lexicography in its own right.
- AL-YAZĪDĪ (d. 237/851).** Compiler of the first extant lexicon of rare words in the Qurʾān.
- YEDIDIA OF RIMINI (YEDIDIYA BEN MOSHE OF RECANATI) (from Pesaro, Duchy of Urbino [now in Italy]; d. 1600×1602).** Teacher and translator; compiler of a Hebrew–Italian biblical glossary.
- YULE, HENRY (Inveresk, East Lothian, 1820 – London 1899).** Soldier and geographer; compiler (with A. BURNELL) of *Hobson-Jobson*, a dictionary of English as used in India.
- AL-ZABĪDĪ (Bilgram, Kanawdj [now in India] 1145/1732 – Cairo 1205/1791).** Compiler of *Tāğ al-ʾarūs*, the most comprehensive Arabic lexicon of the classical period; it was the major source of the *Arabic–English Lexicon* of E. LANE.
- AL-ZAĞĞĀĞĪ (Nahavand, Abbasid Caliphate [now in Iran] late 3rd cent./c. 860×870 – Tiberias, Seljuk empire [now in Israel] 337×340/948×950).** Grammarian; compiler of *al-Lāmāt*, a monograph investigating the Arabic letter *lām*.
- AL-ZAMAḤṢARĪ, ĀBŪ L-QĀSIM MAḤMŪD ʿUMAR (Zamaḥṣar, Khorasan [now in Turkmenistan] 467/1075 – Djurdjāniyya, Khorasan [now Kunya-Urgench, Turkmenistan] 538/1144).** Theologian and exegete; compiler of *Asās al-balāga*, the first Arabic lexicon that arranges roots in full alphabetical order, and of the Arabic–Persian dictionary *Muqaddimat al-adab*.
- ZDANOWICZ, ALEKSANDER MARIA (Igumenskij Uezd, Russian empire [now in Belarus] 1805 or 1808 – Vilnius, Russian empire [now in Lithuania] 1868).** Historian and philologist; one of the editors of the monolingual Polish dictionary known as the Vilnius dictionary.
- ZEISBERGER, DAVID (Zauchenthal, Holy Roman empire [now Suchdol nad Odrou, Czech Republic] 1721 – Goshen, Ohio 1828).** Moravian missionary in North America; compiler of an English–German–Onondaga–Delaware dictionary.
- ZENKER, JULIUS THEODOR (Ehrenfriedersdorf, Saxony [now in Germany] 1811 – Thum, German empire 1884).** Orientalist and linguist; compiler of the *Dictionnaire turc-arabe-persan*.
- ZHĀNG XUĀN 張萱 (from Boluo, China; 1550s–1630s).** Philologist and official; compiler of *Huīyǎ*, which brings together various early works in the *Ēryǎ* lexicographical tradition along with his own commentary and a supplement.
- ZHÀO YÍGUĀNG 趙宦光 (from Taicang, China; 1559–1625).** Literary figure and publisher; author of *Shuōwén chángjiān*, a work in the tradition of the dictionary of script *Shuōwén jiězi*.

ZHŌU DÉQĪNG 周德清 (from Gaoan, China; 1277–1365). Opera librettist; compiler of the rhyme dictionary *Zhōngyuán yīnyùn*.

ZHŪ MÓUWĒI 朱謀瑋 (from Nanchang, China; 1564–1624). Imperial prince; compiler of the dictionary *Piányǎ*, which explains the meaning of various literary disyllabic compounds.

ZIEGENBALG, BARTHOLOMÄUS (Pulsnitz, Electorate of Saxony [now in Germany] 1682 – Tranquebar [now Tharangambadi, India] 1719). Lutheran missionary (Dänisch-Hallesche Mission) in India and Bible translator; compiler of a Tamil dictionary.

ZINGARELLI, NICOLA (Cerignola, Italy 1860 – Milan, Italy 1935). Philologist; compiler of an Italian dictionary.

ZYZANIJ, LAVRENTIJ IVANOVYČ (LAVRENTIJ TUSTANOVSKIJ) (Tustan', Polish-Lithuanian Commonwealth [now in Ukraine] or Potylicz, Polish-Lithuanian Commonwealth [now Potelyč, Ukraine], c. 1570 – possibly Korec', Polish-Lithuanian Commonwealth [now in Ukraine], after 1634). Archpriest; compiler of a wordlist containing Church Slavic words and their equivalents in the East Slavic vernacular.

Primary Sources

- Aasen, I. A. *Norsk ordbog met dansk forklaring* (Oslo: Mallings, 1873).
- Aasen, I. A. *Ordbog over det norske folkesprog* (Oslo: Kongelige Norske Videnskabs-Selskab, 1850).
- Abel, I. *Symphona symphona sive undecim linguarum orientalium discors exhibita concordia Tamulicae, videlicet Granthamicae, Telugicae, Samscrutamicae, Marathicae, Balabandicae, Canaricae, Hindostanicae, Cuncanicae, Gutzaratizae et Peguanicae* (Copenhagen: Hallagerus, 1782).
- Abdel-Nour, J., and S. Idriss. *Al Manhal: Dictionnaire français-arabe* (Beirut: Dar Al-'Adāb, 1979).
- Abramović, T. *Německij i serbskij slovar' / Deutsch und Illyrisches Wörterbuch*, 2 vols. (Vienna: von Kurzbeck, 1791).
- Abū Miṣḥāl, *Nawādir* = Abū Miṣḥāl 'Abdallāh b. Ḥarīš al-A' rābī, *al-Nawādir*, ed. 'I. Ḥasan, 2 vols. (Damascus: Mağma' al-Luġa al-'Arabiyya, 1961).
- Abū l-Ṭayyib, *Ibdāl* = Abū l-Ṭayyib 'Abd alrahmān b. 'Alī al-Luġawī, *al-Ibdāl*, ed. 'I. al-Dīn al-Tanūḥī, 2 vols. (Damascus: al-Mağma' al-'Ilmī al-'Arabī bi-Dimašq, 1960–1).
- Abū 'Ubayd, *Amṭāl* = Abū 'Ubayd al-Qāsim b. Sallām al-Harawī, *al-Amṭāl*, ed. 'A. Qaṭāmiš (Damascus: Dār al-Ma'mūn, 1980).
- Abū 'Ubayd, *Ġarīb* = Abū 'Ubayd al-Qāsim b. Sallām al-Harawī, *al-Ġarīb al-muṣannaf*, ed. M. al-Muḥṭār al-'Ubaydī, 3 vols. (Qartāğ: Bayt al-Hikma, 1989–96).
- Abū Zayd, *Nawādir* = Abū Zayd Sa'īd b. Aws al-Anṣārī, *al-Nawādir fī l-luġa*, ed. M. 'A. Aḥmad (Beirut: Dār al-Šurūq, 1981).
- Addy, S. O. *Sheffield Words* (London: English Dialect Society, 1888).
- Adelung, J. C. *Grammatisch-kritisches Wörterbuch der hochdeutschen Mundart*, 4 vols. (Leipzig: Breitkopf, 1793–1801).
- Aggavaṃsa. *Saddanīti*, ed. H. Smith, 3 parts = Acta Regiae Societatis Humaniorum Litterarum Lundensis 12.1–3 (Lund: Gleerup, 1928–30).
- Aitareya Brāhmaṇa = *The Aitareya Brahmanam of the Rigveda*, vol. I, *Sanscrit Text*, ed. M. Haug (Bombay: Government Central Book Depot; London: Trübner, 1863).
- Aitareya Upaniṣad = *The Upanishads*, vol. V, *The Aitareya and Taittiri'ya Upanishads and Sri Sankara's Commentary*, trans. S. S. Sastri (Madras: Sheshacharri, 1901).
- Ajayapāla. *Nānārthasaṃgraha*, ed. T. R. Chintamani Dikshit (Chennai: Madrapuṇyaviśvavidyālayaḥ, 1937).
- Akbar, M. 'A. A. A. *Muṣṭalahāt-i ṭhagī* (Kolkata: Block, 1839).

- Aktunç, H. *Türkçenin Büyük Argo Sözlüğü (Tanımlarıyla)* (Istanbul: Yapı Kredi Yayınları, 1998).
- Albert, R., and D. L. Shaul. *A Concise Hopi and English Lexicon* (Amsterdam: Benjamins, 1985).
- Alcalá, P. de. *Vocabulista arauigo en letra castellana* (Granada: J. Varela de Salamanca, 1505).
- Alcalay, R. *Milon ivri–angli shalem/The Complete Hebrew–English Dictionary*, 4 vols. (Tel Aviv: Massada, 1963).
- Alcalay, R. *Milon ivri–angli shalem/The Complete Hebrew–English Dictionary*, new enlarged edn, 2 vols. (Givatayim: Massada, 1990).
- Alcedo, A. de. ‘Vocabulario de voces provinciales de la América usadas en el Diccionario geográfico-histórico de ella; y de los nombres propios de plantas, aves y animales’ in Alcedo, *Diccionario geográfico-histórico de las Indias Occidentales ó América* (Madrid: Gonzalez, 1786–9), 5, separately paginated appendix.
- Alcover, A. M., and F. de B. Moll. *Diccionari català–valencià–balear*, 10 vols. (Palma de Mallorca: Miramar, 1930–62), dcvb.iecat.net.
- Aldus Manutius: *The Greek Classics*, ed. and trans. N. G. Wilson (Cambridge, MA, and London: Harvard University Press, 2016).
- Alévêque, C. *Petit dictionnaire français–coréen* (Seoul: Seoul Press, 1901).
- Alfāsi, David ben Abraham, *Jāmi^c al-‘alfāz = The Hebrew–Arabic Dictionary of the Bible Known as Jāmi^c al-‘Alfāz (Agron)*, ed. S. L. Skoss, 2 vols. (New Haven: Yale University Press, 1936–45).
- Allsopp, R. *Dictionary of Caribbean English Usage* (Oxford University Press, 1996).
- Das älteste deutsche Buch: Die ‘Abrogans’-Handschrift der Stiftsbibliothek St. Gallen*, ed. B. Bischoff, J. Duft, and S. Sonderegger, 2 vols. (St Gallen: Zollikofer Fachverlag, 1977).
- Alvarado, F. de. *Vocabulario en lengua mixteca*, facs. ed. W. Jiménez Moreno (Mexico City: Instituto Nacional Indigenista and Instituto Nacional de Antropología e Historia, 1962).
- Alves, A. *Dicionário Etimológico Bundo–Português*, 2 vols. (Lisbon: Tipografia Silvas/Mensário Administrativo, 1951).
- Amarasiṃha. *Amarakośa* (1913) = *The Nāmalingānuśāśāna (Amarakosha) of Amarasiṃha: with the commentary (Amarakoshadghatana) of Kshirasvāmin*, ed. K. G. Oka (Poona: Law Printing Press, 1913).
- Amarasiṃha. *Amarakośa* (1971–83) = *Amarakośa, with the Unpublished South Indian Commentaries Amarapadavivṛti of Liṅgayasūrin, Amarapadapārijāta of Mallinātha, and the Amarapadavivarāṇa of Appayārya*, ed. A. A. Ramanathan, 3 vols. (Chennai: Adyar Library and Research Centre, 1971–83).
- Amarasiṃha. *Cōsha; or, Dictionary of the Sanskrit Language*, trans. H. T. Colebrooke (Serampore: n.p., 1808).
- American Heritage Dictionary of the English Language*, ed. W. Morris (Boston: Houghton Mifflin; [New York]: American Heritage, 1969).
- Ammonius (attrib.). *Liber de adfinium vocabulorum differentia*, ed. K. Nickau (Leipzig: Teubner, 1966).
- Anchieta, Joseph de. *Arte de gramática da língua mais usada na costa do Brasil* (Coimbra: Mariz, 1595).
- AND₁ = *The Australian National Dictionary*, ed. W. S. Ramson et al. (Melbourne: Oxford University Press, 1988).
- AND₂ = *The Australian National Dictionary*, 2nd edn, ed. B. Moore, A. Laugesen, et al., 2 vols. (Melbourne: Oxford University Press, 2016).

- Ange de Saint Joseph. *Gazophylacium linguae Persarum* (Amsterdam: ex officina Jansonio-Waesbergiana, 1684).
- Āṅguttaranikāya = *The Āṅguttara-nikāya*, part I, *Ekanipāta*, *Dukanipāta*, and *Tikanipāta*, ed. R. Morris, 2nd edn, revised A. K. Warder (London: Pali Text Society, 1961).
- [Anshel]. *Mirkevet ha-Mishneh* (Kraków: Halicz, 1534 or 1535).
- Antoine, F., et al. *Central Carrier Bilingual Dictionary* (Fort James, BC: Carrier Linguistic Committee, 1974).
- Apollonius Sophista. *Lexicon Homericum*, ed. I. Bekker (Berlin: Reimer, 1833).
- Ara, D. de. *Vocabulario de lengua tzeldal segun el orden de Copanabastla*, ed. M. H. Ruz (Mexico City: Universidad Nacional Autónoma de México, 1986).
- Aragona, A. de. 'Breve introducción para aprender la lengua guaraní', ed. B. Melià, *Amérindia*, 4 (1979), 23–61.
- Arenas, P. de. *Vocabulario manual de las lenguas castellana y mexicana, en que se contienen las palabras, preguntas, y respuestas mas co[m]unes, y ordinarias que se suelen ofrecer en el trato y comunicacion entre Españoles, é Indios* (Mexico City: Martínez, 1611).
- Aristophanes of Byzantium. *Fragmenta*, ed. A. Nauck (Halle: Lippert and Schmidt, 1848).
- Aristophanes of Byzantium. *Fragmenta*, ed. W. J. Slater (Berlin: De Gruyter, 1986).
- Arroyo de la Cuesta, F. A *Vocabulary or Phrase Book of the Mutsun Language of Alta California* (New York: Cramoisy, 1862).
- Arte de la lengua cahita*, ed. E. Buelna (Mexico City: Imprenta del Gobierno Federal, 1890; facs. reprint 1989; 1737).
- Arte y vocabulario de la lengua chiquita*, ed. L. Adam and V. Henry (Paris: Maisonneuve, 1880).
- Arte y vocabulario de la lengua dohema, heve o eudeva*, ed. C. W. Pennington (Mexico City: Universidad Nacional Autonoma de Mexico, 1982).
- Arte y vocabulario de la lengua Morocosi* (Madrid: n.p., 1699).
- Arte y vocabulario en la lengua general del Peru, llamada quichua, y en la lengua española* (Lima: Ricardo, 1586).
- Arts, T. (ed.). *Oxford Arabic Dictionary* (Oxford University Press, 2014), premium.oxforddictionaries.com/arabic/.
- Ārzū, Sirāj al-Dīn 'Alī Khān. *Navādir al-Alfāz*, ed. S. 'Abdullah (Karachi: Anjuman-e Taraqqi-ye Urdu, 1951).
- Asamīyā Baṛo Śabdakosha*, 2nd edn (Guwahati: Pūrbottara Bhāshā Samiti, 1987).
- 'Askarī, *Furūq* = Abū Hilāl al-Ḥasan b. 'Abdallāh al-'Askarī, *al-Furūq al-luġawiyya*, ed. Ḥ. al-Dīn al-Qudṣī (Cairo: Maktabat al-Qudṣī, 1353 AH).
- Aslanov, M. G. *Afgansko-russkij slovar' (puštu)* (Moscow: Sovetskaja encyklopedija, 1966).
- Aṣma'ī, Ḥayl = Abū Sa'īd 'Abdalmalik b. Qurayb al-Aṣma'ī, *al-Ḥayl*, ed. H. Nāḡī, *al-Mawrid* 12.4 (1983), 177–222.
- Assumpçam, M. da. *Vocabulario em idioma Bengalla e Portuguez* (Lisbon: da Silva, 1743).
- Atharvaveda = Atharva-veda Samhitā*, trans. and ed. W. D. Whitney, and ed. C. R. Lanman, 2 vols. (Cambridge, MA: Harvard University Press, 1905).
- Aubery, J. 'Dictionnaire abnaquis-francois' (manuscript in Musée des Abénakis, Odanak, Quebec).
- Aubery, J. *French Abenaki Dictionary*, ed. and trans. S. Laurent (Portland, ME: Chisholm Bros., 1995).
- Aulete, F. J. de C. *Diccionario contemporaneo da lingua portugueza*, 2 vols. (Lisbon: Imprensa Nacional, 1881;³1948–52).

- Austin, C. (ed.). *Comicorum Graecorum fragmenta in papyris reperta* (Berlin and New York: De Gruyter, 1973).
- Austin, P. K., and D. Nathan. *Kamilaroi / Gamilaraay Web Dictionary* (1998), dnathan.com.
- Avineri, Y. (ed.). *Hekhal Rashi: mekhil bə-seder alef bet et ḥidduše Rashi və-tserufav ha-ləšoniyyim, perušav lə-millim ul-vittuyim, hašva'otav ha-ləšoniyyot, kəlalav / məluqqat mitokh Peruš Rashi la-Miqra ul-Talmud Bavli, və-khen mi-təšuvotav u-slihotav 'al pi hašva'ah šel 'esrim va-ḥamiššah dəfusim, 'arukh, məvo'ar u-muggah*, 2 vols. (Jerusalem: Mosad ha-Rav Kuk, 1979–85).
- Avneyon, E. *Milon sapir: milon ivri-ivri entsiklopedi beshitat hahove* (Tel Aviv: Eitav, 1998; ²2002).
- Azar, M. (ed.). *Milon Even-Shoshan*, 6 vols. (Tel Aviv: Hamilon Hekhadash, 2003).
- Azhari, *Tahdīb* = Abū Mansūr Muḥammad b. Aḥmad al-Azhari, *Tahdīb al-luġa*, ed. 'A. Hārūn et al., 15 vols. (Cairo: al-Mu'assasa al-Miṣriyya al-'Āmma li-l-Ta'lif wa-l-Anbā' wa-l-Naṣr and al-Dār al-Miṣriyya li-l-Ta'lif wa-l-Taḡama, 1964–7).
- Baalbaki, M. *Al-Mawrid: A Modern English–Arabic Dictionary* (Beirut: Dār Al-'ilm li-l-malāyīn, 1967).
- Baalbaki, M. *Al-Mawrid al-Akbar: A Modern English–Arabic Dictionary*, completed by Ramzi Baalbaki (Beirut: Dār Al-'ilm li-l-malāyīn, 2005).
- Baalbaki, Ruhi. *Al-Mawrid Al-Mar'iyy* (Beirut: Dār Al-'ilm li-l-malāyīn, 2002).
- Baalbaki, Ruhi. *Al-Mawrid: Qāmūs 'arabiyy-inklīzī* (Beirut: Dār Al-'ilm li-l-malāyīn, 1987).
- Baalbaki, Ruhi, and M. Baalbaki. *Al Mawrid Al-Wasīt* (Beirut: Dār Al-'ilm li-l-malāyīn, 1990).
- Bacher, W. 'Ein hebräisch-persisches Wörterbuch aus dem vierzehnten Jahrhundert', supplement to *Jahresbericht der Landes-Rabbinerschule für das Schuljahr 1899/1900* (Budapest: Landes-Rabbinerschule, 1900).
- Bahat, S., and M. Mishor. *Milon hahove* (Tel Aviv: Eitav, 1995).
- Bailey, N. *Dictionarium Britannicum* (London: Cox, 1730).
- Balbi, Giovanni. *Catholicon* (Mainz: [Gutenberg?], 1460; facs. edn 1971).
- Baldinger, K., et al. (eds.). *Dictionnaire étymologique de l'ancien français*, multiple fascicles plus online continuation (Tübingen: Niemeyer, 1974–), www.deaf-page.de.
- Bān Gù. *Hànshū* 漢書 (Beijing: Zhōnghuá shūjù, 1992).
- Bandaniġi, *Taqfiya* = Abū Biṣr al-Yamān b. al-Yamān al-Bandaniġi, *al-Taqfiya fī l-luġa*, ed. Ḥ. I. al-'Aṭiyya (Baghdad: Maṭba'at al-'Ānī, 1976).
- Baraga, F. *A Dictionary of the Otchipwe Language, Explained in English*, new edn [by A. Lacombe], 2 vols. (Montreal: Beauchemin & Valois, 1878–80; reprint edn 1966, 1973; ¹1853).
- Barbier de Meynard, C. A. C. *Dictionnaire turc-français: supplément aux dictionnaires publiés jusqu'à ce jour*, 2 vols. (Paris: Leroux, 1881–6).
- Bardhi, F. *Dictionarium Latino-Epiroticum* (Rome: Congregatio de Propaganda Fide, 1635).
- Barkhudarov, A. S., V. M. Beskrovny, G. A. Zograf, and V. P. Liperovsky. *Khindi-russkii slovar' v dvukh tomakh: okolo 75 000 slov*, 2 vols. (Moscow: Sovetskaja enciklopediya, 1972).
- Bartlett, J. R. *Dictionary of Americanisms: Glossary of Words and Phrases Usually Regarded as Peculiar to the United States* (New York: Bartlett and Welford, 1848).
- Barua, G. C. *Ahom–Assamese–English Dictionary* (Kolkata: Baptist Mission, 1920).

- Barua [Baruwā], H. *Hema Kosha; or, An Etymological Dictionary of the Assamese Language*, ed. P. R. T. Gurdon and H. Gosain (Kolkata: Baptist Mission, 1900).
- Basalengu, D. *Arte y vocabulario de la lengua matlaltzinga vuelto a la castellana*, ed. M. E. Bribiesca (Mexico City: Biblioteca Enciclopédica del Estado de México, 1975).
- Basalengu, D. *Vocabulario de la lengua castellana vuelto en la matlaltzinga*, ed. L. Manrique (Mexico City: Biblioteca Enciclopédica del Estado de México, 1975).
- Baṭalyawsī, Muṭallat = Abū Muḥammad ‘Abdallāh b. Muḥammad b. al-Sīd al-Baṭalyawsī, *al-Muṭallat*, ed. Ṣ. M. al-Farṭūsī, 2 vols. (Baghdad: Dār al-Raṣīd, 1981–2).
- Bathe, W. *Janua linguarum* (Salamanca: de Cea Tesa, 1611).
- Battaglia, S. (ed.). *Grande dizionario della lingua italiana*, 21 vols. (Turin: UTET, 1961–2002).
- Bedersi, Abraham ben Isaac. *Ḥotam Tokhnit/Chotam Tochnit: Hebraeische Synonymik*, ed. G. I. Polak (Amsterdam: Levisson, 1865).
- Beesley, K. R., and D. Elzinga (eds.). *An 1860 English–Hopi Vocabulary Written in the Deseret Alphabet* (Salt Lake City: University of Utah Press, 2015).
- Bélisle, L.-A. *Dictionnaire général de la langue française au Canada* (Quebec City: Bélisle, 1957; first issued in fascicles 1954–7).
- Bell, H. I., and W. E. Crum (eds.). ‘A Greek–Coptic glossary’, *Aegyptus*, 6 (1925), 177–226.
- Bell, H. I., and H. Thompson (eds.). ‘A Greek–Coptic glossary to Hosea and Amos’, *Journal of Egyptian Archaeology*, 11 (1925), 241–6.
- Belostenec, I. *Gazophylacium, seu Latino-Illyricorum onomatum aerarium* (Zagreb: Weitz, 1740).
- Belsare, M. B. *The Pronouncing and Etymological Gujarati–English Dictionary* (Ahmedabad: Pathak, 1895).
- Benešovský, M. *Knjžka slov českých* (Prague: n.p., 1587).
- Bentley, W. H. *Dictionary and Grammar of the Kongo Language, as Spoken at San Salvador, the Ancient Capital of the Old Kongo Empire, West Africa* (London: Baptist Missionary Society and Trübner, 1887).
- Bentley, W. H. *Appendix to the Dictionary and Grammar of the Kongo Language, as Spoken at San Salvador, the Ancient Capital of the Old Kongo Empire, West Africa* (London: Baptist Missionary Society and Kegan Paul, Trench, Trübner, 1895).
- Ben-Yehuda, E. *Milon halashon haivrit hayeshana vehakhadasha/Thesaurus totius Hebraeitis et veteris et recentioris/Gesamtwörterbuch der alt- und neuhebräischen Sprache/Dictionnaire complet de la langue hébraïque ancienne et moderne/A Complete Dictionary of Ancient and Modern Hebrew*, 16 vols. plus Prolegomena vol. (Berlin: Langenscheidt (vols. I–VIII); Jerusalem: Benyehuda (vols. IX–XVI), 1908–59).
- Ben-Yehuda, E. *Milon ivri meturgam ashkenazit hameduberet beyn hayehudim verusit* (Vilnius: Pirozhnikov, 1901).
- Bernolák, A. *Slowár Slowenski Česko–Latínsko–Nemecko–Uherski seu Lexicon Slavicum Bohemico–Latino–Germanico–Ungaricum*, 6 vols. (Buda: Typographia Regiae Universitatis Hungaricae, 1825–7).
- Bertonio, L. *Arte breve de la lengua aymara, para introduction del Arte grande de la misma lengua* (Rome: Zannetti, 1603).
- Bertonio, L. *Arte de la lengua aymara, con una silva de phrases de la misma lengua, y su declaracion en Romance* (Juli, Peru: del Canto, 1612).
- Bertonio, L. *Arte y grammatica muy copiosa de la lengua aymara* (Rome: Zannetti, 1603).
- Bertonio, L. *Vocabulario de la lengua aymara* (Juli, Peru: del Canto, 1612).

- Berynda, P. *Leksikon "slavenorossijskij i imen" tl"kovanie* (Kiev: Tip. Pečerskoj Lavry, 1627).
- Bescherelle, L.-N. *Dictionnaire national; ou, Grand dictionnaire classique de la langue française*, 2 vols. in 4 parts (Paris: Simon, 1845–7).
- Beschi, C. G. *Grammatica Latino–Tamulica* (Tharangambadi: typis Missionis Danicae, 1738).
- Beschi, C. G. *Īktu Vīramāmuṇivar ceyta Caturakarāti Ceṇṇaica caṇkattil accuppatittatu* (n.p.: n.p., n.d., but 1824 or later), Pondicherry, EFEO Library, shelfmark EO DIC TA 5.
- Beschi, C. G. *Vīramāmuṇivar aruḷiya Caturakarāti (Āyvuṇ patippu)*, ed. Cū. Inṇāci (Chennai: Mātā patippakam, 1989; 1979).
- Beskrovny, V. M. *Khindi–russkij slovar'* (Moscow: Gos. izd-vo inostrannykh i nacional'nykh slovarej, 1953).
- Bethel, R., P. V. Kroskrity, C. Loether, and G. A. Reinhardt. *A Practical Dictionary of Western Mono* (North Fork, CA: Sierra Mono Museum, 1984).
- Bhagavatī = Aṅgasuttāni*, vol. II, *Bhagavatī*, ed. M. Nathamal (Ladnun: Jain Vishva Bharati, 1974).
- Bhānuji Dīkṣita, *Vyākhyāsudhā = The Nāmalingānuśāsana (Amarakosha) of Amarasimha, with the Commentary (Vyākhyāsudhā or Ramāśramī [sic]) of Bhānuji Dīkṣhit*, ed. Śivadatta, 5th edn, rev. by W. L. Paṇṣīkar (Mumbai: Nirṇaya-Sāgar, 1929).
- Bibbesworth, Walter de. *Le tretiz*, ed. W. Rothwell (London: Anglo-Norman Text Society, 1990), www.anglo-norman.net.
- Bibliotheca Lusitana historica, critica, e cronologica*, ed. D. Barbosa Machado, 4 vols. (Lisbon: da Fonseca, 1741–59).
- Bird, W. 'A short vocabulary of the Chowie-language of the Buccaneer Islanders (Sunday Islanders), north western Australia', *Anthropos*, 10 (1915), 180–6.
- Bleek, W. H. I. *The Languages of Mosambique: Vocabularies of the Dialects of Lourenço Marques, Inhambane, Sofala, Tette, Sena, Quelimane, Mosambique, Cape Delgado, Anjoane, The Maravi, Mudsau, &c.* (London: Harrison, 1856).
- Bloch, O., and W. von Wartburg. *Dictionnaire étymologique de la langue française* (Paris: PUF, 1932; 1968).
- Bluteau, R. *Vocabulário português e latino*, 10 vols. (Coimbra: Colegio das Artes da Companhia da Jesu [imprint varies], 1712–28).
- 'Bocabulario de la lengua sangleya por las letraz de el A.B.C.' (London: British Library, Add. MS 25317).
- Bohorič, A. *Arcticae horulae succisivae de Latinocarniolana literatura ad Latinae linguae analogiam accomodata* (Wittenberg: n.p., 1584).
- Böhtlingk, O. N. von. *Sanskrit-Wörterbuch in kürzerer Fassung*, 7 vols. in 3 (Delhi: Banarsidass, 1991; 1879–89).
- Böhtlingk, O. N. von, and Roth, R. von. *Sanskrit-Wörterbuch*, 7 vols. (St Petersburg: Kaiserliche Akademie der Wissenschaften, 1855–75).
- Bollée, A. (ed.). *Dictionnaire étymologique des créoles français de l'Océan Indien*, 2 vols. in 4 parts (Hamburg: Buske, 1993–2007).
- Bolton, W. F. (ed.). *The English Language: Essays by English and American Men of Letters*, vol. I, 1490–1839 (Cambridge University Press, 1966).
- Bonelli, L. *Lessico turco-italiano* (Rome: Istituto per l'Oriente, 1939).
- Bonet, J. *Dictionnaire annamite-français (langue officielle et langue vulgaire)*, 2 vols. (Paris: Imprimerie Nationale, 1899–1901).

- Boucher, J. *Boucher's Glossary of Archaic and Provincial Words: A Supplement to the Dictionaries of the English Language, Particularly those of Dr. Johnson and Dr. Webster*, ed. J. Hunter and J. Stevenson (London: Black, Young, and Young, 1832).
- Bourke, J. G. *Vocabulary of the Apache or 'Inde' Language of Arizona & New Mexico*, ed. C. J. Condie = Occasional Publications in Anthropology: Linguistics Series, 7 (Greeley, CO: Museum of Anthropology, University of Northern Colorado, 1980).
- Bowrey, T. *A Dictionary: English and Malayo, Malayo and English* (London: for the author, 1701).
- Brambila, D. *Diccionario castellano-raramuri* (Mexico City: Obra Nacional de la Buena Prensa, 1983).
- Brambila, D. *Diccionario raramuri-castellano (tarahumara)* (Mexico City: Editorial Buena Prensa, 1960).
- Bray, D. (ed.), in collaboration with the White Mountain Apache Tribe. *Western Apache-English Dictionary: A Community-Generated Bilingual Dictionary* (Tempe, AZ: Bilingual Press/Editorial Bilingüe, 1998).
- Breton, P. *A Vocabulary of the Names of the Various Parts of the Human Body and of Medical and Technical Terms in English, Arabic, Persian, Hindee and Sanscrit, for the Use of the Members of the Medical Department in India* (Kolkata: Government Lithographic Press, 1825).
- Breton, R. *Dictionnaire caraïbe-françois, meslé de quantité de remarques historiques pour l'esclaircissement de la langue* (Auxerre: Bouquet, 1665).
- Brhatkalpabhāṣya* (Bhavnagar: Jain Atmananda Sabha, 1936).
- Briggs, K., and K. Kilpatrick. *A Dictionary of Suffolk Place-Names* (Nottingham: English Place-Name Society, 2016).
- Bright, W. *Native American Placenames of the United States* (Norman: University of Oklahoma Press, 2004).
- Brinton, D., and A. S. Anthony (eds.). *A Lenape-English Dictionary, From an Anonymous MS. in the Archives of the Moravian Church at Bethlehem, PA* (Philadelphia: Historical Society of Pennsylvania, 1888).
- Brito, Guillelmus. *Summa Britonis, sive, Guillelmi Britonis Expositiones vocabulorum Biblie*, ed. L. W. Daly and B. A. Daly (Padua: in aedibus Antenoreis, 1975).
- Brockhaus-Wahrig: Deutsches Wörterbuch*, ed. G. Wahrig, 6 vols. (Wiesbaden: Brockhaus, 1980-4).
- Brodowski, J. *Lexicon Germanico-Lithvanicum et Lithvanico-Germanicum*, ed. J. Drotvinas (Vilnius: Lietuvių kalbos institutas, 2009).
- Bronson, M. *A Dictionary in Assamese and English* (Sivasagar: American Baptist Mission, 1867).
- Bronson, M. *A Spelling Book and Vocabulary, in English, A'sa'mese, Singpho, and Na'ga* (Jaipur: American Baptist Mission, 1839).
- Brooks, M., and J. Ritchie. *Words from the West: A Glossary of Western Australian Terms* (Melbourne: Oxford University Press, 1994).
- Brown, C. P. *A Dictionary of the Mixed Dialects and Foreign Words Used in Telugu, with an Explanation of the Telugu Alphabet* (Chennai: Christian Society, 1854).
- Bruyas, J. *Radices verborum Iroquaeorum/Radical Words of the Mohawk Language, with their Derivatives* = Shea's Library of American Linguistics, 10 (New York: Cramoisy, 1862).
- Bryant, A. T. *A Zulu-English Dictionary, with Notes on Pronunciation, a Revised Orthography, and Derivations and Cognate Words from Many Languages; Including also a Vocabulary of*

- Hlonipa words, Tribal-names, etc.; a Synopsis of Zulu Grammar, and a Concise History of the Zulu People (Pinetown, Natal: Mariannhill Mission, 1905).
- Buchi, É., and W. Schweickard (eds.). *Dictionnaire étymologique roman* (Nancy: ATILF, 2008–), www.atilf.fr/derom.
- Buckwalter, T., and D. Parkinson. *A Frequency Dictionary of Arabic: Core Vocabulary for Learners* (New York: Routledge, 2011).
- Buddhaghosa. *The Atthasālinī: Buddhaghosa's Commentary on the Dhammasaṅgaṇi*, ed. E. Müller (London: Pali Text Society, 1979; ¹1897).
- Buddhaghosa. *The Expositor-Atthasālinī: Buddhaghosa's Commentary on the Dhammasaṅgaṇi, the First Book of the Abhidhamma Piṭaka*, trans. M. Tin, ed. C. A. F. R. Davids, 2 vols. (London: Pali Text Society, 1920–1).
- Buddhaghosa. *Jātakatṭhakaṭṭhā = The Jātaka Together With Its Commentary, Being Tales of the Anterior Births of Gotama Buddha*, ed. V. Fausbøll, 7 vols. (London: Pali Text Society, 1962–4; ¹1877–97).
- Buddhaghosa. *Sumaṅgalavilāsinī = Buddhaghosa. Sumaṅgala-vilāsinī: Buddhaghosa's Commentary on the Dīgha Nikāya*, ed. T. W. R. Davids and J. E. Carpenter, 3 vols. (London: Pali Text Society, 1968–71; ¹1886–1932).
- A Buddhist Manual of Psychological Ethics of the Fourth Century BC: Being A Translation from the Original Pali of the First Book in the Abhidhamma Piṭaka Entitled Dhamma-saṅgaṇi (Compendium of States or Phenomena)*, trans. C. A. F. R. Davids (London: Royal Asiatic Society, 1900).
- Budé, G. *Commentarii linguae Graecae* (Paris: Badius Ascensius, 1529).
- Buechel, E. *A Dictionary of the Teton Dakota Sioux Language, Lakota–English, English–Lakota, with Considerations Given to Yankton and Santee*, ed. P. Manhart (Pine Ridge, SD: Red Cloud Indian School, 1970).
- Buechel, E., and P. Manhart. *Lakota Dictionary: Lakota–English, English–Lakota*, new edn (Lincoln, NE: University of Nebraska Press, 2002).
- Būga, K. *Lietuvių kalbos žodynas: i sąsiuvinis* (Kaunas: Švietimo ministerija, 1924).
- Bugarin, J. *Diccionario Ibanag–Español*, ed. A. Lobato de Santo Tomás, rev. by R. Rodriguez (Manila: Amigos del País, 1854).
- Burhān-i Qāṭi'*, ed. M. Mu'in, 4 vols. (Tehran: Zavvar, 1951–6).
- Busbecq, O. de. *Legationis Turcicae epistolae quatuor* (Paris: Beys, 1589).
- al-Bustānī, B. *Muḥīṭ al-Muḥīṭ* (Beirut: n.p., 1867–70).
- Buxtorf, J. *Lexicon chaldaicum, talmudicum, et rabbinicum* (Basel: König, 1639–40).
- Buxtorf, J. *Lexicon hebraicum et chaldaicum* (Basel: König, 1615).
- Büyük Türkçe Sözlük, tdk.gov.tr/index.php?option=com_bts&view=bts.
- Bydgoszczy, B. z. *Słownik Bartłomieja z Bydgoszczy: Wersja łacińsko–polska*, ed. E. Kędelska, I. Kwileczka, and A. Łuczak, 5 vols. (Warsaw: Slawistyczny Ośrodek Wydawniczy, 1999–2012).
- Bydgoszczy, B. z. *Słownik łacińsko–polski Bartłomieja z Bydgoszczy*, ed. B. Erzepki (Poznań: Drukarnia Dziennika Poznańskiego, 1900).
- Byington, C. *A Dictionary of the Choctaw Language*, ed. J. R. Swanton and H. S. Halbert = Smithsonian Institution, Bureau of American Ethnology, Bulletin 46 (Washington, DC: Government Printing Office, 1915; reprint 1973 and 1978).
- Calepino, A. *Dictionarium* (Reggio dell'Emilia: Bertochus, 1502).
- Callary, E. *Place Names of Illinois* (Urbana: University of Illinois Press, 2008).

- Callary, E. *Place Names of Wisconsin* (Madison: University of Wisconsin Press, 2016).
- Callimachus. *Fragmenta*, ed. R. Pfeiffer (Oxford: Clarendon, 1949).
- Campbell, W. *A Dictionary of the Amoy Vernacular, Spoken Throughout the Prefectures of Chinchiu, Chiang-chiu, and Formosa* (Tainan: Taiwan Church, 1913).
- Canadian Oxford Dictionary*, ed. K. Barber (Don Mills: Oxford University Press, 1998).
- Candrea, I. A., and O. Densușianu. *Dicționarul etimologic al limbii române: Elementele latine*, 4 fascicles (a–putea) (Bucharest: Socec, 1907–14).
- Canicatti, B. M. de [as Cannecattim]. ‘Diccionario abbreviado da lingua Congueza, a que accresce huma quarta columna, que contém os termos da lingua Bunda, identicos, ou semelhantes à Lingua Congueza’, in Canicatti, *Collecção de observações grammaticaes sobre a lingua Bunda ou Angolense* (Lisbon: Impressão Régia, 1805), 149–216.
- Canicatti, B. M. de [as Cannecattim]. *Diccionario da lingua Bunda ou Angolense, explicada na Portugueza, e Latina* (Lisbon: Impressão Régia, 1804).
- Cantemir, D. *Istoria hieroglifică*, 2 vols. (Bucharest: Editura pentru literatură, 1965).
- Cardoso, J. *Dictionarium Latinolusitanicum et vice versa Lusitanicolatinum*, 2 parts in 1 vol. (Coimbra: de Barreira, 1569–70).
- Carradori, A. *Il dizionario turco-ottomano*, ed. L. Rocchi (Edizioni Università di Trieste, 2011).
- Cassidy, F. G. *Dane County Place Names*, 2nd edn (Madison: University of Wisconsin Press, 2009; ¹1947).
- Cassidy, F. G., and R. B. Le Page. *Dictionary of Jamaican English* (Cambridge University Press, 1967; ²1980).
- Cauzim, M. M. *A Dictionary, Goojratee and English*, ed. N. Furdoonjee (Mumbai: Courier, 1846).
- Cavigneaux, A., H. G. Güterbock, M. T. Roth, and G. Farber (eds.). *The Series Erim-Huš = Anantu and An-Ta-Gál = Šaqû = Materials for the Sumerian Lexicon*, 17 (Rome: Pontificium Institutum Biblicum, 1985).
- Cawdrey, R. *A Table Alphabeticall, Conteyning and Teaching the Understanding of Hard Vsual English Wordes* (London: Weaver, 1604).
- Chaliha, D. *Candrakānta abhidhāna: asamīyā śabdara mūla, pada āru udāharanere saite asamīyā-ṁgrājī dui bhāṣāra artha thakā br̥hat abhidhāna/A Comprehensive Dictionary of the Assamese Language with Etymology and Illustrations of Words with their Meanings both in Assamese and English* (Jorhat: Candrakānta Handikai Hal, 1933; rev. edn 1962).
- Changuion, A. N. E. *Proeve van Kaapsch Taaleigen*, separately paginated supplement to Changuion, *De nederduitsche taal in Zuid-Afrika hersteld* (Cape Town: Collard, 1844).
- Chao, E. (ed.). *Diccionario enciclopédico de la lengua española*, 2 vols. (Madrid: Gaspar y Roig, 1853–5).
- Chén Chún. *Neo-Confucian Terms Explained*, trans. W. T. Chan (New York: Columbia University Press, 1986).
- Chén Péngnián. *Sòng běn Yùpiān 宋本玉篇* (Beijing: Zhōngguó shūdiàn, 1983).
- Chen Xiang. *Han yu Wuerdu yu fen lei ci dian* (Beijing: Shang wu yin shu guan, 2013).
- Chimhundu, H. *Duramazwi Guru reChiShona* (Harare: College, 2001).
- Chimhundu, H. *Duramazwi reChiShona* (Harare: College, 1996).
- Chirino, P. ‘*Dictionarium Sino Hispanicum*’ (Rome: Biblioteca Angelica, ms. 60).
- Choeroboscus, George. *Epimerismi in Psalmos*, ed. T. Gaisford (Oxford: Typographeus Academicus, 1842).

- Cholinus, P., and J. Frisius. *Dictionarium Latino–Germanicum* (Zürich: Froschauer, 1541).
- Choueka, Y. *Rav-Milim: hamilon hashalem laivrit hakhadasha* (Tel Aviv: Center for Educational Technology, 1996), www.ravmilim.co.il/.
- Cigale, M. *Deutsch–slovenisches Wörterbuch*, 2 vols. (Ljubljana: Wolf, 1860).
- Cihac, A. de. *Dictionnaire d'étymologie daco-romane, éléments slaves, magyars, turcs, grecs-modernes, et albanais* (Frankfurt am Main: St Goar, 1870–9).
- Cioranescu, A. *Diccionario etimológico rumano*, 7 fascicles (Tenerife: Universidad de La Laguna, 1958–66).
- [Ciudad Real, A. de.] *Calepino maya de Motul*, ed. R. Arzápalo Marín et al., 3 vols. (Mexico City: Universidad Autónoma de México, 1995).
- Clodius, J. C. *Compendiosum lexicon Latino–Turcico–Germanicum* (Leipzig: Deer, 1730).
- Coate, H. H. J., and A. P. Elkin. *Ngarinjin–English Dictionary*, 2 vols. (Sydney: Oceania, 1974).
- Codex Cumanicus bibliothecae ad templum divi Marci Venetiarum*, ed. G. Kuún (Budapest: Scientiarum Academia Hungaricae, 1880).
- Codex Cumanicus: Cod. Marc. Lat. DXLIX*, with introduction by K. Grønbech (Copenhagen: Levin & Munksgaard, 1936).
- Cohen de Lara, D. *‘Ir David, sive, De convenientia vocabulorum rabbinicorum cum Graecis, & quibusdam aliis linguis Europaeis* (Amsterdam: Ravestein, 1648).
- Colección Mutis: Conjunto documental de lenguas indígenas de la Biblioteca del Palacio Real de Madrid, coleccionmutis.cubun.org/Coleccion_Mutis.
- Colenso, J. W. *Zulu–English Dictionary* (Pietermaritzburg: Davis, 1861).
- Collado, D. *Dictionarium sive thesauri linguae Iaponicae compendium* (Rome: Congregatio de Propaganda Fide, 1632).
- Collard, H., and E. S. Collard. *Castellano–mayo; mayo–castellano = Vocabularios indígenas*, 6 (Mexico City: Instituto Lingüístico de Verano, 1962).
- Collins *COBUILD English Language Dictionary*, ed. J. Sinclair (London: HarperCollins, 1987).
- Collins *Cobuild Learner's Dictionary, Concise Edition*, ed. C. Rammell (Glasgow: HarperCollins, 2003).
- Collins *Concise English Dictionary*, 7th edn (Glasgow: Collins, 2008).
- Colloquia of the Hermeneumata Pseudodositheana*, ed. E. Dickey, 2 vols. (Cambridge University Press, 2012–15).
- Comenius, J. A. *Januae linguae Latinae vestibulum* (Košice: Seiderlich, 1676).
- Comenius, J. A. *Janua linguarum reserata, or, A Seed-Plot of all Languages and Sciences* (London: Slater, 1636).
- Comenius, J. A. *Janua linguarum reserata, sive seminarium linguarum et scientiarum omnium* (Leszno: [press of the Unitas Fratrum], 1631).
- Comenius, J. A. *Janua linguarum reserata, sive seminarium linguarum et scientiarum omnium* (Leipzig: Grosius, 1632).
- Comenius, J. A. *Orbis sensualium pictus* (Nuremberg: Endter, 1658).
- Comenius, J. A. *Porta linguarum trilinguis reserata* (London: for the translator, 1631).
- Commentaria et lexica Graeca in papyris reperta*, ed. G. Bastianini et al., 8 vols. planned (Munich and Leipzig: Saur [imprint varies], 2004–).
- Concise Oxford Dictionary*, 7th edn, ed. J. B. Sykes (Oxford: Clarendon, 1982).
- Concise Oxford Dictionary*, 10th edn, ed. J. Pearsall (Oxford: Clarendon, 2001).
- Concise Oxford Dictionary*, 12th edn, ed. A. Stevenson and M. Waite (Oxford University Press, 2011).

- The Concise Oxford Dictionary of Current English*, ed. H. W. Fowler and F. G. Fowler (Oxford: Clarendon, 1911).
- The Connected Discourses of the Buddha: A New Translation of the Saṃyutta Nikāya*, trans. Bhikkhu Bodhi, 2 vols. (Boston, MA: Wisdom, 2000).
- Cooper, T. *Thesaurus linguae Romanae et Britannicae* (London: Berthelet, 1565).
- Córdova, J. de. *Arte en lengua zapoteca* (Mexico City: Balli, 1578).
- Córdova, J. de. *Vocabulario Castellano–Zapoteco*, facs. ed. W. Jiménez Moreno (Mexico City: Instituto Nacional de Antropología e Historia, 1942).
- Córdova, J. de. *Vocabulario en lengua çapoteca* (Mexico City: Ocharte and Ricardo, 1578).
- Corneille, T. *Dictionnaire des arts et des sciences*, 2 vols. (Paris: Coignard, 1694).
- Corominas, J., and J. A. Pascual. *Diccionario crítico etimológico castellano e hispánico*, 6 vols. (Madrid: Gredos, 1980–91).
- Coromines, J., et al. *Diccionari etimològic i complementari de la llengua catalana*, 10 vols. (Barcelona: Curial, 1980–2001).
- Corpus glossariorum Latinorum*, ed. G. Goetz. 7 vols. (Leipzig: Teubner, 1888–1923).
- Corriente, F. *Diccionario avanzado árabe*, vol. I, *Árabe–español*, with I. Ferrando, and vol. II, *Español–árabe*, with A.-S. Ould Mohamed Baba (Barcelona: Herder, 2005–10).
- Cortés y Zedeño, J. T. de A. *Arte, vocabulario, y confessorario en el idioma mexicano, como se usa en el Obispado de Guadalajara* (Puebla: Colegio Real de San Ignacio, 1765).
- Cotgrave, R. *Dictionarie of the French and English Tongues* (London: Islip, 1611).
- Cotton, J. *Vocabulary of the Massachusetts (or Natick) Indian Language*, ed. J. Pickering (Cambridge, MA: Metcalf, 1829).
- Courtois, V. J. *Diccionario Cafre-Tetense–Portuguez ou Idioma fallado no Districto de Tete e na vasta região do Zambeze Inferior: Mafara akukonkédú a chizungu na mu chisendzi cha ku Nyungue* (Coimbra: Imprensa da Universidade, 1900).
- Courtois, V. J. *Diccionario Portuguez–Cafre-Tetense ou idioma fallado no Districto de Tete e na vasta região do Zambeze Inferior: Mafara akukonkédú a chizungu na mu chisendzi cha ku Nyungue* (Coimbra: Imprensa da Universidade, 1899).
- Covarrubias Horozco, S. *Tesoro de la lengua castellana o española* (Madrid: Sanchez, 1611).
- Covarrubias Horozco, S. *Tesoro de la lengua castellana o española, edición integral* (Madrid: Iberoamericana, 2006).
- Cox, B. *The Place-Names of Leicestershire*, 7 vols. (Nottingham: English Place-Name Society, 1998–2016).
- Crastoni, G. [*Lexicon graeco–latinum*] (Milan: Buonaccorso, 1478?).
- Crowther, S. A. *Grammar and Vocabulary of the Yoruba Language* (London: Seeleys, 1852).
- Crowther, S. A. *Grammar of the Yoruba Language* (London: Seeleys, 1852).
- Crowther, S. A. *Vocabulary of the Yoruba Language* (London: Seeleys, 1852).
- [Crowther, S. A.]. *A Vocabulary of the Yoruba Language &c. &c.* (London: n.p., n.d. [probably between 1867 and 1872]).
- Crowther, S. A. *Vocabulary of the Yoruba Language: Part I, English and Yoruba, Part II, Yoruba and English; to which are Prefixed the Grammatical Elements of the Yoruba Language* (London: Church Missionary Society, 1843).
- Crowther, S. A., with J. F. Schön. *Vocabulary of the Ibo Language*, 2 parts (London: Society for Promoting Christian Knowledge, 1882–3).
- CSD = *The Concise Scots Dictionary*, ed. M. A. Robinson (Aberdeen University Press, 1985; ²2017).

- Cũa, P. *Đại Nam quốc âm tự vị*, with 'Lời giới thiệu' ('Introduction') by Bùi Đ. T., 2 vols. (Ho Chi Minh City: NXB Trẻ, 1998; ¹1895–6).
- Cūla Moggallāna. *Abhidhānappāḍīpikā: A Study of the Text and Its Commentary*, ed. Bhikkhu Medagama Nandawansa (Pune: Bhandarkar Oriental Research Institute, 2001).
- Cunha, A. G. da (ed.). *Índice do vocabulário do português medieval*, 3 vols. (Rio de Janeiro: Fundação Casa de Rui Barbosa/Ministério da Cultura, 1986–94).
- Cunha, A. G. da (ed.). *Vocabulário histórico-crônológico do português medieval*, DVD (Rio de Janeiro: Fundação Casa de Rui Barbosa/Ministério da Cultura, 2002; ³2014).
- Cuoq, J.-A. *Lexique de la langue algonquine* (Montreal: Chapleau & Fils, 1886).
- Cuoq, J.-A. *Lexique de la langue iroquoise* (Montreal: Chapleau, 1882).
- Curr, E. M. *The Australian Race: Its Origin, Languages, Customs, Place of Landing in Australia, and the Routes by which it Spread itself over that Continent*, 4 vols. (Melbourne: Ferres; London: Trübner, 1886–7).
- Curtis, E. S. *The North American Indian*, 20 vols. ([Seattle]: Curtis, 1907–30).
- DA = *A Dictionary of Americanisms*, ed. M. M. Mathews, 2 vols. (University of Chicago Press, 1951).
- Ḍabbī, *Amṭāl* = Abū l-‘Abbās al-Mufaḍḍal b. Muḥammad al-Ḍabbī, *Amṭāl al-‘Arab*, ed. I. ‘Abbās (Beirut: Dār al-Rā’id al-‘Arabī, 1981).
- DAE = *Dictionary of American English*, ed. W. A. Craigie, J. R. Hulbert, et al., 4 vols. (University of Chicago Press, 1938–44).
- Dagut, M. *Milon ivri–angli/Hebrew–English Dictionary* (Jerusalem: Kiryat Sefer, 1986).
- Daijirin, ed. Matsumura Akira (Tokyo: Sanseidō, 1988; ³2006).
- Daijisen, ed. Matsumura Akira (Tokyo: Shōgakkān, 1995; ²2012).
- Daijiten, ed. Shimonaka Yasaburō, 26 vols. (Tokyo: Heibonsha, 1934–6).
- Daijiten, ed. Ueda Kazutoshi, Okada Masayuki, Iijima Tadao, Sakaeda Takei, and Iida Den’ichi (Tokyo: Keiseisha, 1917).
- Dai kanwa jiten, ed. Morohashi Tetsuji, 12 vols. plus index (Tokyo: Taishūkan, 1955–60; rev. edn 1984–6; suppl. 1990 and 2000).
- Dai-Nihon kokugo jiten, ed. Matsui Kanji and Ueda Kazutoshi, 4 vols. (Tokyo: Fuzanbō, 1915–19).
- Dal’, V. *Tolkovyj slovar’ živogo velikorusskago jazyka*, 4 vols. (Moscow: [imprints vary], 1863–6).
- Dal’, V. *Tolkovyj slovar’ živogo velikorusskago jazyka*, 3rd edn, ed. J. Baudouin de Courtenay, 4 vols. (St Petersburg: Vol’f, 1903–9).
- D’Alberti di Villanuova, F. *Dizionario universale critico enciclopedico della lingua italiana*, 6 vols. (Lucca: Marescandoli, 1797–1805).
- D’Alberti di Villanuova, F. *Nuovo dizionario italiano–francese* (Marseille: Mossy, 1772).
- Daniil of Moscopole. *Eisagōgikē didaskalia* ([Constantinople: Patriarchal press?], 1802).
- Den Danske Ordbog, ed. E. Hjorth and K. Kristensen, 6 vols. (Copenhagen: Danske Sprog- og Litteraturselskab, 2003–5), ordnet.dk/ddo.
- Dansk Ordbog, udgivet under Videnskabernes Selskabs Bestyrelse, 8 vols. (Copenhagen: [various imprints], 1793–1905).
- Dard, J. *Dictionnaire français–wolof et français–bambara, suivi du dictionnaire wolof–français* (Paris: Imprimerie Royale, 1825).
- DARE = *Dictionary of American Regional English*, ed. F. G. Cassidy, J. H. Hall, et al., 6 vols. (Cambridge, MA: Belknap Press of Harvard University Press, 1985–2013).

- Dās, Ś. S. *Hīndī śabdāsāgara*, 4 vols. (Varanasi: Kāśī-Nāgarī-Pracārīṇī Sabhā, 1916–28, ²1965–75).
- Dasypodius, P. *Dictionarium* (Strasburg: Rihelius, 1535; ²1536).
- Dasypodius, P. *Dictionarium Latinogermanicum*, trans. A. Schorus (Antwerp: Montanus, 1542).
- Database of Latin Dictionaries*, clt.brepolis.net/dld.
- Database of the Old Prussian Linguistic Legacy* (Vilnius University, 2012–13), prusistika.flf.vu.lt/zodynas/apie/.
- Dave, N. L. *Narmakośa: Gujarātī bhāṣhāno prathama śabdārtha kośa*, new edn (Surat: Kavi Narmada Yūgāvarta Ṭraṣṭa, 1998; ¹1861–73).
- David ben Yesha^c. *Millon ‘ivriy-‘araviy al-Jāmi^c – Ha-M’assef*, ed. S. Gamliel (Jerusalem: Makhon Shalom le-shivtei Yeshurun, 1988).
- Davis, W. J. *A Dictionary of the Kaffir Language, Including the Xosa and Zulu Dialects: Part I, Kaffir–English* (London: Wesleyan Mission House, 1872).
- Davis, W. J. *An English–Kaffir Dictionary, Principally of the Xosa–Kaffir but Including Also Many Words of the Zulu–Kaffir Dialect* (London: Wesleyan Mission House, 1877).
- DCHP = *A Dictionary of Canadianisms on Historical Principles*, ed. W. S. Avis et al. (Toronto: Gage, 1967).
- DCHP-1 Online: *A Dictionary of Canadianisms on Historical Principles Online*, prepared by S. Dollinger, L. J. Brinton, and M. Fee (Vancouver: University of British Columbia, 2013), dchp.ca/DCHP-1.
- DCHP-2 = *A Dictionary of Canadianisms on Historical Principles*, 2nd edn, ed. S. Dollinger with M. Fee (Vancouver: University of British Columbia, 2017), www.dchp.ca/dchp2.
- DeFrancis, J. *ABC Chinese–English Comprehensive Dictionary* (Honolulu: University of Hawai‘i Press, 2003).
- De Lange, N. R. M. (ed.). *Greek Jewish Texts from the Cairo Genizah* (Tübingen: Mohr, 1996).
- Della Bella, A. *Dizionario italiano, latino, illirico* (Venice: Zanne, 1728).
- De Man, Louis, ed. *Middeleeuwse systematische glossaria* (Brussels: n.p., 1964).
- Deshpande, P. G., and B. Deshpande. *Universal English–Gujarati Dictionary* (Mumbai: Oxford University Press, 1988).
- Deutsches Wörterbuch*, ed. J. Grimm and W. Grimm et al., 33 vols. (Leipzig: Hirzel, 1852–1971), woerterbuchnetz.de/DWB/.
- The Dhammasaṅgaṇi*, ed. E. Müller (London: Pali Text Society, 1978; ¹1885).
- Dhanapāla, Pāṭiyalacchī = G. Bühler (ed.), ‘The Pāṭiyalacchī Nāmamālā, a Prakrit Kosha, by Dhanapāla’, *Beiträge zur Kunde der indogermanischen Sprachen*, 4 (1878), 70–166.
- Dharaṇidāsa, *Anekārthasāra* = *Dharanikośa of Dharaṇidāsa*, part 1, ed. E. D. Kulkarni (Pune: Deccan College Postgraduate and Research Institute, 1968).
- Díaz, F. ‘Vocabulario de letra China con la explication castellana hecho con gran propiedad y abundancia de palabras’ (Kraków: Jagiellonian Library, Ms. Berol. Ms. Sin 13. Hisz. chin., XVII).
- Dicionario da lingua portuguesa* (Lisbon: Academia das Ciências, 1793; reprint 1976 as *Dicionário da língua portuguesa*).
- ‘Dicionario da Lingua Geral do Brazil’ (Universidade de Coimbra, MS 69, pp. 237–353), digitalis-dsp.uc.pt/bgz/UCBG-Ms-69/UCBG-Ms-69_item1.
- Diccionario de la lengua castellana, compuesta por la Real Academia Española*, 6 vols. (Madrid: Hierro, 1726–39).

- Diccionario de la lengua castellana, reducido a un tomo para su más fácil uso* (Madrid: Ibarra, 1780; ²³2014), dle.rae.es.
- Diccionario grande de la lengua de Michuacán*, ed. J. B. Warren, 2 vols. (Morelia, Mexico: Fimax Publicistas, 1991).
- Diccionario k'iche' de Berlín: el vocabulario en lengua 4iche otlatecas*, ed. M. Dürr and F. Sachse (Berlin: Mann, 2017).
- Diccionario português-brasiliano e brasileiro-português*, ed. P. M. da S. Ayrosa (São Paulo: Imprensa official do estado, 1934).
- Diccionario português, e brasileiro, obra necessaria aos ministros do altar* (Lisbon: Officina Patriarcal, 1795).
- Diccionario y gramatica chibcha: manuscrito anónimo de la Biblioteca Nacional de Colombia*, ed. M. S. González de Pérez (Bogotá: Instituto Caro y Cuervo, 1987).
- Dicionário da língua portuguesa contemporânea*, 2 vols. (Lisbon: Verbo, 2001).
- Dickey, E., and R. Ferri (eds.). 'A new edition of the Latin-Greek glossary on P. Sorb. inv. 2069 (verso)', *Zeitschrift für Papyrologie und Epigraphik*, 175 (2010), 177–87.
- Dictionarium Latino Lusitanicum, ac Iaponicum, ex Ambrosii Calepini volumine depromptum* (Amakusa: in Collegio Iaponico Societatis Iesu, 1595).
- Dictionarium tetraglotton, seu voces Latinae omnes, et Graecae eis respondentes, cum Gallica et Teutonica (quam passim Flandricam vocant) earum interpretatione*, ed. C. Kilian (Antwerp: Plantin, 1562).
- Dictionarul etimologic al limbii române* (Bucharest: Editura Academiei Române, 2011–).
- Dictionarul explicativ al limbii române* (Bucharest: Editura Academiei Republicii Socialiste România, 1975).
- Dictionarul limbii române*, 5 vols. (Bucharest: [various imprints], 1913–49; first fascicle 1907).
- Dictionarul limbii române: serie nouă*, 14 vols. (Bucharest: EARSR/Editura Academiei române, 1965–2010).
- Dictionnaire de l'Académie française*, 2 vols. (Paris: Coignard, 1694).
- Dictionnaire de l'occitan médiéval*, ed. W.-D. Stempel et al., 7 fascicles plus 1 supplement, continued online (Tübingen: Niemeyer [imprint varies], 1996–2013, continued 2013–), dom-en-ligne.de.
- Dictionnaire du moyen français*, ed. R. Martin et al. (Nancy: CNRS/Université Nancy 2 [imprints vary], 2003–), www.atilf.fr/dmf.
- Dictionnaire Usito* (Quebec City: Editions Delisme, 2009–), usito.com.
- Diderot, D., and J. d'Alembert (eds.). *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts, et des métiers*, 35 vols. (Paris: Panckouke, 1751–80).
- Diez, F. *Grammatik der romanischen Sprachen*, 3 vols. (Bonn: Weber, 1836–43).
- Diez, F. *Lexicon etymologicum linguarum Romanarum, Italicae, Hispanicae, Gallicae/ Etymologisches Wörterbuch der romanischen Sprachen* (Bonn: Marcus, 1853).
- Dīghanikāya = The Dīgha Nikāya*, ed. T. W. R. Davids and J. E. Carpenter, 3 vols. (London: Pali Text Society, 1966–76; ¹1890–1911).
- Digital Corpus of Cuneiform Lexical Texts*, ed. N. C. Veldhuis (2003–), oracc.museum.upenn.edu/dcclt/.
- Digitales Wörterbuch der deutschen Sprache*, www.dwds.de.
- Dihkhuda, 'Alī Akbar, with Muḥammad Mu' īn and Ja'far Shahīdī (eds.). *Lughatnāmah*, 222 fascicles in 50 vols. (Tehran: Government Printing Press and Tehran University, 1946–80)

- [revised and reissued in 15 vols. (Tehran University, 1993–4); revised and reissued in 16 vols. (Tehran: Dihkhuda Dictionary Foundation, 1998); available on CD from 1994.]
- Dihlavi, S. A. *Farhang-i āṣāfiyah*, reprint edn, 4 vols. (New Delhi: National, 1974; ¹1918).
- Dīnawarī, Nabāt = Abū Ḥanīfa Aḥmad b. Dāwūd al-Dīnawarī, *al-Nabāt*, ed. B. Lewin (Wiesbaden: Steiner, 1974).
- Diouf, J.-L. *Dictionnaire wolof-français et français-wolof* (Paris: Karthala, 2003).
- Dixon, R. M. W. *The Dyirbal Language of North Queensland* (Cambridge University Press, 1972).
- Dixon, R. M. W., W. S. Ramson, and M. Thomas, *Australian Aboriginal Words in English: Their Origin and Meaning* (Melbourne: Oxford University Press, 1990; ²2006).
- ‘Dizionario’ [links page] (Florence: Accademia della Crusca, 2011), www.accademiadella-crusca.it/it/link-utili/dizionari-sincronici.
- DNE = *Dictionary of Newfoundland English*, ed. G. M. Story, W. J. Kirwin, and J. D. A. Widdowson (University of Toronto Press, 1982; ²1990).
- Dobson, S. *The Geordie Dictionary* (Newcastle: Graham, 1974).
- Doctrina Christiana y catecismo para instruccion de los Indios, y de las de mas personas, que han de ser enseñadas en nuestra sancta Fé, con un confessorario, y otras cosas necesarias para los que doctrinan . . . traducido en las dos lenguas generales, de este Reyno, Quichua, y Aymara* (Lima: Ricardo, 1584).
- Doke, C. M., D. McK. Malcolm, and J. M. A. Sikakana. *English–Zulu Dictionary* (Johannesburg: Witwatersrand University Press, 1958).
- Doke, C. M., and B. W. Vilakazi. *Zulu–English Dictionary* (Johannesburg: Witwatersrand University Press, 1948; ²1972).
- Doke, C. M., et al. *English–Zulu/Zulu–English Dictionary* (Johannesburg: Witwatersrand University Press, 1990).
- Dolan, T. P. *Dictionary of Hiberno-English* (Dublin: Gill and Macmillan, 1988; ²2004).
- Domínguez, R. J. *Diccionario nacional, o gran diccionario clásico de la lengua española*, 2 vols. (Madrid: Dominguez, 1846–7).
- Doniach, N. S. *The Oxford English–Arabic Dictionary of Current Usage* (Oxford: Clarendon, 1972).
- Doroszewski, W. *Słownik języka polskiego*, 11 vols. (Warsaw: Państwowe Wydawnictwo ‘Więda Powszechna’ and Państwowe Wydawnictwo Naukowe, 1958–69).
- DOST = *Dictionary of the Older Scottish Tongue*, ed. W. A. Craigie, A. J. Aitken, J. A. C. Stevenson, M. G. Dareau, et al., 12 vols. (University of Chicago Press; Aberdeen University Press, 1931–2002).
- Douglas, C. *Chinese–English Dictionary of the Vernacular or Spoken Language of Amoy: With the Principal Variations of the Chang-Chew and Chin-Chew Dialects* (London: Trübner, 1873).
- Douglas, W. H. *Bardi Language Word-Book: A Revision of ‘Word Gems from Iwanya-Sunday Island’* (Broome: Jawa, 1992).
- Douglas, W. H. *Illustrated Topical Dictionary of the Western Desert Language, Warburton Ranges Dialect* (Perth: United Aborigines Mission Language Department, 1959).
- Douglas, W. H. *An Introduction to the Western Desert Language* (University of Sydney, 1957; ²1964).
- Douglas, W. H. *An Introductory Dictionary of the Western Desert Language* (Perth: Institute of Applied Language Studies, Western Australian College of Advanced Education, 1988).

- Douglas, W. H. 'Phonology of the Australian Aboriginal language spoken at Ooldea, South Australia, 1951–1952', *Oceania*, 25.3 (1955), 216–29.
- Dozy, R. P. A. *Supplément aux dictionnaires arabes*, 2 vols. (Leiden: Brill, 1881).
- Dror, M. *Milon ivri–rusi/Evrejsko (ivrit)–russkij slovar'* (Tel Aviv: Am Oved, 1975).
- Drummond, R. *Illustrations of the Grammatical Parts of the Guzerattee, Mahratta, & English Languages* (Mumbai: Courier, 1808).
- du Cange, C. *Glossarium ad scriptores mediae et infimae graecitatis* (Lyon: Anisson, 1688).
- du Cange, C. *Glossarium ad scriptores mediae et infimae latinitatis*, 3 vols. (Paris: Billaine, 1678).
- Duden: *Das große Wörterbuch der deutschen Sprache*, ed. G. Drowsdowski et al., 6 vols. (Mannheim: Bibliographisches Institut, 1976–81; ²1993–5).
- Duden: *Das große Wörterbuch der deutschen Sprache*, ed. W. Scholze-Stubenrecht et al., 3rd edn, 10 vols. (Mannheim: Dudenverlag, 1999).
- Duden: *Deutsches Universalwörterbuch*, ed. G. Drowsdowski (Mannheim, Bibliographisches Institut, 1983; ⁸2015), www.duden.de/.
- Duncan, A. 'Index vocum' in Duncan, *Latinae grammaticae pars prior* (Edinburgh: Waldegrave, 1595), appendix, sigs. A7 v–F7 v.
- Dunglison, R. 'Dunglison's glossary (1829–1830)', ed. A. W. Read, *Dialect Notes*, 5 (1927), 422–32.
- Duo glossaria: Anonymi Montepessulanensis dictionarius*, ed. A. Grondeux, and *Glossarium Gallico–Latinum*, ed. B. Merrilees and J. Monfrin (Turnhout: Brepols, 1998).
- Durbin, M. E. 'A componential analysis of the San Carlos dialect of Western Apache: A study based on the analysis of the phonology, morphophonics, and morphemics', unpublished PhD dissertation, State University of New York, Buffalo, 1964.
- Đurđević, B. *De afflictione tam captivorum quam etiam sub Turcae tributo viventium Christianorum* (Antwerp: Copenius, 1544).
- Durgācāryakṛtavṛttisametam Niruktam: Pūrvaṣaṭkātmakaḥ prathamō bhāgaḥ*, ed. V. K. Rājvāde, 2 vols. (Poona: Anandashram, 1921–6).
- Early Irish Glossaries Database*, ed. P. Russell, S. Arbutnot, and P. Moran, www.asnc.cam.ac.uk/irishglossaries/.
- Edkins, J. *A Vocabulary of the Shanghai Dialect* (Shanghai: Presbyterian Mission, 1869).
- Egede, P. *Dictionarium Grönlandico–Danico–Latinum* (Copenhagen: Orphanotrophium Regii, 1750).
- Eighteenth Century Xebero: Mss. Add. 25,323 and 25,324 of the British Library, London (UK)*, ed. A. Alexander-Bakkerus (Munich: Lincom, 2016).
- Ekārthak kośa*, ed. M. Dulaharāja and S. Kusumaprajñā, 2nd edn (Ladnun: Jain Vishva Bharati, 2003).
- El'azar ben Ḥalfon. *Pitron ha-ʿAnaq me-et El'azar ben Ḥalfon: be'ur millot ha-tsimud be-Sefer ha-ʿAnaq le-Moshe ibn 'Ezra/ʿAnaq Commentary by Eleazar ben Ḥalfon: Interpretation of the Homonyms in Moses ibn Ezra's 'Anaq*, ed. A. Dotan and N. Basal (Jerusalem: Ben-Zvi Institute, 2011).
- The Electronic Text Corpus of Sumerian Literature*, ed. J. A. Black et al. (1997–2006), etcsl.orinst.ox.ac.uk/.
- Elger, G. *Dictionarium Polono–Latino–Lottavicum* (Vilnius: typis academicis Societatis Iesu, 1683).
- Elias, E. A. *Elias' Modern Dictionary Arabic–English* (Cairo: Elias, 1922).

- Elias, E. A. *Elias' Modern Dictionary Arabic-English*, 9th edn, ed. E. E. Elias (Cairo: Elias, 1962).
- Elias, E. A. *al-Qāmūs al- 'aṣṣī, Inkilīzī- 'Arabī/The new dictionary, English-Arabic* (Cairo: al-Muqtaṭaf, 1913).
- Elliott, W. A. *Dictionary of the Tebele & Shuna Languages* ([London]: for the author, 1897).
- Elyot, T. *The Dictionary* (London: Berthelet, 1538).
- Endzelīns, J., and E. Hauzenberga. *Ergänzungen und Berichtigungen zu K. Mühlenbachs Lettisch-deutschem Wörterbuch/Papildinājumi un labojumi K. Mühlenbacha Latviešu valodas vārdnīcai*, 2 vols. (Riga: Lettisches Kulturfonds/Grāmatu apgāds, 1934–46).
- English-Tamil Dictionary*, ed. A. Chidambaranatha Chettiar with K. Appadurai et al., 3 vols. (Chennai: University of Madras, 1963–5).
- Epimerismi Homerici*, ed. A. R. Dyck, 2 vols. (Berlin and New York: De Gruyter, 1983–95).
- The Épinal, Erfurt, Werden, and Corpus Glossaries*, ed. B. Bischoff et al. (Copenhagen: Rosenkilde and Bagger, 1988).
- Ėryā gūlin 爾雅詁林, ed. Zhū Zūyán 朱祖延, 5 vols. (Wuhan: Húběi jiàoyù, 1996).
- Es'ad Efendi, M. *Lehceü'l-luğât* (Istanbul: n.p., 1795).
- Es'ad Efendi, M. *Lehceü'l-luğât*, 2nd edn (Istanbul: Darü't-tibaatü'l-mamure, 1801).
- Es'ad Efendi, M. *Lehceü'l-luğât*, edited in Latin script by A. Kirkkılıç (Ankara: Türk Dil Kurumu, 1999).
- Estienne, H. *Thesaurus Graecae linguae*, 5 vols. (Geneva: Estienne, 1572–3).
- Estienne, R. *Dictionarium Latinogallicum* (Paris: Estienne, 1538).
- Estienne, R. *Dictionarium, seu Latinae linguae thesaurus* (Paris: Estienne, 1531; ²1536).
- Estienne, R. *Dictionarium, seu Latinae linguae thesaurus . . . editio secunda* [really ed. 3], 3 vols. (Paris: Estienne, 1543).
- Estienne, R. *Dictionnaire françoislatin* (Paris: Estienne, 1539).
- An Ethnologic Dictionary of the Navaho Language* (St Michaels, AZ: Franciscan Fathers, 1910).
- Etymologicum graecae linguae Gudianum*, ed. F. W. Sturz (Leipzig: Weigel, 1818).
- Etymologicum Gudianum quod vocatur*, ed. E. L. de Stefani, 2 vols. [*a-zeiai*] (Leipzig: Teubner, 1909–20).
- Etymologicum magnum genuinum; Symeonis etymologicum cum magna grammatica; Etymologicum magnum auctum*, ed. F. Lasserre and N. Livadaras, 2 vols. [*a-b*] (Rome: Ateneo, 1976; Athens: Parnassos, 1992).
- Etymologicum magnum graecum* (Venice: Kallierges, 1499).
- Etymologicum magnum, seu verius, Lexicon saepissime vocabulorum origines indagans*, ed. T. Gaisford (Oxford: Typographeus Academicus, 1848).
- Etymologicum parvum quod vocatur*, ed. R. Pintaudi (Milan: Istituto Editoriale Cisalpino, 1973).
- Etymologicum Symeonis*, ed. D. Baldi [*g-e*, continuing the edition begun in *Etymologicum magnum genuinum* (1976–1992)] (Turnhout: Brepols, 2013).
- Even-Shoshan, A. *Hamilton hekhdash*, 4 vols. plus supplement (Jerusalem: Kiryat Sefer, 1966–70, 1983).
- Even-Shoshan, A. *Milon khadash*, 4 vols. plus supplement (Jerusalem: Kiryat Sefer, 1947–52, 1958).
- Faber, B. *Thesaurus eruditionis scholasticae* (Leipzig: Rhamba, 1572).
- Fabra, P. *Diccionari general de la llengua catalana* (Barcelona: Llibreria Catalònia, 1932).
- Fabricius, O. *Den grønlandske Ordbog forbedret og forøget* (Copenhagen: Kongelige Vaisenhus, 1804).

- Fabvre, B. *Racines montagnaises*, ed. L. Angers and G. McNulty (Quebec City: Presses de l'Université Laval, 1970).
- Fallon, S. W. *A Hindustani–English Law and Commercial Dictionary* (Varanasi: Lazarus, 1879).
- Fallon, S. W. *A New Hindustani–English Dictionary, with Illustrations from Hindustani Literature and Folk-Lore*, 2 vols. (Varanasi: Medical Hall Press for Trübner, 1879–83).
- Fang Yizhi. *Tongya* 通雅 = *Siku quanshu*, vol. 857 (Taipei: Taiwan shangwu yinshuguan, 1983).
- Fāngyán jiào jiān* 方言校箋, ed. Zhōu Zǔmó 周祖謨 (Beijing: Kēxué chūbānshè, 1956).
- Fārābī, Dīwān = Abū Ibrāhīm Ishāq b. Ibrāhīm al-Fārābī, *Dīwān al-adab*, ed. A. M. 'Umar, 5 vols. (Cairo: al-Hay'a al-'Āmma li-Ṣu'ūn al-Maṭābī' al-Amīriyya, 1974–9).
- Farāhī, A. N. M. B. *Niṣāb al-ṣibyān* (Tehran: Cāpḵhānah-yi Shirkat-i Kāviyānī, 1923).
- Farmer, J. S. *Americanisms Old and New* (London: for the author, 1889).
- Faruqī, H. S. *al-Mu'jam al-qānūnī: 'Arabī–Inkilizī/Faruqī's Law Dictionary Arabic–English* (Beirut: Librairie du Liban, 1972).
- Faruqī, H. S. *al-Mu'jam al-qānūnī, Iniklīzī–'Arabī/Faruqī's Law Dictionary English–Arabic* (Tripoli: Dār al-Nashr al-Lībīyah, 1962).
- Favorino, Guarino. *Magnum ac perutile dictionarium* (Rome: Kallierges, 1523).
- Favorino, Guarino. *Magnum ac perutile dictionarium*, 2nd edn, ed. J. Camerarius (Basel: Winter, 1538).
- Febrés, A. *Arte de la lengua general del Reyno de Chile, con un dialogo chileno–hispano muy curioso: a que se añade . . . por fin un vocabulario hispano–chileno, y un calepino chileno–hispano mas copioso* (Lima: En la calle de la Encarnación, 1765).
- Fenton, J. *The Hamely Tongue* (Belfast: Ullans Press, 1995; 2014).
- Féraud, J.-F. *Dictionnaire critique de la langue françoise*, 3 vols. (Marseille: Mossy, 1787).
- Ferraguto, P. *Il 'Dittionario della Lingua Turchesca'*, ed. L. Rocchi (Edizioni Università di Trieste, 2012).
- Ferreira, A. Buarque de Holanda. *Novo Dicionário da Língua Portuguesa* (Rio de Janeiro: Nova Fronteira, 1975).
- FEW = W. von Wartburg et al. (eds.). *Französisches etymologisches Wörterbuch: eine Darstellung des galloromanischen Sprachschatzes*, 25 vols. (Bonn etc.: [imprint varies], 1922–2002).
- Figueiredo, C. de. *Nôvo dicionário da língua portuguesa*, 2 vols. (Lisbon: Tavares Cardoso & Irmão, 1899).
- Finkel, I. L. (ed.). *The Series SIG₇.ALAN = Nabnītu = Materials for the Sumerian Lexicon*, 16 (Rome: Pontificium Institutum Biblicum, 1982).
- Firūzābādī, Qāmūs = Maḡd al-Dīn Abū l-Ṭāhīr Muḥammad b. Ya'qūb al-Firūzābādī, *al-Qāmūs al-muḥīṭ*, 4 vols. (Cairo: Muṣṭafā l-Bābī al-Ḥalabī, 1952).
- Flierl, J. 'German vocabulary of native tribes, North East South Australia' (Sydney: State Library of New South Wales, Mitchell Library, MS 161, Box 1, Folder 2, Item 2).
- Florio, J. *A Worlde of Wordes* (London: Blount, 1598).
- Francard, M., et al. *Dictionnaire des belgicisms*, 2nd edn (Louvain-la-Neuve: De Boeck Duculot, 2015; 2010).
- Franciosini, L. *Vocabolario italiano e spagnolo* (Rome: Ruffinelli and Manni, 1620).
- Frantext (ATILF-CNRS and Université de Lorraine, 1998–), frantext.fr.
- Fraser, A. (ed.). *Huron Manuscripts from Rev. Pierre Potier's Collection* = Fifteenth Report of the Bureau of Archives for the Province of Ontario (Toronto: James, 1920).

- Freire-Marreco, B. 'Tonto Apache vocabulary and information' (Washington, DC: National Anthropological Archives, Smithsonian Institution, MS 1747).
- French–Onondaga Dictionary, from a Manuscript of the Seventeenth Century, ed. J. G. Shea = [Shea's] Library of American Linguistics, 1 (New York: Cramoisy, 1860).
- Freytag, G. W. *Lexicon Arabico–Latinum, praesertim ex Djeuharii Firuzabaditque et aliorum Arabum operibus, adhibitis Golii quoque et aliorum libris confectum*, 4 parts in 2 vols. (Halle: Schwetschke, 1830–7).
- Frisch, J. L. *Teutsch–Lateinisches Wörter-Buch* (Berlin: Nicolai, 1741).
- Frisch, J. L. *Teutsch–Lateinisches Wörter-Buch*, facs. edn, ed. G. Powitz (Hildesheim: Olms, 1977).
- Fruchtman, M., and D. Sivan. *Milon Ariel Hamakif* (Kiryat Gat: Korim, 2007).
- Fürecker, C. *Lettisches und Teutsches Wörterbuch* (1997) = *Fürecker's Dictionary: The First Manuscript*, ed. T. G. Fennell (Riga: Latvijas Akadēmiskā Bibliotēka, 1997).
- Fürecker, C. *Lettisches und Teutsches Wörterbuch* (1998) = *Fürecker's Dictionary: The Second Manuscript*, ed. T. G. Fennell (Riga: Latvijas Akadēmiskā Bibliotēka, 1998).
- Furetière, A. *Dictionnaire universel*, 3 vols. (The Hague and Rotterdam: Leers, 1690).
- Gale, J. S. *A Korean–English Dictionary* (Yokohama: Kelly & Walsh, 1897).
- Gale, J. S. *The Unabridged Korean–English Dictionary*, ed. A. A. Pieters (Seoul: Christian Literature Society of Korea, 1931).
- Gangler, J.-F. *Lexicon der Luxemburger Umgangssprache* (Luxemburg: Hoffman, 1847).
- Ġarbādaqānī, *Rawḥa* = Muḥaddab al-Dīn Muḥammad b. al-Ḥasan al-Ġarbādaqānī, *al-Rawḥa (fi ḥarfay al-dād wa-l-zā)*, 2 vols. (Frankfurt: Ma'had Tārīḥ al-'Ulūm al-'Arabiyya wa-l-Islāmiyya, 1985).
- Gardiner, A. H. (ed.). *Ancient Egyptian Onomastica*, 3 vols. (Oxford University Press, 1947).
- Garzoni, M. *Grammatica e vocabolario della lingua Kurda* (Rome: Congregazione de Propaganda Fide, 1787).
- Gatschet, A. S. 'Classification into seven linguistic stocks of western Indian dialects contained in forty vocabularies' in G. M. Wheeler, *Report on United States Geographical Surveys West of the One Hundredth Meridian*, vol. VII, *Archaeology, Appendix, Linguistics* (Washington, DC: Government Printing Office, 1879).
- Gatschet, A. S. 'Pinal Apache: vocabulary and texts with interlinear translation' (Washington: National Anthropological Archives, Smithsonian Institution, MS 1567).
- Gatschet, A. S. *Zwölf Sprachen aus dem Südwesten Nordamerikas (Pueblos- und Apache-Mundarten; Tonto, Tonkawa, Digger, Utah)* (Weimar: Böhlau, 1876).
- Ġawālīqī, Mu'arrab = Abū Maṣṣūr Mawḥūb b. Aḥmad al-Ġawālīqī, *al-Mu'arrab min al-kalām al-aḡamī 'alā ḥurūf al-mu'ḡam*, ed. A. M. Šākir (Cairo: Dār al-Kutub al-Miṣriyya, 1361 AH).
- Ġawḥarī, *Šiḥāḥ* = Abū Naṣr Ismā'īl b. Ḥammād al-Ġawḥarī, *al-Šiḥāḥ, Tāğ al-luḡa wa-šihāḥ al-'Arabiyya*, ed. A. 'Aṭṭār, 2nd edn, 6 vols. (Beirut: Dār al-'Ilm li-l-Malāyīn, 1979).
- Gelenius, S. *Lexicon symphonum, quo quattuor linguarum Europae familiarium, Graecae scilicet, Latinae, Germanicae ac Slavonicae concordia consonantiaque indicatur* (Basel: Frobenius and Episcopus, 1537).
- Ġemāl-al-Dīn ibnī Mühennā. *Moğolca İbni Mühenna Lüğati*, ed. B. Gül (Ankara: Türk Kültürünü Araştırma Enstitüsü, 2016).
- Génibrel, J. F. M., *Dictionnaire annamite-français*, 2nd edn (Saigon: Imprimerie de la Mission à Tân Định, 1898).

- Gerov, N. *Rěčnik na bǎlgarskij jazyk*, 5 vols. (Plovdiv: Družestvena Pečetinca 'Sǎglasie', 1895–1904).
- Gesenius, W. *Hebräisch–Deutsches Handwörterbuch über die Schriften des Alten Testaments*, 2 vols. (Leipzig: Vogel, 1810–12).
- Gesenius, W. *Thesaurus philologicus criticus linguae Hebraeae et Chaldaeae Veteris Testamenti* (Leipzig: Vogel, 1829–58).
- Geytenbeek, B., and H. Geytenbeek. *Gidabal Grammar and Dictionary* (Canberra: Australian Institute of Aboriginal Studies, 1971).
- Ghiyās al-Dīn. *Ghiyās al-Lughāt*, ed. M. Dabīr Siyāqī (Tehran: Kānūn-i Maʿrifat, 1958).
- Gilberti, M. *Vocabulario en lengua de Mechuacan* (Mexico City: Bressano, 1559).
- Gilberti, M. *Vocabulario en lengua de Mechuacán*, ed. A. J. Zavala, rev. by C. Martínez and J. B. Warren (Zamora: Colegio de Michoacán; Mexico City: Fideicomiso Teixidor, 1997).
- Gilchrist, J. B. *A Dictionary, English and Hindoostanee, in which the Words are Marked with their Distinguishing Initials; as Hinduwee, Arabic, and Persian, whence the Hindoostanee, or what is Vulgarly, but Improperly Called the Moor Language, is Evidently Formed*, 2 vols. (Kolkata: Stuart and Cooper, 1787–90 [really 1798]).
- Giles, H. *A Chinese–English Dictionary* (London: B. Quaritch; Shanghai: Kelly & Walsh, 1892).
- Gilliéron, J., and E. Edmont. *Atlas linguistique de la France*, 10 vols. (Paris: Champion, 1902–10).
- [Giorda, J., et al.]. *A Dictionary of the Kalispel or Flat-Head Language*, 2 parts (Montana: St Ignatius Mission, 1877–9).
- Glasgow, K. *Burarra-Gun-Nartpa Dictionary with English Finder List* (Darwin: Summer Institute of Linguistics, 1994).
- Glass, A. *Ngaanyatjarra Word List* (Kalgoorlie: United Aborigines Mission Language Department, 1975; 21988).
- Glass, A., and D. Hackett. *Ngaanyatjarra & Ngaatjatjarra to English Dictionary* (Alice Springs: IAD, 2003).
- Glosas Emilianenses*, ed. H. J. Wolf (Universidad de Sevilla, 1996).
- Le glossaire de Bâle*, ed. M. Banitt (Jerusalem: Israel Academy of Science, 1972).
- Glossaire des patois de la Suisse romande*, ed. L. Gauchat et al., 126 fascicles to date (Neuchâtel and Paris: Attinger, 1924–).
- Glossaire du parler français au Canada* (Québec: L'Action Sociale, 1930).
- Glossaria Latina*, ed. W. M. Lindsay et al., 5 vols. (Paris: Les Belles Lettres, 1926–31).
- Godefroy, F. *Dictionnaire de l'ancienne langue française et de tous ses dialectes du IXe au XVe siècle*, 10 vols. (Paris: Vieweg-Bouillon, 1880–1902).
- Goi, ed. Kimura Masakoto and Yokoyama Yoshikiyo, 4 vols. (Tokyo: Henshūyō, 1871–84).
- Goliuz, J. 'Dictionarium Persico–Latinum', in E. Castell, *Lexicon heptaglotton* (London: Roycroft, 1669), vol. II, separately registered appendix.
- Goliuz, J. *Lexicon Arabico–Latinum* (Leiden: Elsevir, 1653).
- González Holguín, D. *Vocabulario de la lengua general de todo el Peru llamada lengua qquichua o del Inca* (Lima: del Canto, 1608).
- Goodwin, G. 'Word lists, phonology, conjugations' = Grenville Goodwin Papers, series 1 (Tucson: Arizona State Museum Archives, University of Arizona, Ms. 17, Folders 1–23).
- Goryeodae Hangukeodaesajeon*, 3 vols. (Seoul: Korea University Press, 2009).
- Government of Tamil Nadu. *Tamiḷnāṭṭu āṭci corṅkaḷ* (Chennai: Aṇṇā Tamiḷc Curukkēḷuttuk Kallūri, 1968).

- Gramatica Marastta: a mais vulgar que se pratica nos reinos do Nizamaxà e Idalxà* (Rome: Congregação de Propaganda Fide, 1778).
- Gramática, vocabulario, catecismo, i confesionario de la lengua chibcha*, ed. E. Uricoechea (Paris: Maisonneuve, 1871).
- Grand dictionnaire Ricci de la langue chinoise*, 7 vols. (Paris: Desclée de Brouwer, 2001).
- [Gravier, J. (attrib.)]. 'Illinois-French dictionary' (manuscript in the American Indian Collection of the Watkinson Library, Trinity College, Hartford, CT).
- Grazovski, Y. *Milon hasafa haivrit*, 3 vols. plus supplement (Tel Aviv: Dvir, 1934–6, 1941).
- Grazovski, Y., and J. Klausner. *Milon shel kis meivrit lerusit veashkenazit umerusit leivrit veashkenazit, khelek rishon: ivri–rusi–ashkenazi* (Warsaw: Tushiya, 1900).
- Grazovski, Y., and D. Yellin. *Hamilton haivri* (Jaffa: Eyntan and Shoshani, 1919).
- The Great Tzotzil Dictionary of Santo Domingo Zinacantan, with Grammatical Analysis and Historical Commentary*, ed. and trans. R. M. Laughlin with John B. Haviland, 3 vols. (Washington, DC: Smithsonian Institution, 1988).
- Grey, G. A *Vocabulary of the Dialects of South-Western Australia* (Perth: for the author, 1839; ²1840).
- Grossman, R., and M. Segal. *Milon ivri–angli shalem/Compendious Hebrew–English Dictionary* (Tel Aviv: Dvir, 1938).
- Gu Yanwu. *Rizhi lu* 日知錄 = *Siku quanshu*, vol. 858 (Taipei: Taiwan shangwu yinshuguan, 1983).
- Gù Yěwáng. *Yuánběn Yùpiān cánjuàn* 原本玉篇殘卷 (Beijing: Zhōnghuá shūjù, 1985).
- Guadalajara, T. de. *Compendio del arte de la lengua de los tarahumares y guazápare* (Puebla de los Angeles: de Leon, 1683).
- Guadalajara, T. de. *Gramática Tarahumara*, ed. A. Rodríguez López (Universidad Autónoma de Ciudad Juárez, 2010).
- Guðmundur Andrésson. *Lexicon Islandicum*, ed. P. J. Resen (Copenhagen: Gerhard, 1683).
- Guevara, M. de. 'Arte doctrinal y modo general para aprender la lengua Matlaltzinga', *Boletín de la Sociedad Mexicana de Geografía y Estadística*, 9 (1862–3), 198–260.
- Gug-eo Saesajeon*, ed. by the Society of Korean Language and Literature (Seoul: Dongchulpansa, 1958).
- Guignes, C. L. J. de. *Dictionnaire chinois, français, et latin* (Paris: Imprimerie Impériale, 1813).
- Günther, J. W. 'Grammar and vocabulary of the Aboriginal dialect called the Wirradhuri' in Threlkeld, *Australian Language* (1892), Appendix D, 56–120.
- Guóyīn chángyòng zìhuì* 國音常用字彙 (Shanghai: Shāngwù Yīnshūguǎn, 1932).
- Guóyīn zìdiǎn* 國音字典 (Shanghai: Shāngwù Yīnshūguǎn, 1919).
- Gur [Grazovski], Y. *Milon ivri* (Tel Aviv: Dvir, 1946).
- Haack, F. W. *Vocabularium Lithuanico-Germanicum et Germanico-Lithuanicum* (Halle: Orban, 1730).
- Habdelić, J. *Dictionar, ili réchi szlovenszke z vexega ukup zebrane, u red posztaulyene ... na pomoch napredka u diachkom navuku skolneh mladenczeu horvatszkoga i szlovenszkoga naroda* (Graz: Widmanstad, 1670).
- Haensch, G., and R. O. Werner (eds.). *Nuevo diccionario de argentinismos* (Bogotá: Instituto Caro y Cuervo de Bogotá, 1993).
- Haensch, G., and R. O. Werner (eds.). *Nuevo diccionario de colombianismos* (Bogotá: Instituto Caro y Cuervo de Bogotá, 1993).

- Haensch, G., and R. O. Werner (eds.). *Nuevo diccionario de uruguayismos* (Bogotá: Instituto Caro y Cuervo de Bogotá, 1993).
- Haile, B. *A Stem Vocabulary of the Navaho Language*, 2 vols. (St Michaels, AZ: St Michaels, 1950–1).
- Halāyudha, *Abhidhānaratnamāla* = *Halayudha's Abhidhanaratnamala: A Sanskrit Vocabulary*, ed. T. Aufrecht (London: Williams & Norgate, 1861).
- Halbertsma, J. *Lexicon Frisicum*, ed. T. Halbertsma, vol. I only (A–Feer) (The Hague: Nijhoff, 1872).
- Hale, H. E. *Ethnography and Philology = United States Exploring Expedition During the Years 1838, 1839, 1840, 1841, 1842, Under the Command of Charles Wilkes, USN*, vol. VI (Philadelphia: Sherman, 1846).
- Ḥalīl, 'Ayn = Abū 'Abd al-Rahmān al-Ḥalīl b. Aḥmad al-Farāhīdī, *Kitāb al-'Ayn*, ed. M. al-Maḥzūmī and I. al-Sāmarrā'ī, 8 vols. (Baghdad: Dār al-Rašīd, 1980–5).
- Halliwell-Phillipps, J. O. *A Dictionary of Archaic and Provincial Words, Obsolete Phrases, Proverbs, and Ancient Customs from the Fourteenth Century*, 2 vols. (London: Smith, 1847; 1889).
- Hansen, K. C. *Luritja Picture Dictionary* (Alice Springs: IAD, 2011).
- Hansen, K. C., and L. E. Hansen. *Pintupi Dictionary* (Darwin: Summer Institute of Linguistics, Australian Aborigines Branch, 1974; 1977 and 1992, both as *Pintupi/Luritja Dictionary*).
- Hanxleden, J. E. 'Vocabolario malabarico lusitano' (Barcelona: Biblioteca de Catalunya, Ms. 1484, fos. 37–175).
- Hànyǔ dà cídiǎn* 汉语大词典, ed. Luó Zhúfēng, 13 vols. (Shanghai: Hànyǔ dà cídiǎn chūbǎnshè, 1986–94).
- Hànyǔ dà zìdiǎn* 汉语大字典, 8 vols. (Chengdu: Sìchuān císhū chūbǎnshè; Wuhan: Húběi císhū chūbǎnshè, 1986–90).
- Harawī, *Ġarībayn* = Abū 'Ubayd Aḥmad b. Muḥammad al-Harawī, *K. al-Ġarībayn fī l-Qur'ān wa-l-Ḥadīṭ*, ed. A. F. al-Mazyadī, 6 vols. (Mecca: Maktabat Nizār Muṣṭafā l-Bāz, 1999).
- Ḥarbī, *Ġarīb* = Abū Ishāq Ibrāhīm b. Ishāq, *Ġarīb al-Ḥadīṭ*, ed. S. b. Ibrāhīm b. Muḥammad al-'Āyid, 3 vols. (Mecca: Dār al-Madanī, 1985).
- Ḥarīrī, *Durra* = Abū Muḥammad al-Qāsim b. 'Alī al-Ḥarīrī, *Durrat al-ḡawwās fī awhām al-ḥawāṣṣ*, ed. M. Abū l-Faḍl Ibrāhīm (Cairo: Dār Nahḍat Miṣr, 1975).
- The Harley Latin–Old English Glossary*, ed. R. T. Oliphant (The Hague: Mouton, 1966).
- Harṣakīrti. *Śāradyākhyā-nāmamālā*, ed. M. M. Patkar (Pune: Deccan College Postgraduate and Research Institute, 1951).
- Hartmann, A. M. *English–Mashona Dictionary with Appendix of Some Phrases* (Cape Town: Juta, 1894).
- Ḥasan, N. *Nūr al-lughāt*, reprint edn, 4 vols. (Karachi: Janral, 1959; 1922–31).
- Hatzfeld, A., and A. Darmesteter, with A. Thomas. *Dictionnaire général de la langue française*, 2 vols. (Paris: Delagrave, 1889–1901).
- Ḥayyūj, Yehuda. *Two Treatises on Verbs Concerning Feeble and Double Letters by R Jehuda Ḥayyūj of Fez, Translated into Hebrew from the Original Arabic by R. Moses Ġikātilia of Cordova, to which is added the Treatise on Punctuation by the same Author, Translated by Aben Azra*, ed. and trans. J. W. Nutt (London and Berlin: Asher, 1870).
- Ḥayyūj, Yehuda. *The Weak and Geminative Verbs in Hebrew by Abu Zakariyya Yahya ibn Dawūd of Fez known as Hayyug: The Arabic text Published for the First Time*, ed. M. Jastrow (Leiden: Brill, 1897).

- Heath, J. *Nunggubuyu Dictionary* (Canberra: Australian Institute of Aboriginal Studies, 1982).
- Heckewelder, J. G. E. 'Mahicanni words, taken down from the mouth of one of that nation, who had been born in Connecticut' (Philadelphia: American Philosophical Society, Mss. 497.V85, item 19).
- Hemacandra, *Abhidhānacintāmaṇi* = *Hemakāndra's Abhidhānakintāmaṇi*, ed. O. Boehtlingk and C. Rieu (St Petersburg: Kaiserliche Akademie der Wissenschaften, 1847).
- Hemacandra, *Abhidhānacintāmaṇīkā* = *The Abhidhanachintamani of Kalikala Sarvagna Shri Hemachandracharya (with His Own Notes)*, ed. Hargovindas and Behechardas, 2 vols. (Bhavnagar: Vidya Vijay, 1914–19).
- Hemacandra. *Deśināmamālā*, ed. R. Pischel, 2nd edn, revised by P. K. Ramanujaswami (Pune: Bhandarkar Oriental Research Institute, 1938; 1880).
- Herensch, G. *Teütsche Sprach vnd Weißheit* (Augsburg: Frank, 1616).
- Herennius Philo. *De diversis verborum significationibus*, ed. V. Palmieri (Naples: D'Auria 1988).
- Hershberger, H. D., and R. Hershberger. *Kuku-Yalanji Dictionary* (Darwin: Summer Institute of Linguistics, 1982).
- Hesychius. *Dictionary*, ed. M. Musurus (Venice: Manutius, 1514).
- Hesychius. *Lexicon*, vol. I, ed. K. Latte (Copenhagen: Munksgaard, 1953; 2nd edn by I. C. Cunningham (Berlin and Boston: De Gruyter, 2018); vol. II, ed. K. Latte (Copenhagen: Munksgaard, 1966); vol. III, ed. P. A. Hansen (Berlin and New York: De Gruyter, 2005); vol. IV, ed. P. A. Hansen and I. C. Cunningham (Berlin and New York: De Gruyter, 2009).
- Higgins, N. S. 'Ethnographic notes and vocabulary of Apache tribes of Arizona Territory ... collected from captives belonging to the Chiricahua, Sierra Blanca, Pinal and Coyotero tribes' (Washington, DC: National Anthropological Archives, Smithsonian Institution, MS 180).
- [Hill, F., et al.]. *Western Apache Dictionary* (Fort Apache, AZ: White Mountain Apache Culture Center, 1972).
- Hilton, K. S. *Diccionario tarahumara de Samachique* = *Vocabularios indígenas*, 101 (Tucson: Instituto Lingüístico de Verano, 1993).
- Hilton, K. S., with R. López and E. Carrasco. *Tarahumera y español* = *Vocabularios indígenas*, 1 (Mexico City: Instituto Lingüístico de Verano, 1959).
- Hince, B. *The Antarctic Dictionary: A Complete Guide to Antarctic English* (Collingwood, Australia: CSIRO, 2000).
- Hindī-Asamīyā kośa; Hindī-Asamīyā-Hindī kā eka mātra śabda kośa* (Guwahati: Asama Rāshṭrabhāṣā Pracāra Samiti, 1965).
- Hinds, M., and El-S. M. Badawi. *A Dictionary of Egyptian Arabic: Arabic-English* (Beirut: Librairie du Liban, 1986).
- Hitti, Y. K., and A. S. al-Khatib. *Hitti's New Medical Dictionary: English-Arabic, with an Arabic-English Glossary* (Beirut: Librairie du Liban, 1989).
- Hodge, J. W. *Corean Words and Phrases: Being a Handbook and Pocket Dictionary for Visitors to Corea and New Arrivals in the Country*, 2nd edn (Seoul: Seoul Press, 1902).
- Hoff, A. B. C., and L. C. Hoff. *The Hoff Vocabularies of Indigenous Languages, From the Far West Coast of South Australia* (Kingswood, Australia: Hoff, 2004).
- Hogarth, R. (ed. and trans.). *Gazophylacium Anglicanum* (London: Taylor, 1689).

- Hojjer, H. 'San Carlos Apache linguistic notes', 4 vols. (manuscript in the estate of Keith H. Basso; copy in the Arizona State Museum Library, University of Arizona).
- Holdermann, J. B. *Grammaire Turque ou methode courte & facile pour apprendre la langue Turque* (Constantinople: [Mütefferika], 1730).
- Holm, J., and A. W. Shilling. *Dictionary of Bahamian English* (Hyde Park, NY: Lexik House, 1982).
- Hoogland, J. C. *Woordenboek Arabisch–Nederlands* (Amsterdam: Bulaaq, 2003).
- Hoogland, J. C. *Woordenboek Nederlands–Arabisch* (Amsterdam: Bulaaq, 2003).
- Hopi Dictionary/Hopiikwa Laväytutuveni: *A Hopi Dictionary of the Third Mesa Dialect with an English–Hopi Finder List and a Sketch of Hopi Grammar*, ed. K. C. Hill et al. (Tucson: University of Arizona Press, 1998).
- Houaiss, A., and M. de S. Villar. *Dicionário Houaiss da Língua Portuguesa* (Rio de Janeiro: Objetiva, 2001).
- Hough, W. (ed.). 'Vocabulary collected from the White Mountain Apache by an Army Hospital sergeant about 1880' (Washington: National Anthropological Archives, Smithsonian Institution, MS 3316; typed copy no. 1783).
- Houtman, F. de. *Spraecck ende Woord-Boeck inde Maleysche ende Madagaskarsche Talen* (Amsterdam: Cloppenburch, 1603).
- Houtsma, M. Th. (ed.). *Ein türkisch-arabisches Glossar* (Leiden: Brill, 1894).
- Hrinčenko, B. *Slovar' ukrains'koji movy*, 4 vols. (Kiev: Kievskaya Starina, 1907–9).
- Hu Yinglin. 'Huayang boyi' 華陽博議, in *Shaoshi shanfang bicong* 少室山房筆叢 (Beijing: Zhonghua shu ju, 1958).
- Hudson, J., and E. Richards, with P. Siddon, P. Skipper, et al. *The Walmatjari: An Introduction to the Language and Culture* (Darwin: Summer Institute of Linguistics, Australian Aborigines Branch, 1978).
- Hughes, E. J. *Nunggubuyu–English Dictionary*, 2 vols. (Sydney: Oceania, 1971).
- Huguet, E. *Dictionnaire de la langue française du seizième siècle*, 7 vols. (Paris: Champion, 1925–67).
- Humphreys, D. *The Yankey in England: A Drama in Five Acts* ([Connecticut: n.p., 1815?]).
- Hunter, J. *The Hallamshire Glossary* (London: Pickering, 1829).
- Ibn al-Anbārī, *Aḍḍād* = Abū Bakr Muḥammad b. al-Qāsim al-Anbārī, *al-Aḍḍād*, ed. M. Abū l-Faḍl Ibrāhīm (Kuwait: Dā'irat al-Maṭbū'āt wa-l-Našr, 1960).
- Ibn al-Anbārī, *Nuzha* = Abū l-Barakāt 'Abdalrahmān b. Muḥammad al-Anbārī, *Nuzhat al-alibbā' fī ṭabaqāt al-udabā'*, ed. I. al-Sāmarrā'ī (Baghdad: Maktabat al-Andalus, 1970).
- Ibn Durayd, *Ġamhara* = Abū Bakr Muḥammad b. al-Ḥasan ibn Durayd, *Ġamhrat al-luġa*, ed. R. M. Ba'albakī, 3 vols. (Beirut: Dār al-'Ilm li-l-Malāyīn, 1987–8).
- Ibn Durustawayhi, *Taṣḥīḥ* = Abū Muḥammad 'Abdallāh b. Ġa'far ibn Durustawayhi, *Taṣḥīḥ al-Faṣīḥ*, ed. 'A. al-Ġubūrī, vol. I. (Baghdad: Maṭba'at al-Irṣād, 1975).
- Ibn Fāris, *Maqāyīs* = Abū l-Ḥusayn Aḥmad ibn Fāris, *Muġam Maqāyīs al-luġa*, ed. 'A. M. Hārūn, 6 vols. (Cairo: Dār l-ḥyā' al-Kutub al-'Arabiyya, 1946–).
- Ibn Fāris, *Muġmal* = Abū l-Ḥusayn Aḥmad ibn Fāris, *Muġmal al-luġa*, ed. Z. 'A. Sulṭān, 4 vols. in 2 [beginning of vol. IV not indicated] (Beirut: Mu'assasat al-Risāla, 1984).
- Ibn Fāris, *Ṣāḥibī* = Abū l-Ḥusayn Aḥmad ibn Fāris, *al-Ṣāḥibī fī fiqh al-luġa wa-sunan al-'Arab fī kalāmihā*, ed. M. l-Šuwaymī (Beirut: Mu'assasat Badrān, 1963).
- Ibn Ḥālawayhi, *Laysa* = Abū 'Abdallāh al-Ḥusayn b. Aḥmad ibn Ḥālawayhi, *Laysa fī kalām al-'Arab*, ed. A. 'A. 'Aṭṭār, 2nd edn (Beirut: Dār al-'Ilm li-l-Malāyīn, 1979).

- ibn Janāḥ, Yona, *Kitāb al-Mustalḥaq* = *Sefer ha-Hassaga hu Kitāb al-Mustalḥaq lə-Rabbi Yona ibn Janāḥ, bə-targumo šel 'Ovadia ha-Sep̄haradi, mahdura biqqortit 'im mavo, he'arot u-mafteah pəsuqim*, ed. D. Téné (Jerusalem: Academy of Hebrew Language, 2006).
- ibn Janāḥ, Yona, *Kitāb al-ʿUṣūl* = *The Book of Hebrew Roots by Abu 'L-Walid Marwān ibn Janāḥ or Rabbi Yōnāḥ*, ed. A. Neubauer (Oxford: Clarendon, 1875).
- ibn Janāḥ, Yona. *Opusculs et traités d'Abou'l-Walid Merwan ibn Janah*, ed. J. Dérenbourg and H. Dérenbourg (Paris: Imprimerie Nationale, 1880).
- ibn Janāḥ, Yona, *Sefer ha-šorašim* = *Sepher Haschoraschim: Wurzelwörterbuch der hebräischen Sprache von Abulwalid Merwān ibn Ġanāḥ (R. Jona), aus dem Arabischen in's Hebräische übersetzt von Jehuda ibn Tibbon*, ed. W. Bacher (Berlin: Itskovski, 1896).
- Ibn Mālik, *I'timād* = Ġamāl al-Dīn Abū 'Abdallāḥ Muḥammad b. 'Abdallāḥ ibn Mālik, *al-I'timād fī naẓā'ir al-zā' wa-l-ḍād*, ed. Ḥ. Ṣ. al-Ḍāmin, 2nd edn (Beirut: Mu'assasat al-Risāla, 1984).
- Ibn Manẓūr, *Lisān* = Ġamāl al-Dīn Abū l-Faḍl Muḥammad b. Mukram ibn Manẓūr, *Lisān al-'Arab*, 15 vols. (Beirut: Dār Ṣādir, 1968).
- ibn Parḥon, Shelomo, *Maḥberet* = *Salomonis ben Abrahami Parchon Aragonensis, Lexicon Hebraicum*, ed. S. G. Stern (Bratislava: Schmid, 1844).
- Ibn al-Qaṭṭā', *Af'āl* = Abū l-Qāsim 'Alī b. Ġa'far ibn al-Qaṭṭā', *al-Af'āl*, 4 vols. (Hyderabad: Dā'irat al-Ma'ārif al-'Uṭmāniyya, 1360–4 AH).
- ibn Quraysh, Yehuda. *Risāla*, ed. D. Becker (Tel-Aviv University, 1977).
- Ibn al-Šaġarī, *Mā ttafaqa* = Ḍiyyā' al-Dīn Abū l-Sa'ādāt Hibatallāḥ b. 'Alī ibn al-Šaġarī, *Mā ttafaqa lafẓuhu wa-ḥtalafa ma'nāhu*, ed. 'A. Rizq (Stuttgart: Steiner, 1992).
- Ibn Sallām, *Ṭabaqāt* = Abū 'Abdallāḥ Muḥammad ibn Sallām al-Ġumaḥī, *Ṭabaqāt fuḥūl al-šu'arā*, ed. M. M. Šākir, 2nd edn, 2 vols. (Ġudda: Dār al-Madani, 1980).
- Ibn Sida, *Muḥaṣṣaṣ* = Abū l-Ḥasan 'Alī b. Ismā'īl ibn Sida, *al-Muḥaṣṣaṣ*, 17 vols. (Būlāq: n.p., 1316–21 AH).
- Ibn Sida, *Muḥkam* = Abū l-Ḥasan 'Alī b. Ismā'īl ibn Sida, *al-Muḥkam wa-l-muḥīt al-a'zam*, ed. 'A. Hindāwī, 11 vols. (Beirut: Dār al-Kutub al-'Ilmiyya, 2000).
- Ibn al-Sikkīt, *Iṣlāḥ* = Abū Yūsuf Ya'qūb b. Iṣḥāq ibn al-Sikkīt, *Iṣlāḥ al-manṭiq*, ed. A. M. Šākir and 'A. M. Hārūn, 2nd edn (Cairo: Dār al-Ma'ārif, 1956).
- Ibn Wallād, *Maqṣūr* = Abū l-'Abbās Aḥmad b. Muḥammad ibn Wallād, *al-Maqṣūr wa-l-mamdūd*, ed. P. Brönnle (London: Luzac; Leiden: Brill, 1900).
- Inbal, S. *Milon ivri–angli/Hebrew–English Dictionary* (Jerusalem: Zack, 1988).
- Inshā, M. I. K. *Daryā-yi laṭāfat*, trans. P. B. D. 'Kaifi' (New Delhi: Anjuman-i Taraqqī-yi Urdū (Hind), 1988).
- Introito e porta* (Venice: von Rottwil, 1478).
- Iohannis Zonarae lexicon*, ed. J. A. H. Tittmann et al., 2 vols. (Leipzig: Crusius, 1808).
- Iṣfahānī, *Durra* = Abū 'Abdallāḥ Ḥamza b. al-Ḥasan al-Iṣfahānī, *al-Durra al-fāḥira fī l-amṭāl al-sā'ira*, ed. 'A. Qaṭāmiš, 2 vols. (Cairo: Dār al-Ma'ārif, 1972–6).
- Ishikawa Masamochi. *Gagen shūran* = *Zōho gagen shūran*, ed. Nakajima Hirotsu, 57 fascicles (Tokyo: Nakajima Tadaichi, 1887; 1826–49 [incomplete]; reprint, 3 vols., Tokyo: Kōeki tosho, 1903; reprint, 4 vols., Kyoto: Rinsen shoten, 1965).
- Isidore of Seville. *Etymologies*, ed. and trans. S. Barney et al. (Cambridge University Press, 2006).
- Islex*, ed. Þ. Úlfarsdóttir, islex.is.

- Iversen, E. (ed.), *Papyrus Carlsberg No. VII: Fragments of a Hieroglyphic Dictionary* = Kongelige Danske Videnskabernes Selskab, Historisk-filologiske Skrifter, II. 2 (Copenhagen: Kongelige Danske Videnskabernes Selskab, 1958).
- Jacobson, S. A. *Yup'ik Eskimo Dictionary*, 2nd edn, 2 vols. (Fairbanks: Alaska Native Language Center, University of Alaska, 2012; 1984).
- Jakobsen, J. *Etymologisk Ordbog over det norrøne sprog på Shetland*, 4 vols. (Copenhagen: Prior, 1908–21).
- Jakobsen, J. 'Ordsamling' in V. U. Hammershaimb (ed.), *Færøsk anthologi*, 2 vols. (Copenhagen: Møller, 1891), II.1–418.
- Jakubaš, F. *Hornjoserbski-němski slownik* (Budyšin: Domowina, 1954).
- Jambūdvīpaprājñaptīkā* (Mumbai: Naginbhai Ghelabhai Jhaveri, 1920).
- Jamieson, J. An *Etymological Dictionary of the Scottish Language*, 2 vols. (Edinburgh: Creech et al., 1808).
- Jamieson, J. *Supplement to the Etymological Dictionary of the Scottish Language*, 2 vols. (Edinburgh: Tait, 1825).
- Jauncey, D. *Bardi Grubs and Frog Cakes: South Australian Words* (South Melbourne: Oxford University Press, 2004).
- Jenssøn, C. *Den Norske dictionarium eller Glosebog* (Copenhagen: Anderssøn, 1646).
- Jetté, J., and E. Jones. *Koyukon Athabaskan Dictionary* (Fairbanks: Alaska Native Language Center, University of Alaska, 2000).
- Jidaietsu kokugo daijiten: Jōdai hen* (Tokyo: Sanseidō, 1967).
- Jidaietsu kokugo daijiten: Muromachi jidai hen*, 5 vols. (Tokyo: Sanseidō, 1985–2000).
- Jñānaśrīmitra. *Vṛttamālāstuti, eine Beispielsammlung zur altindischen Metrik nach dem tibetischen Tanjur zusammen mit der mongolischen Version*, ed. M. Hahn (Wiesbaden: Harrassowitz, 1971).
- Johnson, Francis (ed.). *A Dictionary, Persian, Arabic, and English* (London: Allen, 1852).
- Johnson, Frederick. *Kamusi ya Kiswahili yaani kitabu cha maneno ya Kiswahili* (London: Sheldon, 1935).
- Johnson, S. *A Dictionary of the English Language*, 2 vols. (London: Knapton et al., 1755).
- Johnson on the English Language*, ed. G. J. Kolb and R. DeMaria (New Haven: Yale University Press, 2005).
- Johnson's English Dictionary in Miniature*, ed. J. Hamilton (London: Bell, 1836).
- Jones, W. *A Grammar of the Persian Language* (London: Richardson, 1771).
- Joseon-eo Sajeon* (Seoul: Joseonchongdogbu, 1920).
- Joseonmal Keunsajeon*, 6 vols. [vols. IV–VI published as *Keunsajeon*] (Seoul: Eul-yumunhwasa, 1947–57).
- Jungmann, J. *Slovník česko-německý*, 5 vols. (Prague: w kněžecí arcibiskupské knihtiskárně, u Josefy wdowy Fetterlowé, řizenjm Wáclawa Špinky, 1835–9).
- Junius, F. 'Glossarium gothicum', separately paginated supplement to F. Junius and T. Marshall (eds.), *Quatuor D. N. Jesu Christi euangeliorum versiones perantiquae duae* (Dordrecht: Essaei, 1665).
- Junod, H. A. *Grammaire ronga, suivie d'un manuel de conversation et d'un vocabulaire ronga-portugais-français-anglais, pour exposer et illustrer les lois du ronga, langage parlé par les indigènes du district de Lourenço-Marques* (Lausanne: Bridel, 1896).
- Juška, A. *Litovskij slovar' A. Juškeviča s tolkovaním slov na ruskom i pol'skom jazykax*, 3 vols. (St Petersburg: [imprint varies], 1897–1922).

- Kaṅkēyar. *Uric col nikaṇṭu* (Chennai: Cānti Cātaṇā, 2003).
- Karadžić, V. *Dodatak k sanktpeterburgskim sravniteljnim rječnicima sviju jezika i narječija i s osobitim ogledima bugarskog jezika* (Vienna: n.p., 1822).
- Karadžić, V. *Srpski rječnik istolkovan njemackim i latinskim riječima* (Vienna: bei den P.P. Armeniern, 1818).
- Karadžić, V. *Srpski rječnik istumačen njemackijem i latinskijem riječima* (Vienna: Štamparija jermenskog manastira, 1852; ³1898).
- Karłowicz, J. A. L. *Słownik gwar polskich*, 6 vols. (Kraków: Nakładem Akademii umiejętności, 1900–11).
- Karłowicz, J. A. L., A. A. Kryński, and W. M. Niedźwiedzki. *Słownik języka polskiego*, 8 vols. (Warsaw: Nakładem prenumeratorów i Kasy Mianowskiego, 1900–27).
- Kāśikā = Kāśikavṛtti [sic] of Jayāditya-Vāmana: *Along with Commentaries Vivaraṇapañcīkā-Nyāsa of Jinendrabuddhi and Padamañjarī of Haradatta Miśra*, ed. S. Misra, 6 vols. (Varanasi: Ratna, 1985).
- Katakī, D. *Natuna Asamīyā abhidhāna* (Guwahati: Laṭyārcha Buka Shṭāla, 1996).
- Kayātaram*, ed. S. Vaiyāpuri Pillai = Bulletin of the Tamil Department no. 4 (Chennai: Madras University, 1939).
- Kelham, R. *A Dictionary of the Norman or Old French Language* (London: Brooke, 1779).
- Kennedy, V. *Dictionary of the Marat,ha Language in Two Parts* (Mumbai: Courier Press, 1824).
- Kenrick, W. *A New Dictionary of the English Language* (London: Rivington et al., 1773).
- Khan, M. M. T. *Dictionary of Law: English to Urdu* (Lahore: Kashmir Law Times, 2004).
- Khān, Z. A.-D. A. *Akḥbār lughāt ma ‘rūf bēh kalīd-i akḥbār bīnī* (Delhi: Delhi Printing Works, 1915).
- al-Khatib, A. S. *A New Dictionary of Scientific and Technical Terms, English–Arabic* (Beirut: Librairie du Liban, 1971).
- Kholar, V. L., and D. U. Maslai. *Tivā mātpāḍiḥ: A Tri-Lingual Dictionary, Tiwa–Assamese–English* (Guwahati: Tiwa Mathonlai Tokhra [Tiwa Sahitya Sabha], 2004).
- Kieffer, J. D., and T. X. Bianchi. *Dictionnaire turc–français*, 2 vols. (Paris: Imprimerie Royale, 1835–7).
- Kilham, C., M. Pamulkan, J. Pootchemunka, and T. Wolmby. *Dictionary and Source Book of the Wik-Mungkan Language* (Darwin: Summer Institute of Linguistics, Australian Aborigines Branch, 1986).
- Kiliaan, C. *Dictionarium Teutonico–Latinum* (Antwerp: Plantin, 1574).
- Kiliaan, C. *Etymologicum Teutonicae linguae, sive Dictionarium Teutonico–Latinum* (Antwerp: Plantin, 1599).
- Kimchi, David. *Sefer ha-šorašim/Radicum liber sive Hebraeum Bibliorum lexicon*, ed. J. H. R. Biesenthal and F. Lebrecht (Berlin: Bethge, 1847).
- Kircher, A. *Lingua Aegyptiaca restituta* (Rome: Scheus, 1643).
- Kisā’ī, Mā talḥan = Abū l-Ḥasan ‘Alī b. Ḥamza al-Kisā’ī, *Mā talḥan fīhi l-‘amma*, ed. R. ‘Abdaltawwāb (Cairo: Maktabat al-Ḥānṣī; Riyad: Dār al-Rifā’ī, 1982).
- Kleinschmidt, S. P. *Den grønlandske Ordbog*, ed. H. F. Jorgensen (Copenhagen: Klein, 1871).
- Knaani, Y. *Otsar halashon haivrit litkufoteyha hashonot*, 18 vols. (Jerusalem [place of publication varies]: Massada, 1960–89).
- Koelle, S. W. *Polyglotta Africana, or, A Comparative Vocabulary of Nearly Three Hundred Words and Phrases, in More than One Hundred Distinct African Languages* (London: Church Missionary House, 1854).

- Kogo taikan, ed. Tsukishima Hiroshi et al., 2 vols. of a projected 4 (Tokyo: Tokyo daigaku shuppankai, 2011–).
- Kōjien, ed. Shinmura Izuru (Tokyo: Iwanami shoten, 1955; ⁷2018).
- Kokovzov, P. *Novye materialy dlja xarakteristiki Exudi Xajūza, Samuila Nagida i nekotorix drugix predstavitelej evrejskoj filologičeskoj nauki v-x, xi i xii veki* (St Petersburg: Imperatorskaja Akademija nauk, 1916).
- Kong Julian. *Wuerdu yu Han yu ci dian* (Beijing: Gao deng jiao yu chu ban she, 2014).
- Krahl, G. *Deutsch–Arabisches Wörterbuch* (Leipzig: VEB, 1964).
- Krahl, G., and G. M. Gharieb. *Wörterbuch Arabisch–Deutsch* (Munich: Hueber, 1984).
- Kramer, J. (ed.). *Etymologisches Wörterbuch des Dolomitenladinischen*, 8 vols. (Hamburg: Buske, 1988–98).
- Kramer, J. (ed.). *Glossaria bilingua altera* = Archiv für Papyrussforschung und verwandte Gebiete, Beiheft 8 (Munich and Leipzig: Saur, 2001).
- Kramer, J. (ed.). *Glossaria bilingua in papyris et membranis reperta* = Papyrologische Texte und Abhandlungen, 30 (Bonn: Habelt, 1983).
- Kramer, M. *Het koninglyk neder–hoog-duitsch en hoog–neder-duitsch Dictionnaire/Das Königliche Nieder–Hoch-Teutsch, und Hoch–Nieder-Teutsch Dictionarium*, 2 vols. (Nürnberg: for the author, 1719).
- Kramuvunt [Kramvant], J. S., with B. Ghagve et al. *A Dictionary of the Maratha Language* (Mumbai and Pune: Bombay Native Education Society, 1829).
- Kramuvunt [Kramvant], J. S., T. Candy, and J. T. Molesworth. *Supplement to the Dictionary of the Maratha Language* (Pune: Department of Public Instruction, 1831).
- Krapf, J. L. *A Dictionary of the Suahili Language, with Introduction Containing an Outline of a Suahili Grammar* (London: Trübner, 1882).
- Krapf, J. L. *Vocabulary of Six East-African Languages (Kisúaheli, Kiníka, Kikámbe, Kipokómo, Kihíáu, Kigállá)* (Tübingen: Fues, 1850).
- Krapf, J. L. *Vocabulary of the Galla Language*, ed. C. W. Isenberg (London: Church Missionary Society, 1842).
- Kriol Dictionary, with Kriol to English and English to Kriol* (Darwin: Summer Institute of Linguistics, Australian Aborigines and Islanders Branch, 1986).
- Kriyāvin̄ Tarkālat Tamil̄ akarāti: Tamil̄–Tamil̄–Āṅkilam/Dictionary of Contemporary Tamil (Tamil–Tamil–English), ed. S. Ramakrishnan (Madras: Cre-A, 1992; ²2008).
- Kro, B. S. *Akemi Karbi lamthe amarjong: A Karbi to Karbi–English–Assamese Dictionary* (Diphu: Krohemtun, 2002).
- Kropf, A. *A Kaffir–English Dictionary* (Lovedale, South Africa: Lovedale Mission, 1899).
- Kropfitch, L. *Langenscheidt Handwörterbuch Arabisch–Deutsch* (Berlin: Langenscheidt, 1996).
- Kropfitch, L. *Langenscheidt Taschenwörterbuch Arabisch* (Berlin: Langenscheidt, 1998).
- Krūnitz, J. G. *Oeconomische Encyclopädie, oder allgemeines System der Land-, Haus-, und Staats-Wirthschaft, in alphabetischer Ordnung* [title varies], 242 vols. (Berlin: Pauli [imprint varies], 1773–1858), www.kruenitz1.uni-trier.de/.
- Kun dga' rgyal mtshan, *Tshig gi gter* = mNgon brjod kyi bstan bcos *Tshig gi gter zhes bya ba'i 'grel pa rgyal cher don gsal ba* (Gangtok: Gonpo Tseten, 1977), tbrc.org.
- Kurā', *Munaḡḡad* = Abū l-Ḥasan 'Alī b. al-Ḥusayn al-Hunā' ī Kurā' al-Naml, *al-Munaḡḡad fi l-luḡa*, ed. A. M. 'Umar and Ḍ. 'Abdalbāqī, 2nd edn (Cairo: 'Ālam al-Kutub, 1988).
- Kurmin, J. *Słownik polsko łacinsko lotewski* (Vilnius: Zymelowicz, 1858).

- Kurschat, A. *Litauisch-deutsches Wörterbuch/Thesaurus linguae Lituanicae*, ed. W. Wissmann, E. Hofmann, et al., 4 vols. (Göttingen: Vandenhoeck & Ruprecht, 1968–73).
- Kurschat, F. *Grammatik der litauischen Sprache* (Halle: Buchhandlung des Waisenhauses, 1876).
- Kurschat, F. *Wörterbuch der litauischen Sprache*, part 1, *Deutsch-litauisches Wörterbuch*, 2 vols., and part 2, *Litauisch-deutsches Wörterbuch* (Halle: Buchhandlung des Waisenhauses, 1870–4 and 1883).
- Labèrnia, P. *Diccionari de la llengua catalana*, 2 vols. (Barcelona: Hereus de la Vidua Pla, 1839–40).
- Lacombe, A. *Dictionnaire de la langue des Cris* (Montreal: Beauchemin & Valois, 1874).
- Lagadeuc, J. *Le Catholicon*, ed. C.-J. Guyonvarc'h, *Celticum* 22 (supplement to *Ogam-Tradition Celtique* 27), 1975 (reprint Brest: Éditions Armeline, 2005).
- Lagunas, J. B. de. *Arte y diccionario con otras obras en lengua michuacana* (Mexico City: Balli, 1574).
- Lane, E. W. *An Account of the Manners and Customs of the Modern Egyptians*, 2 vols. (London: Knight, 1836).
- Lane, E. W. *An Arabic-English Lexicon, Derived from the Best and the Most Copious Eastern Sources*, 8 vols., vols. VI–VIII ed. S. Lane-Poole (London: Williams & Norgate, 1863–93; reprint Beirut: Librairie du Liban, 1997).
- Lane, E. W. (trans.). *The Thousand and One Nights, Commonly Called, in England, The Arabian Nights' Entertainments*, 3 vols. (London: Knight, 1839–41).
- Lange, J. *Vollständiges deutschlettisches und lettischdeutsches Lexicon, nach den Hauptdialekten in Lief- und Curland ausgefertigt* (Mitau: Steffenhagen, 1777).
- Lara, L. F. (ed.). *Diccionario del Español de México* (2012–), dem.colmex.mx.
- Larousse, P. *Grand dictionnaire universel du XIXe siècle*, 17 vols. (Paris: Larousse, 1865–90).
- Larramendi, M. *Diccionario trilingue del castellano, bascuense, y latin*, 2 vols. (San Sebastian: Riesgo y Montero, 1745).
- Las Casas, C. de. *Vocabulario de las dos lenguas toscana y castellana* (Seville: de Aguilar, 1570).
- Latif, M., A. Hasan, S. Baqal, and A. Rizvi. *A Dictionary of Technical Terms: Sociology, English-Urdu* (Karachi: Bureau of Composition, Compilation, and Translation, University of Karachi, 1970).
- Latinsk–Svenskt glossarium efter Cod. Ups. C20*, ed. E. Neuman and B. Tjäder, 7 fascicles (Uppsala: Svenska Fornskrift-Sällskapet, 1918–94).
- Latviešu valodas literārās vārdnīca*, 8 vols. (Riga: Zinātne, 1972–7).
- Laughlin, R. M. *The Great Tzotzil Dictionary of San Lorenzo Zinacatán* (Washington, DC: Smithsonian Institution, 1975).
- Lebedev, K. A., L. S. Jacevič, and Z. M. Kalinina. *Russko-afganskij slovar' (puštu)* (Moscow: Sovetskaja enciklopedija, 1973).
- [LeBoullenger, A. R.] 'French and Miami-Illinois dictionary' (manuscript at the John Carter Brown Library, Brown University, Providence, RI).
- Leeding, V. J. 'Anindilyakwa dictionary (first draft)' (Canberra, Australian Institute of Aboriginal and Torres Strait Islander Studies, MS 1092).
- Legrand de la Liraye, T. M. *Prononciation figurée des caractères chinois en mandarin annamite*, with calligraphy by Tràn N. H. (Saigon: Collège des Stagiaires, 1875).
- Le Guennec, G., and J. F. Valente. *Dicionário Português-Umbundu* (Luanda: Instituto de Investigação Científica de Angola, 1972).

- Lemoine, G. *Dictionnaire français-algonquin* (Chicoutimi: Delisle, 1909; reissued Quebec City: L'Action Sociale, 1911).
- Lemoine, G. *Dictionnaire français-montagnais, avec un vocabulaire montagnais-anglais, une courte liste de noms géographiques, et une grammaire montagnaise* (Boston: Cabot and Cabot, 1901).
- Lěng Yùlóng and Wéi Yīxīn (eds.). *Zhōnghuá zìhǎi* 中华字海 (Beijing: Zhōnghuá shūjù, Zhōngguó yǒuyì chūbǎn gōngsī, 1994).
- Lengua de Maynas: Ms Egerton 2881 de la British Library*, ed. A. Alexander-Bakkerus (Munich: Lincom, 2014).
- Lenguas de América: manuscritos de la Real Biblioteca* (Madrid: n.p., 1928).
- Le Roux, J.-M. *Essai de dictionnaire français-haoussa et haoussa-français, précédé d'un essai de grammaire de la langue haoussa* (Algiers: Jourdan, 1886).
- Lexicon románescu-látinescu-ungurescu-nemțescu/Lexicon valachico-latino-ungarico-germanicum* (Buda: typographia Regiae Universitatis Ungaricae, 1825).
- LeSourd, P. S. *Kolusuwakonol peskotomuhkati-wolastokewi naka ikolisomani latuwewakon/ Passamaquoddy-Maliseet and English Dictionary*, ed. R. M. Leavitt and D. A. Francis (Fredericton: Micmac-Maliseet Institute, University of New Brunswick, 1984).
- Lessico etimologico italiano*, ed. M. Pfister and W. Schweickard, 128 fascicles to date (Wiesbaden: Reichert, 1979–).
- Le Talleur, Guillaume. *Dictionarius familiaris et compendiosus*, ed. W. Edwards and B. Merrilees (Turnhout: Brepols, 2002).
- Lettres édifiantes et curieuses, écrites des missions étrangères*, ed. Y. M. M. T. de Querbeuf et al., new edn, 14 vols. (Lyon: Vernarel, 1819).
- Le Ver, Firmin. *Dictionarius*, ed. B. Merrilees and W. Edwards (Turnhout: Brepols, 1994).
- Levita, E. *Meturgeman: Lexicon Chaldaicum* (Isny: Fagius, 1541).
- Levita, E. *Tishbi* (Isny: Fagius, 1541).
- Levy, E. *Provenzalisches Supplement-Wörterbuch*, 8 vols. (Leipzig: Reisland, 1894–1924).
- Lexica Graeca minora*, ed. K. Latte and H. Erbse (Hildesheim: Olms, 1965).
- Lexicon Graecolatinum, seu Thesaurus linguae Graecae, post eos omnes qui in hoc commentandi genere hactenus excelluerunt, ex ipsius demum G. Budaei manu scripto Lexico, magna cum dictionum tum elocutionum accessione auctus* (Geneva: Crespin, 1554).
- Lexicon of Cyril = Das sogenannte 'Kyrill'-Lexikon in der Fassung der Handschrift E* (Codex Bremensis G [really 'C'] 11), ed. U. Hagedorn, 2005, kups.ub.uni-koeln.de/1813/1/Hagedorn_Kyrillos_Hauptdatei.pdf.
- Lexicon sive Dictionarium graecolatinum*, ed. R. Constantin (Geneva: Crespin, 1562).
- Lexicons of Early Modern English*, ed. I. Lancashire, leme.library.utoronto.ca/.
- Lexicon Vindobonense*, ed. A. Nauck (St Petersburg: Eggers; Leipzig: Voss, 1867).
- Lǐ Lùxīng. *Xīn HSK 5000 cí fēnjí cídiǎn (1–3 jí)* 新HSK5000词分级词典(1~3级)/*Dictionary of 5000 Graded Words for New HSK* (Beijing: Běijīng Yǔyán Dàxué chūbǎnshè, 2013).
- Liang Shih-ch'iu. *A New Practical English-Chinese Dictionary* (Taipei: Far East Book, 1971).
- Liber glossarum*, ed. A. Grondeux and F. Cinato (2016), liber-glossarum.huma-num.fr.
- Liber ordinis rerum*, ed. P. Schmitt. 2 vols. (Tübingen: Niemeyer, 1983).
- Lietuvių kalbos žodynas*, 20 vols. (Vilnius: Mokslo/Mokslo ir enciklopedijų leidybos institutas, 1941–2002; first 2 fascicles 1924–5).
- Lin Yutang. *Chinese-English Dictionary of Modern Usage* (Chinese University of Hong Kong, 1972).

- Linde, S. B. *Słownik języka polskiego*. 6 vols. (Warsaw: Drukarnia XX. Pilarów, 1807–14).
- Linde, S. B. *Słownik języka polskiego*, 2nd edn, 6 vols. (Lwów: Zakład Narodowy im. Ossolińskich, 1854–60).
- Lisboa, M. de. *Vocabulario de la lengua Bicol* (Sampaloc: Convento de Nuestra Señora de Loreto, 1754).
- Litke, R. L. (ed.). *A Reconstruction of the Assyro-Babylonian God-Lists*, AN: ⁴A-nu-um and AN: Anu ša amēli = Texts from the Babylonian Collection, 3 (New Haven: Yale Babylonian Collection, 1998).
- The Little Oxford Dictionary of Current English*, ed. G. Ostler (Oxford: Clarendon, 1930).
- Litré, É. *Dictionnaire de la langue française*, 2 vols. in 4 parts plus supplement (Paris: Hachette, 1863–72, 1877).
- Lloyd, W. 'An alphabetical dictionary' in J. Wilkins, *Essay Towards a Real Character and a Philosophical Language* (London: Gellibrand and Martin, 1668), sigs. aaaI r–4 v, AaaI r–Ttt3 r.
- Lobscheid, W. *English and Chinese Dictionary: With the Punti and Mandarin Pronunciation* (Hong Kong: Daily, 1866).
- Loderecker, P. *Dictionarium septem diversarum linguarum, videlicet Latine, Italice, Dalmatice, Bohemice, Polonice, Germanice, et Ungarice* (Prague: for the author, 1605).
- Loderus, A. L. (ed.). *Maleische Woord-Boek Sameling/Collectanea Malaica vocabularia*, 2 vols. (Jakarta: Loderus, 1707–8).
- Lombardo, N. *Arte de la lengua teguima, vulgarmente llamada Opata* (Mexico City: de Ribera, 1702).
- Lombardo, N. *Arte de la lengua tegüima, vulgarmente llamada Opata*, ed. I. Guzman Betancourt (Mexico City: Instituto Nacional de Antropología e Historia, 2009).
- Lomonosov, M. 'Predislovie o pol'ze knig cerkovnyx v rossijskom jazyke' in *Sobranie raznyx sočinenij v stixax i prozě* 1, 2nd edn (Moscow: pri Imperatorskom Moskovskom universitete, 1758), 3–10.
- Lomonosov, M. *Rossijskaja grammatika* (St Petersburg: Imperatorskoi Akademii Nauk, 1755).
- The Long Discourses of the Buddha: A Translation of the Dīgha Nikāya*, trans. M. O'C. Walshe (Boston, MA: Wisdom, 1995; 1st edn 1987, as *Thus Have I Heard: The Long Discourses of the Buddha*).
- Loughridge, R. M., and D. M. Hodge. *English and Muskokee Dictionary* (St Louis: Smith, 1890; reprint 1964).
- Love, J. R. B. 'English–Worora vocabulary' (Canberra, Australian Institute of Aboriginal and Torres Strait Islander Studies, MS 17b).
- Love, J. R. B. 'Grammatical structure of the Worora language of north-western Australia', unpublished MA thesis, University of Adelaide (1934).
- Love, J. R. B. *Stone-Age Bushmen of Today: Life and Adventure among a Tribe of Savages in North-Western Australia* (London and Glasgow: Blackie, 1936).
- Love, J. R. B. 'Worora–English vocabulary' (Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies, MS 17a).
- Lù Ěrkuí (ed.). *Ciyuán 辭源*, 2 vols. (Shanghai: Shāngwù Yīnshūguǎn, 1915).
- Lù Gǔsūn (ed.). *Yīng–Hàn dà cídiǎn/The English–Chinese Dictionary*, 2 vols. (Shanghai: Shànghǎi yìwén chūbǎnshè, 1989–91; ²2007).
- Lu Rong. *Shuyuan zaji 菽園雜記* (Beijing: Zhonghua shuju, 1985).

- Ludolf, H. *Lexicon aethiopico-latinum* (London: Roycroft, 1661).
- Lughāt kabīr-i Urdū*, ed. 'A. Ḥaq et al., 22 vols. (Karachi: Anjuman-i Taraqqī-yi Urdū, 1973–2010).
- Luxemburger Wörterbuch*, 5 vols. (Luxemburg: Linden, 1950).
- Lyon, R. M. 'A glance at the manners and language of the Aboriginal inhabitants of Western Australia, with a short vocabulary', *Perth Gazette and Western Australian Journal*, 1 (1833), 51–2, 56, 59–60, 63–4.
- Maaler, J. *Die Teütsch Spraach* (Zürich: Froschauer, 1561).
- Macafee, C. *Concise Ulster Dictionary* (Oxford University Press, 1996).
- MacGowan, J. *English and Chinese Dictionary of the Amoy Dialect* (Xiamen: Marcal, 1883).
- Machoni de Cerdeña, A. *Arte y vocabulario de la lengua lule, y tonocote* (Madrid: Infanzón, 1732).
- MacIver, D. *An English–Chinese Dictionary in the Vernacular of the Hakka People in the Canton Province* (Shanghai: American Presbyterian Mission, 1905).
- Maclay, R. S., and C. C. Baldwin. *An Alphabetic Dictionary of the Chinese Language in the Foochow Dialect* (Fuzhou: Methodist Episcopal Mission, 1870).
- The Macquarie Dictionary*, 7th edn, ed. S. Butler, 2 vols. (Sydney: Macquarie, 2017).
- Maćczyński, J. *Lexicon Latino–Polonicum ex optimis Latinae linguae scriptoribus concinnatum* (Königsberg: Daubmannus, 1564).
- Madan, A. C. *English–Swahili Dictionary, compiled for the use of the Universities' Mission to Central Africa* (Oxford: Clarendon; London: Society for Promoting Christian Knowledge, 1894).
- Madan, A. C. *Swahili–English Dictionary* (Oxford: Clarendon, 1903).
- Magnús Ólafsson, *Specimen lexici runici*, ed. O. Worm (Copenhagen: Martzan, 1650); issued as part of Worm (ed.), *Antiquitates danicae* (Copenhagen: Martzan, [1651]).
- Mahāniddesa*, ed. L. de La Vallée Poussin and E. J. Thomas (London: Pali Text Society, 1978; 1916–17).
- Mahārāṣṭra śābdakośa*, ed. Y. R. Dāte, 7 vols. and supplement (Pune: Mahārāṣṭra Kōsamanḍaḷa, 1932–8 and 1950).
- The Mahāvvyutpatti: Sanskrit–Tibetan–Mongolian Dictionary of Buddhist Terminology*, ed. Yumiko Ishihama and Yoichi Fukuda = *Studia Tibetica*, 16 (Tokyo: Toyo Bunko, 1989).
- Maheśvara. *Viśvaprakāśa*, ed. Ś. Sthavira (Varanasi: Chowkhambā Sanskrit Book Depot, 1911).
- Maḥmūd al-Kāšgarī. *Compendium of the Turkic Dialects (Dīwān Luḡāt at-Turk)*, ed. and trans. R. Dankoff with J. Kelly, 3 vols. (Cambridge, MA: Harvard University, 1982–5).
- Majid, 'A. *Jāmi' al-lughāt*, 4 vols. (Lahore: Malik Dīn Muḥammad Aīnḍ Sanz, 1933).
- Mālaqī, *Raṣf* = Abū Ḡa'far Aḥmad b. 'Abdalnūr al-Mālaqī, *Raṣf al-mabānī fī ṣarḥ ḥurūf al-ma'ānī*, ed. A. M. al-Ḥarrāṭ (Damascus: Maṭba'at Zayd b. Ṭābit, 1975).
- Malaret, A. *Diccionario de americanismos* (Mayaguez: n.p., 1925).
- Mancelius, G. *Lettus, das ist Wortbuch sampt angehengtem täglichem Gebrauch der Lettischen Sprache . . . Phraseologia lettica, das ist täglicher Gebrauch der Lettischen Sprache* (Riga: Schröder, 1638).
- Maṅkha. *Maṅkhakośa*, ed. T. Zachariae (Mumbai: Education Society, 1897).
- Manuale Lettico–Germanicum*, ed. T. G. Fennell, 2 vols. (Riga: Latvijas Akadēmiskā Bibliotēka, 2001).

- Mao Qiling. *Yuèyǔ kěngqīng lù* 越語肯綮錄 = *Xihe heji* 西河合集, vol. 75 (n.p., c. 1662–1722).
- Marak, H. W. *Ku-bidik: A Garo–English–Assamese Dictionary* (Guwahati: Assam Academy for Cultural Relations in association with Dept. of Arts and Culture, Govt. of Meghalaya, 2000).
- Marbán, P. *Arte de la lengua moxa, con su vocabulario, y cathecismo* ([Lima]: de Contreras, 1702).
- Marinoni, A. *Dal ‘Declarus’ di A. Senisio i vocaboli Siciliani* (Palermo: Centro di Studi Filologici e Linguistici Siciliani, 1955).
- Marsh, J. *Martu Wangka–English Dictionary* (Darwin: Summer Institute of Linguistics, 1992).
- Martin, B. *Lingua Britannica reformata* (London: Hodges et al., 1749).
- Masthay, C. (ed.). *Kaskaskia Illinois-to-French Dictionary* [an edition of Gravier (attrib.), ‘Illinois–French dictionary’] (St Louis, MO: for the editor, 2002).
- Materials for the Sumerian Lexicon* [previously published as *Materialien zum sumerischen Lexikon*], ed. B. Landsberger et al., 19 vols. to date (Rome: Pontificium institutum biblicum, 1937–).
- Mathews, R. H. *A Chinese–English Dictionary Compiled for the China Inland Mission* (Shanghai: China Inland Mission, 1931).
- Matta, J. D. C. da. *Ensaio de dicionario Kimbúndu–Portuguez* (Lisbon: Pereira, 1893).
- Mauro, T. de. *Grande dizionario italiano dell’uso*, 8 vols. (Turin: UTET, 1999–2007).
- Mavor, W. F. *Mavor’s Spelling Book, with Its Goojrathee Translation* (Mumbai: Summachar, 1837).
- Mawer, A., and F. M. Stenton. *The Place-Names of Buckinghamshire* (Cambridge University Press for the English Place-Name Society, 1925).
- Maydānī, Maġma ‘ = Abū l-Faḍl Aḥmad b. Muḥammad al-Maydānī, *Maġma ‘ al-amṭāl*, ed. M. Muḥyī l-Dīn ‘Abdalḥamīd, 2 vols. (Cairo: Maṭba‘at al-Sunna al-Muḥammadiyya, 1955).
- Mažiulis, V. (ed.). *Prūsų kalbos paminklai*, 2 vols. (Vilnius: Mokslas, 1966–81).
- Medhurst, W. H. *Chinese and English Dictionary: Containing All the Words in the Chinese Imperial Dictionary, Arranged According to the Radicals*, 2 vols. (Parapattan: [London Missionary Society], 1842–3).
- Medhurst, W. H. *A Dictionary of the Hok-këen Dialect of the Chinese Language, According to the Reading and Colloquial Idioms, Containing About 12,000 Characters* (Macao: Honorable East India Company, 1832).
- Medhurst, W. H. *English and Chinese Dictionary*, 2 vols. (Shanghai: Mission, 1847–8).
- Medinikara. *Nānārtha Śabda Kośa or Medinī Kośa*, ed. J. Ś. Hoṣhīng, 3rd edn (Varanasi: Chowkhamba Sanskrit Series, 1968).
- Megiser, H. *Dictionarium quatuor linguarum, videlicet, Germanicae, Latinae, Illyrica (quae vulgo Sclavonica appellatur) et Italicae, sive Hetruscae* (Graz: Faber, 1592).
- Megiser, H. *Thesaurus polyglottus vel dictionarium multilingue* (Frankfurt am Main: for the author, 1603).
- Mehta, B. B., and B. N. Mehta. *The Modern Gujarati–English Dictionary* (Vadodara: M. C. Kothari, 1925).
- Meikai kogo jiten*, ed. Kindaichi Kyōsuke and Kindaichi Haruhiko (Tokyo: Sanseidō, 1953).
- Meikai kokugo jiten*, ed. Kindaichi Kyōsuke (Tokyo: Sanseidō, 1943).

- Menāḥem ben Saruq. *Maḥberet*, ed. Á. Sáenz-Badillos (Universidad de Granada, 1986).
- Meninski, F. *Thesaurus linguarum orientalium Turcicae-Arabicae-Persicae*, ed. M. Ölmez (Istanbul: Simurg, 2000).
- Méntrida, A. de. *Vocabulario de la lengua bisaya, hiligueyna y haraya de la isla de Panay y Sugbú y para las demás islas*, ed. J. García-Medall (IEIP-Universidad de Valladolid, 2004).
- Merriam-Webster, www.merriam-webster.com/.
- Mertens-Pack3, web.philo.ulg.ac.be/cedopal/database-mp3/.
- Meyer, H. A. E. *Vocabulary of the Language Spoken by the Aborigines of the Southern and Eastern Portions of the Settled Districts of South Australia . . . Preceded by a Grammar* (Adelaide: Allen, 1843).
- Meyer-Lübke, W. *Romanisches etymologisches Wörterbuch* (Heidelberg: Winter, 1911).
- Mgharfaoui, K., A. Chekayri, and A. Mabrou. *Qāmūs al-dārīḡa al-maḡribīya* (Casablanca: Centre de promotion de la darija, fondation Zakoura, 2017).
- Micul dicționar academic, ed. M. Sala and I. Dănăilă, 4 vols. (Bucharest: Univers enciclopedic, 2001–3).
- Middle English Dictionary, ed. H. Kurath, S. M. Kuhn, R. E. Lewis, et al., 10 vols. (Ann Arbor: University of Michigan Press, 1952–2001).
- The Middle Length Discourses of the Buddha: A New Translation of the Majjhima Nikāya*, trans. Bhikkhu Ñāṇamoli, ed. and revised by Bhikkhu Bodhi, 3rd edn (Kandy, Sri Lanka: Buddhist Publication Society, 2002).
- Mielcke, C. G. *Littauisch-deutsches und deutsch-littauisches Wörterbuch* (Königsberg: Hartung, 1800).
- Mikalja, J. *Blago jezika slovinskoga illi Slovník u komu izgovarajuse rjeci slovinske, latinski, i diacki/Thesaurus linguae Illyricae, sive dictionarium Illyricum, in quo verba Illyrica Italice et Latine redduntur*, 2 parts (part 1, Loreto: Serafini, 1649; part 2, Ancona: Beltrano, 1651).
- Miklosich, F. *Lexicon Palaeoslovenico-Graeco-Latinum*, new edn (Vienna: Braumueller, 1862–5; 1850).
- Miller, E. (ed.). *Mélanges de littérature grecque, contenant un grand nombre de textes inédits* (Paris: Imprimerie Imperiale, 1868).
- Milon *sapir mekuvan*, www.milononline.net/.
- Minsheu, J. *Ductor in linguas: The Guide into Tongues* (London: [for the author], 1617).
- Mischlich, A. *Wörterbuch der Hausasprache* (Berlin: Reimer, 1906).
- Mishra, B. V. *An English-Hindi Vocabulary in Verse*, 2nd edn (Kolkata: Gunsindhu, 1902).
- Mistral, F. *Lou tresor dóu Felibrige ou Dictionnaire provençal-français*, 2 vols. (Aix-en-Provence: Remondet-Aubin, 1879–86).
- Molesworth, J. T., with G. Candy and T. Candy. *A Dictionary, Maráṭhi and English*, 2nd edn (Mumbai: Government [of Bombay], 1857).
- Molesworth, J. T., and T. Candy. *A Dictionary, English and Maráṭhi, Compiled for the Government of Bombay* (Mumbai: American Mission, 1847).
- Molesworth, J. T., with T. Candy and G. Candy. *A Dictionary, Murat,hee & English: Compiled for the Government of Bombay* (Mumbai: n.p., 1831).
- Molesworth's English and Marathi Dictionary* [facs. edn of Molesworth et al., *A Dictionary, Maráṭhi and English* (1857)], with a preface by N. G. Kalelkar (Pune: Shubhada-Saraswat, 1992; 1st facs. edn 1975).
- Molina, A. de. *Aquí comienza un vocabulario en la lengua castellana y mexicana*, ed. M. Galeote (Malaga: Universidad de Málaga, 2001).

- Molina, A. de. *Vocabulario en lengua castellana y mexicana* (Mexico City: de Spinola, 1571).
- Moliner, M. *Diccionario de uso del español*, 2 vols. (Madrid: Gredos, 1966–7).
- Molossi, L. *Nuovo elenco di voci e maniere di dire biasimate e di altre che sembrano di buona ragione e mancano ne' vocabolarj italiani* (Parma: Carmignani, 1839–41).
- Monet, P. *Invantaire des deus langues, françoise et latine* (Lyon: Rigaud & Borde, 1635).
- Montgomery, M. B. *From Ulster to America: The Scotch-Irish Heritage of American English* (Belfast: Ulster Historical Foundation, 2006).
- Montgomery, M. B., and J. S. Hall. *Dictionary of Smoky Mountain English* (Knoxville: University of Tennessee Press, 2004).
- Moon Seyoung. *Joseon-eo Sajeon* (Seoul: Joseon-eosajeonganhaenghoe, 1938).
- Moore, G. F. *A Descriptive Vocabulary of the Language in Common Use amongst the Aborigines of Western Australia; with Copious Meanings, Embodying much Interesting Information Regarding the Habits, Manners, and Customs of the Natives, and the Natural History of the Country* (London: Orr, 1842).
- Morfix, www.morfix.co.il/.
- Morice, A.-G. *The Carrier language (Déné Family): A Grammar and Dictionary Combined*, 2 vols. [with slightly differing titles and imprints] (St-Gabriel-Mödling, Austria: Anthropos; Winnipeg: for the author, 1932).
- Morohashi Tetsuji. *Dai Kan-Wa jiten* 大漢和辭典, revised and enlarged edn, 13 vols. (Tokyo: Taishukan, 1984–6; 1955–60).
- Morrison, R. *Dictionary of the Chinese Language in Three Parts*, 3 vols. (Macao: East India Company, 1815–23).
- Morrison, W. T. *An Anglo-Chinese Vocabulary of the Ningpo Dialect* (Shanghai: American Presbyterian Mission, 1876).
- Moser, E., and M. B. Moser. *Vocabulario seri: seri-castellano, castellano-seri = Vocabularios indígenas*, 5 (Mexico City: Instituto Lingüístico de Verano, 1961).
- Moser, M. B., and S. A. Marlett. *Comcáac quih yaza quih hant ihiip hac: Diccionario seri-español-inglés* (Hermosillo: Universidad de Sonora; Mexico City: Plaza y Valdés, 2005; 2010).
- Mu'arriḡ, *Amṭāl* = Abū Fayḍ Mu'arriḡ b. 'Amr al-Sadūsī, *al-Amṭāl*, ed. R. 'Abdaltawwāb (Beirut: Dār al-Nahḍa al-'Arabiyya, 1983).
- Mühlenbach, K., and J. Endzelins. *Lettisch-deutsches Wörterbuch/Latviešu valodas vārdnīca*, 4 vols. (Riga: Lettisches Bildungsministerium, 1923–32).
- Muka, E. *Słownik dolnosorbskeje řečy a jej narěcow*, 3 vols. (St Petersburg: Akademičeskaja dvēnadcataja gosudarstvennaja tipografija, 1911–15 [vol. I]; Prague: Verlag der böhmischen Akademie für Wissenschaften und Kunst, 1925–8 [vols. II–III]).
- Mukhlis, Ānand Rām. *Mir'āt al-Iṣṭilāḥ*, ed. C. Shekhar, H. Ghelichkani, and H. Yousefdahi, 2 vols. (Delhi: National Mission for Manuscripts, 2013).
- Múlji, K. *A Pocket Dictionary, Gujarati and English* (Mumbai: Union, 1862).
- Murādī, *Ġanā* = Badr al-Dīn al-Hasan b. Qāsim al-Murādī, *al-Ġanā l-dānī fī ḥurūf al-ma'ānī*, ed. F. Qabāwa and M. N. Fāḍil, 2nd edn (Beirut: Dār al-Āfāq al-Ġadīda, 1983).
- Muret-Sanders = *Encyklopädisches englisch-deutsches und deutsch-englisches Wörterbuch/Encyclopaedic German-English and English-German Dictionary*, ed. E. Muret, D. Sanders, et al., 4 vols. (Berlin: Langenscheidt, 1891–1901).
- Murko, A. J. *Slovensko-nemshki in nemshko-slovenski rozhni besednik; kako se slovenshina govori na shtajerskim, koroshkim, krajnskim in v' sahodnih straneh na vogerskim* (Graz: Ferstli, 1833).

- Murmellius, J. *Dictionarius variarum rerum, tum pueris tam adultis utilissimus, cum Germanicae atq[ue] Polonicae interpretatione* (Kraków: Vieto[r], 1528).
- Muṣṭafā, I. Mu 'jam al-waṣīf, 2 vols. (Cairo: Maṭba'at Miṣr, 1960–1).
- Mymmer, F. *Dictionarium trium linguarum* (Kraków: Scharfenberg, 1528).
- Naembouck van allen naturelicken, ende ongheschuumden vlaemschen woirden (Gent: Lambrecht, 1546).
- Nagasawa Kikuya. *Min Shin zokugo jisho shūsei* 明清俗語辭書集成 (Tokyo: Kyūko shoin, 1974).
- Nágera Yanguas, D. de. *Doctrina y enseñanza en la lengua maçahua de cosas muy vtils, y prouechosas para los Ministros de Doctrina, y para los naturales que hablan la lengua Maçahua* (Mexico City: Ruyz, 1637).
- Nakhat [Niyāz 'Alī Beg]. *Maḥẓan-i favā'id: Urdū muṣṭalahāt, muḥāvarāt aur amṣāl kā ek nādir luḡhat*, ed. M. Z. Ḥusain (Patna: Khuda Bakhsh Oriental Public Library, 1998; 11845).
- Nannūl* = Pavananti Muṇivar, *Nannūl* with the commentary of Mayilai Nātar, ed. U. Vē. Cāminātaiyar, 2nd edn, prepared by S. Kaliyāṇa Cuntaraiyar (Chennai: Kapīr Accukkūtam, 1946; 1918).
- Nascimento, J. P. do. 'Diccionario Portuguese–Umbundu' and 'Diccionario Umbundu–Portuguez' in Nascimento, *Grammatica do Umbundu ou Língua de Benguela* (Lisbon: Imprensa Nacional, 1894), separately paginated appendix.
- Nathan ben Yeḥiel. *Aruch completum, sive lexicon vocabula et res, quae in libris Targumicis, Talmudicis, et Midrashicis continentur*, ed. A. Kohut, 8 vols. (Vienna: Börg, 1878–92).
- Nebrija, E. A. de. *Dictionarium ex hispaniensi in latinum sermonem* (Salamanca: n.p., 1495).
- Nebrija, E. A. de. *Lexicon: hoc est dictionarium ex sermone latino in hispaniense[m]* (Salamanca: n.p., 1492).
- Nekes, H., and E. A. Worms. *Australian Languages*, 1 microfilm reel = Micro-Bibliotheca Anthropos, 10 (Fribourg: Anthropos Institut, 1953).
- Nekes, H., and E. A. Worms. *Australian Languages*, ed. W. B. McGregor (Berlin and New York: Mouton de Gruyter, 2006).
- Nesselmann, G. H. F. *Wörterbuch der litauischen Sprache* (Königsberg: Bornträger, 1851).
- New Oxford Dictionary of English*, ed. J. Pearsall and P. Hanks (Oxford: Clarendon, 1998).
- New Redhouse Turkish–English Dictionary/Redhouse Yeni Türkçe–İngilizce Sözlük*, ed. A. Tietze et al. (Istanbul: Redhouse Yayınevi, 1968).
- New Redhouse Turkish–English Dictionary/Redhouse Yeni Türkçe–İngilizce Sözlük*, ed. A. Tietze et al., 2nd edn, with supplement (Istanbul: Redhouse Yayınevi, 1974).
- The New Zealand Oxford Dictionary*, ed. G. D. Kennedy and T. Deverson (South Melbourne and Oxford: Oxford University Press, 2005).
- Ngag dbang 'jig rten dbang phyug grags pa'i rdo rje. *Mkhas pa'i ma rgyan = mNgon brjod kyi bstan bcos mkhas pa'i ma rgyan* (Beijing: Mi rigs dpe skrun khang, 1999).
- Nicot, J. *Thresor de la langue francoyse* (Paris: Douceur, 1606).
- Niedermann, M., A. Senn, F. Brender, and A. Salys. *Wörterbuch der litauischen Schriftsprache, Litauisch–Deutsch*, 5 vols. (Heidelberg: Winter, 1932–68).
- Nihon kokugo daijiten*, ed. Matsui Shigekazu et al., 20 vols. (Tokyo: Shogakkan, 1972–6; 2000–2).
- Nilambikai Ammaiyar, T. *Ṽaṭacorramiḷ Akaravaricai* (Chennai: Tenṇintiya Caivacittānta Nūṛpatippuk Kaḷakam, 1937).

- Nirukta* = *The Nighaṇṭu and the Nirukta: The Oldest Indian Treatise on Etymology, Philology, and Semantics*, [part 1] *Sanskrit Text, With an Appendix Showing the Relation of the Nirukta with Other Sanskrit Works*, and [part 2] *Introduction, English Translation, and Notes*, ed. and trans. L. Sarup, 2nd reprint (Delhi: Banarsidass, 1966–7).
- Nişanyan, S., *Sözlerin Soyağacı. Çağdaş Türkçenin Etimolojik Sözlüğü* (Istanbul: Adam Yayınları, 2002).
- Nişanyan, S., *Sözlerin Soyağacı. Çağdaş Türkçenin Etimolojik Sözlüğü*, 4th edn (Istanbul: Everest Yayınları, 2009), www.nisanyansozluk.com/.
- Nizolius, M. *Observationes in M. T. Ciceronem prima [et secunda] pars* (Pralboino: Gambarà, 1535).
- Noceda, J. de, and P. de San Lúcar. *Vocabulario de la lengua tagala* (Manila: Imprenta de la Compañía de Iesus, 1754).
- Nogueira, R. de S. *Dicionário Ronga–Português* (Lisbon: Junta de Investigações do Ultramar, Centro de Estudos Políticos e Sociais, 1960).
- Nonius Marcellus. *De compendiosa doctrina*, ed. R. Mazzacane et al., with an introduction by P. Gatti, 3 vols. projected (Florence: SISMEL, 2014–).
- Norsk Ordbok*, ed. A. Hellevik et al., 12 vols. (Oslo: Norske Samlaget, 1950–2016), no2014.uio.no/perl/ordbok/no2014.cgi.
- Nowell, L. *Vocabularium Saxonium*, ed. A. H. Marckwardt = University of Michigan Publications: Language and Literature, 25 (Ann Arbor: University of Michigan Press, 1952).
- Nyembezi, C. L. S., and O. E. H. M. Nxumalo. *Inqolobane yesizwe* (Pietermaritzburg: Shuter & Shooter, 1966).
- Oates, L. F. *Muruwari Dictionary* (Albury, NSW: Van Brummelen Freelance Desktop Publishing, 1992).
- Oates, W. J., and L. F. Oates. ‘Gugu-Yalanji vocabulary’ in Oates and Oates et al., *Gugu-Yalanji and Wik-Munkan Language Studies* (Canberra: Australian Institute of Aboriginal Studies, 1964), 79–146.
- Ó Cléirigh, M. *Foclóir nó sanasán nua* (Leuven: [press of the Irish Franciscans], 1643).
- ‘O’Davoren’s glossary’, ed. W. Stokes, *Archiv für Celtische Lexicographie* 2 (1904), 197–504.
- OED1* = *A New English Dictionary on Historical Principles*, ed. J. A. H. Murray, H. Bradley, W. A. Craigie, and C. T. Onions, 12 vols. (Oxford: Clarendon, 1884–1928).
- OED1* (1933) = *The Oxford English Dictionary, Being a Corrected Re-issue, with an Introduction, Supplement, and Bibliography, of A New English Dictionary on Historical Principles*, ed. J. A. H. Murray, H. Bradley, W. A. Craigie, and C. T. Onions, 13 vols. (Oxford: Clarendon, 1933).
- OED2* = *The Oxford English Dictionary*, ed. J. A. Simpson and E. S. C. Weiner, 20 vols. (Oxford: Clarendon, 1989).
- OED3* = *The Oxford English Dictionary*, online edn, ed. M. Proffitt, www.oed.com.
- OED Supplement* = *A Supplement to the Oxford English Dictionary*, ed. R. W. Burchfield, 4 vols. (Oxford: Clarendon, 1972–86).
- Oghaniryukti* (Mumbai: Agamoday Samiti, 1919).
- Old Babylonian Model Contracts*, ed. G. Spada, oracc.museum.upenn.edu/obmc/.
- Old English Glosses to the Épinal-Erfurt Glossary*, ed. J. D. Pheifer (Oxford: Clarendon Press, 1974).
- Olmos, A. de. *Arte de la lengua mexicana concluido en el convento de San Andrés de Ueytlalpan en la Provincia de la Totonacapan que es en la Nueva España el 1º de enero de 1547*, ed. A.

- Hernández de León-Portilla and M. León-Portilla, 2 vols. (Madrid: Ediciones de Cultura Hispánica, 1993).
- Ó Muirithe, D. *A Dictionary of Anglo-Irish: Words and Phrases from Gaelic in the English of Ireland* (Dublin: Four Courts, 1996).
- Oprea, I., et al. *Noul dicționar universal al limbii române* (Bucharest and Chișinău: Litera Internațional, 2006).
- ordabók.is, www.ordabok.is/.
- Ordbog over det danske Sprog, ed. V. Dahlerup et al., 28 vols. (Copenhagen: Gyldendal, 1918–56), ordnet.dk/ods_en?set_language=en.
- Ordbog over det danske Sprog: Supplement, ed. A. Duekilde, 5 vols. (Copenhagen: Gyldendal, 1992–2005).
- Ordbok över Svenska språget utgiven av Svenska Akademien/Svenska Akademiens Ordbok, 37 vols. to date (Lund: Gleerup, 1893–), www.saob.se/in-english/.
- Orion of Thebes. *Etymologicon*, ed. F. W. Sturz (Leipzig: Weigel, 1820).
- Osberno. *Derivazioni*, ed. P. Busdraghi (Spoleto: Centro italiano di studi sull'Alto Medioevo, 1996).
- Osing, J. (ed.). *Hieratische Papyri aus Tebtunis I*, 2 vols. = The Carlsberg Papyri, 2; Carsten Niebuhr Institute Publications, 17 (Copenhagen: Museum Tusculanum, 1998).
- Ōta Zensai. *Rigen shūran* = *Zōho Rigen shūran*, ed. Inoue Yorikuni and Kondō Heijō, 3 vols. [here misattributed to Murata Ryōa] (Tokyo: Kōten kōkyūjo insatsubu, 1899–1900; reprint Tokyo: Meichō kankōkai, 1965–6; reprint as *Rigen shūran: Jihitsu kōhonban*, with facs. of manuscript, Tokyo: Kuresu shuppan, 1992–3).
- Ōtsuki Fumihiko. *Daigenkai*, 4 vols. plus index (Tokyo: Fuzanbō, 1932–7).
- Ōtsuki Fumihiko. *Genkai*, 4 vols. (Tokyo: Ōtsuki Fumihiko, 1889–1991; facs. edns Tokyo: Taishūkan, 1979 and Tokyo: Chikuma shobō, 2004).
- Oudin, C. *Tesoro de las dos lenguas francesca y española* (Paris: Orry, 1607).
- OWID: Online-Wortschatz-Informationssystem Deutsch (Mannheim: Institut für deutsche Sprache), www.owid.de/.
- Oxford Dictionaries (Oxford University Press), en.oxforddictionaries.com/.
- Oxford Dictionary of English, 2nd edn, ed. C. Soanes and A. Stevenson (Oxford University Press, 2005).
- The Oxyrhynchus Papyri*, 81 vols. to date (London: Egypt Exploration Fund/Egypt Exploration Society, 1898–).
- Ožegov, S. I. *Slovar' russkogo jazyka* (Moscow: GNS, 1949).
- Pahl, H. W., et al. *The Greater Dictionary of Xhosa*, 3 vols. (Alice, South Africa: University of Fort Hare, 1989–2006).
- Palmer, E. 'Pinella and Avipa Apache vocabulary' (Washington, DC: National Anthropological Archives, Smithsonian Institution, MS 139a).
- Palsgrave, J. *Lesclarcissement de la langue francoyse* (London: Pynson and Haukyns, 1530).
- Pančev, T. *Dopālнение na bālgarskija rēčnik ot N. Gerov* (Plovdiv: Pečetinca 'Trud', 1908).
- Pañjāb Yūniwarasiṭi Aṅgrezī-Pañjābī kosha/English-Punjabi Dictionary* (Patiala: Punjab University, 1968).
- Papias. *Elementarium*, ed. V. de Angelis. 3 fascicles, A–azoni (Milan: Cisalpino-Goliardica, 1977–80).
- Parvum vocabularium Latino-Coreanum ad usum studiosae juventutis Coreanae* (Hong Kong: typis societatis missionum ad exteros, 1891).

- Pasor, G. *Lexicon Graeco-Latinum in Novum Domini nostri Jesu Christi Testamentum* (Herborn: [Corvinus], 1619).
- Patriot woordeboek Afrikaans–Engels/Patriot Dictionary Cape Dutch–English, 2 vols. (Paarl: du Toit, 1902–4).
- Paul, H. *Deutsches Wörterbuch* (Halle (Saale): Niemeyer, 1897).
- Paul, H. *Deutsches Wörterbuch*, 10th edn, ed. H. Henne et al. (Tübingen: Niemeyer, 2002).
- Paulinus a Sancto Bartholomaeo. *Amarasinha, sectio prima: de caelo* (Rome: Fulgonius, 1798).
- Paulinus a Sancto Bartholomaeo. *Centum adagia Malabarica* (Rome: Fulgonius, 1791).
- Pausch, O. (ed.). *Das älteste italienisch–deutsche Sprachbuch* (Vienna: Böhlau Nachfolgern, 1972).
- Peacock, E. *A Glossary of Words in Use in the Wapentakes of Manley and Corringham, Lincolnshire* (London: for the English Dialect Society, 1877).
- Peciar, Št. (ed.). *Slovník slovenského jazyka*, 6 vols. (Bratislava: Slovenská akadémia vied, 1959–68).
- Peile, A. R. *Body and Soul: An Aboriginal View* (Perth: Hesperian, 1997).
- Peile, A. R. 'Kukatja botanical terms and concepts' in McGregor, *Studies in Kimberley Languages*, 71–115.
- Pèiwén yùnfǔ 佩文韻府 (Beijing: [imperial household administration], 1711).
- Pellas, S. A. *Dictionnaire provençal et français* (Avignon: Offray, 1723).
- Pelleprat, P. *Introduction a la langue des Galibis, sauvages de la terre ferme de l'Amerique Meridionale* (Paris: Cramoisy, 1655).
- Perotti, N. *Cornucopiae linguae Latinae* (Venice: de Paganinis, 1489).
- Perrin, J. *An English–Kafir Dictionary of the Zulu-Kafir Language, as Spoken by the Tribes of the Colony of Natal* (Pietermaritzburg: May & Davis, 1855).
- Perry, W. *The Royal Standard English Dictionary* (Edinburgh: Wilkie et al., 1775).
- Petit Larousse illustré: *Nouveau dictionnaire encyclopédique*, ed. C. Augé (Paris: Larousse, 1905).
- Perrin, J. *A Kafir–English Dictionary of the Zulu-Kafir Language, as Spoken by the Tribes of the Colony of Natal* (London: n.p. [but sponsored by the Society for promoting Christian Knowledge], 1855).
- Petitot, E. *Dictionnaire de la langue Dènè-dinjié: dialectes Montagnais ou Chippewayan, Peaux de Lièvre ou Loucheux* (Paris: Leroux; San Francisco: Bancroft, 1876).
- Petitot, E. *Vocabulaire Français–Esquimau, dialecte des Tchiglit des Bouches du Mackenzie et de l'Anderson* (Paris: Leroux; San Francisco: Bancroft, 1876).
- Petrocchi, P. *Novo dizionario universale della lingua italiana*, 2 vols. (Milan: Treves, 1884–1891).
- Petter, R. *English–Cheyenne Dictionary* (Kettle Falls, WA: Petter, 1913–15).
- Pfeiffer, A. *Critica sacra* (Dresden: Hübner, 1680).
- Pfuhl, Ch. T. *Lausitzisch-wendisches Wörterbuch* (Bautzen: Mačica Serbska, 1866).
- Phillips, E. *The New World of English Words, or, A General Dictionary* (London: Brooke, 1658).
- Philostratus. *Lives of the Sophists*, trans. W. C. Wright = Loeb Classical Library, 134 (Cambridge, MA: Harvard University Press, 1921).
- Philoxenos. *Die Fragmente*, ed. C. Theodoridis (Berlin and New York: De Gruyter, 1976).
- Photius. *Amphilochia*, ed. L. G. Westerink (Leipzig: Teubner, 1986).
- Photius. *Lexicon*, ed. C. Theodoridis, 3 vols. to date (Berlin etc. [imprint varies]: De Gruyter, 1982–2013).

- Photius. *Lexicon*, transcribed R. Porson and ed. P. P. Dobree (London: for Trinity College, Cambridge, 1822).
- Photius *On Line*, ed. C. Theodoridis, dctthree.github.io/photios/.
- Phrynichus. *Die Ekloge des Phrynichos*, ed. E. Fischer = Sammlung griechischer und lateinischer Grammatiker, 1 (Berlin and New York: De Gruyter, 1974).
- Phrynichus. *Praeparatio sophistica*, ed. J. von Borries (Leipzig: Teubner, 1911).
- Pianzola, B. *Dizionario grammatiche e dialoghi per apprendere le lingue italiana, latina, greca-volgare, e turca*, 2 vols. (Padua: Conzatti, 1781; 21789 in 3 vols.).
- Pianzola, B. *Grammatica, dizionarioj e colloqui per imparare le lingue italiana, greca-volgare, e turca*, 4 vols. (Venice: Zatta, 1801).
- Pichardo, E. *Diccionario provincial de voces cubanas* (Matanzas: imprenta de la Real Marina, 1836).
- Piekarski, E. *Slovar' jakutskago jazyka* (St Petersburg/Petrograd: Imperatorskaja Akademija nauk [etc.], 1907–30).
- [Pinet, P.-F.], 'French–Miami-Illinois dictionary' (St-Jérôme, QC, Canada: Archives de la Compagnie de Jésus, Province du Canada français, MS 004662).
- Piṅkalantai eṇṇum Piṅkala Nikaṇṭu (Chennai: TTSS, 1968).
- Pittau, M. *Nuovo vocabulario della lingua sarda*, 2 vols. (Sestu: Domus de Janas, 2014).
- Pleteršnik, M. *Slovenski–nemški slovar*, 2 vols. (Ljubljana: Wolf, 1894–5).
- Plocher, J. 'English Apache dictionary, containing a vocabulary of the San Carlos Apache, also some White Mount. terms and many sentences illustrating the use of the words' (Chicago: Newberry Library, MS Ayer 1972).
- Le plus ancien dictionnaire bantu/Het oudste Bantu-woordenboek: Vocabularium P. Georgii Gelensis*, ed. J. van Wing and C. Penders (Louvain: Kuly-Otto, 1928).
- The Pocket Oxford Dictionary of Current English*, ed. F. G. Fowler and H. W. Fowler (Oxford: Clarendon, 1924).
- Podolsky, B. *Ivrit–russkij slovar'* (Tel Aviv: Rolnik, 1993).
- Podolsky, B. *Ivruv: Novejšij ivrit–russkij slovar'* (Tel Aviv: Ivruv, 2007).
- Poirier, C. (ed.). *Dictionnaire historique du français québécois* (Quebec City: Presses universitaires Laval, 1998).
- Pollux, Julius. *Onomasticon*, ed. E. Bethe, 3 vols. (Leipzig: Teubner, 1900–37).
- Pompeius Festus. *De verborum significatu*, ed. W. M. Lindsay (Leipzig: Teubner, 1913).
- Pott, A. *Die Zigeuner in Europa und Asien*, 2 vols. (Halle: Heynemann, 1844–5).
- Praśāsaniya paribāshā, or A Glossary of Administrative and Other Terms* (Shillong: Assam Government Press, 1966).
- Pratt, T. K. *Dictionary of Prince Edward Island English* (University of Toronto Press, 1988).
- Pratt, T. K., and S. Burke. *Prince Edward Island Sayings* (University of Toronto Press, 1998).
- Příruční slovník jazyka českého*, 9 vols. (Prague: [imprints vary], 1935–57).
- Proença, *Vocabulario Tamulico* (1966) = *Antaõ de Proença's Tamil–Portuguese Dictionary*, AD 1679, facs. edn, ed. X. S. Thani Nayagam (Kuala Lumpur: Department of Indian Studies, University of Malaya, 1966).
- Promptorium parvulorum* (London: Pynson, 1499).
- Puruṣottamaśa. *Trikāṇḍaśeṣa*, ed. R. S. Bhattacharya (Varanasi: Ratna, 1995).
- Putsillo, M. P. *Opyt russko–korejskago slovarja* (St Petersburg: Gogenfelden, 1874).
- Pyojun Gug-eo Daesajeon* (Seoul: Dusandong-a, 1999).
- Qafisheh, H. A. *The NTC's Gulf Arabic–English Dictionary* (Lincolnwood, IL: NTC, 1997).

- Qālī, Bārī' = Abū 'Alī Ismā'īl b. al-Qāsim al-Qālī, *al-Bārī' fī l-luġa*, ed. H. al-Ṭa'ān (Baghdad: Maktabat al-Nahḍa; Beirut: Dār al-Ḥaḍāra al-'Arabiyya, 1975).
- Al-Qāmūs/The Dictionary*, CD-ROM version 2 (Cairo: Sakhr Software, 1996).
- Quay, M., et al. *Western Apache-English Dictionary* (Whiteriver, AZ: White Mountain Apache Tribe, Tribal Education Department, 1987).
- Quintana, A. de. *Instruccion christiana, y guia de ignorantes para el Cielo en lengua mixe* (Puebla: Ortega Bonilla, 1729).
- Quintana, A. de. *Confessionario en lengua mixe, con una construccion de las oraciones de la Doctrina Christiana, y un compendio de voces mixes, para enseñarse a pronunciar la dicha lengua* (Puebla: Ortega, 1733).
- Quintão, J. L. *Dicionários Xironga-Português e Português-Xironga* (Lisbon: Agência Geral das Colónias, 1951).
- Qurub, Muṭallatāt = Abū 'Alī Muḥammad b. al-Mustanīr Qurub, *al-Muṭallatāt*, ed. R. al-Suwaysī (Libya and Tunis: al-Dār al-'Arabiyya li-l-Kitāb, 1978).
- Radloff, F. *Opyt slovarja tjurkskix narechij/Versuch eines Wörterbuches der Türk-Dialecte* (St Petersburg: Imperatorskaja Akademija nauk, 1893–1911).
- Rangīn, S. Y. K., and M. I. K. *Inshā. Dīvān-i Rangīn va Inshā* (Budaun: Nizami, 1924).
- Raphelengius, F. *Lexicon arabicum* (Leiden: ex officina auctoris, 1613).
- Rasles [Rale], S. 'A dictionary of the Abnaki language in North America', ed. J. Pickering, *Memoirs of the American Academy of Arts and Sciences*, new ser. 1 (1833), 370–574.
- Rastell, J. *Exposiciones terminorum legum anglorum* (London: Rastell, c. 1524).
- Ray, A *Collection of English Words* (1874) = *Ray's Collection of English Words Not Generally Used, from the Edition of 1691; together with Thoresby's Letter to Ray*, ed. W. W. Skeat (London: for the English Dialect Society, 1874).
- Ray, J. A *Collection of English Words Not Generally Used, with Their Significations* (London: Barrell, 1674).
- Rāyamukuṭa. *Padacandrikā*, ed. K. K. Dutta (Calcutta: Sanskrit College, 1966).
- Raynouard, F.-J.-M. *Lexique roman, ou Dictionnaire de la langue des troubadours*, 6 vols. (Paris: Silvestre, 1838–4).
- Reagan, A. B. 'Grammar treatise of the White Mountain Apache Indian language; also a vocabulary of one thousand words of the same language with their English equivalents appended thereto' (Washington, DC: National Anthropological Archives, Smithsonian Institution, MS 1982).
- Rečnik na makedonskiot jazik. 3 vols. (Skopje: Institut za makedonski jazik, 1961–6).
- Rečnik srpskohrvatskog književnog i narodnog jezika. 19 vols. to date (Belgrade: Institut na srpskohrvatski jezik, 1959–).
- Reda, Y. M. *Al-Kamel Al-Kabir plus: Dictionnaire du francais classique et contemporain français-arabe* (Beirut: Librairie du Liban, 1996).
- Redḍi, T. R. T. *Shabdārtha-Chintāmaṇi* (Chennai: n.p., 1906).
- Redhouse, J. W. A *Turkish and English Lexicon* (Istanbul: Boyajian, 1890).
- The Redhouse Contemporary Turkish-English Dictionary/Redhouse Çağdaş Türkçe-İngilizce Sözlüğü*, ed. A. Tietze et al. (Istanbul: Redhouse Yayınevi, 1983).
- The Redhouse Turkish-English Dictionary/Türkçe-İngilizce Redhouse Sözlüğü*, ed. S. Bezmez and C. H. Brown (Istanbul: SEV Matbaacılık ve Yayıncılık – Redhouse, 1999).
- Reece, L. *Dictionary of the Walbri (Warlpiri, Walpiri) language*, 2 vols. (University of Sydney, 1975–9).

- Die Reichenauer Glossen*, vol. I, *Einleitung, Text, vollständiger Index und Konkordanzen*, ed. H.-W. Klein with A. Labhardt (Munich: Hueber, 1968).
- Reig, D. *Larousse as-sabil, arabe-français, français-arabe* (Paris: Larousse, 1983).
- Remacle, L., et al. *Atlas linguistique de la Wallonie*, 10 vols. (Université de Liège, 1953–).
- Reuchlin, J. *De rudimentis hebraicis* (Pforzheim: Anselm, 1506).
- Reuther, J. G. 'Reuther manuscript', 13 vols. (Adelaide: South Australian Museum, series AA/266/09).
- Reyes, A. de los. *Arte en lengua mixteca* (Mexico City: Balli, 1593).
- Rézeau, P. (ed.). *Dictionnaire des régionalismes de France* (Brussels: De Boeck/Duculot, 2001).
- Rhodes, A. de. *Dictionarium Annamiticum, Lusitanum, et Latinum* (Rome: Congregatio de Propaganda Fide, 1651).
- Rhodes, R. A. *Eastern Ojibwa-Chippewa-Ottawa Dictionary* = Trends in Linguistics: Documentation, 3 (Berlin: Mouton de Gruyter, 1985).
- Ribeiro, D. *Vocabulário da língua canarim*, ed. T. Maruyama (Nagoya: Department of Japanese Studies, Nanzan University, 2005).
- Richards, E., and J. Hudson. *Walmajarri-English Dictionary, with English Finder List* (Darwin: Summer Institute of Linguistics, 1990).
- Richardson, C. *A New Dictionary of the English Language*, 2 vols. (London: Pickering, 1836–7).
- Richardson, J. *A Dictionary, Persian, Arabic, and English*, 2 vols. (Oxford: Clarendon, 1777–80).
- Richardson, J. *A Dictionary, Persian, Arabic, and English*, ed. C. Wilkins, 2 vols. (London: Richardson [et al.], 1806–10).
- Richelet, P. *Dictionnaire françois*, 2 vols. (Geneva: Widerhold, 1679–80).
- [Ridel, F.-C., et al.] *Dictionnaire coréen-français* (Yokohama: Levy, 1880).
- Ridley, M. *A Dictionarie of the Vulgar Russe Tongue*, ed. G. C. Stone (Cologne: Böhlau, 1996).
- Ridley, W. *Kámilarói, and Other Australian Languages, with Comparative Tables of Words from Twenty Australian Languages, and Songs, Traditions, Laws, and Customs of the Australian Race*, 2nd edn (Sydney: T. Richards, 1875; 1866 as *Kamilaroi, Dippil, and Turrubul*).
- Riggs, S. R. *A Dakota-English Dictionary*, ed. J. O. Dorsey = Contributions to North American Ethnology, 7 (Washington, DC: Government Printing Office, 1890; reprint 1992).
- Riggs, S. R. (ed.). *Grammar and Dictionary of the Dakota Language* = Smithsonian Contributions to Knowledge, 4 (Washington, DC: Smithsonian Institution, 1852).
- Rinaldini, B. *Arte de la lengua tepeguana con vocabulario, confessionario, y catechismo* (Mexico City: de Hogal, 1743; facs. reprint 1994).
- Rishi, W. R. *Rūṣī-Hindī śābdakośa* (New Delhi: Sāhitya Akādemī, 1957).
- Rivola, F. *Dictionarium armeno-latinum* (Milan: Collegium Ambrosianum, 1621).
- Rječnik hrvatskoga ili srpskoga jezika*, 23 vols. (Zagreb: Jugoslavenska akademija znanosti i umjetnosti, 1880–1976).
- Robert, P. *Dictionnaire alphabétique et analogique de la langue française*, 6 vols. (Paris: Société du Nouveau Littré, 1951–64).
- Robert, P. *Le Petit Robert* (Paris: Société du Nouveau Littré, 1967).
- Robinson, J. *Voices of Queensland: Words from the Sunshine State* (Melbourne: Oxford University Press, 2001).
- Rocha, E., et al. 'Dilzhe'e Apache learners' dictionary', ed. W. J. de Reuse (manuscript, Yavapai-Apache Cultural Resource Center, Camp Verde, Arizona).

- Roques, M. (ed.). *Recueil général des lexiques français du moyen age*. 2 vols. (Paris: Champion, 1936–8).
- Roth, H. *The Sanskrit Grammar and Manuscripts*, ed. A. Camps and J.-C. Muller (Leiden: Brill, 1988).
- Rothwell, W., et al. (eds.). *Anglo-Norman Dictionary*, 7 fascicles (London: Modern Humanities Research Association, 1977–92; 22005).
- Ruby, C. 'Chiricahua Apache vocabulary' (Washington, DC: National Anthropological Archives, Smithsonian Institution, MS 122).
- Ruddiman, T. 'Glossary' in G. Douglas (trans.), *Virgil's Aeneid* (Edinburgh: Simpson and Freebairn, 1710), appendix, sigs. B1 r–Z1 v.
- Ruggieri, M., and M. Ricci (attrib.). *Dicionário Português–Chinês/葡漢辭典/Portuguese–Chinese Dictionary*, facs. edn J. W. Witek (Lisbon: Biblioteca Nacional Portugal, Instituto Português do Oriente; San Francisco: Ricci Institute for Chinese-Western Cultural History, 2001).
- Ruiz Blanco, M. *Conversion de Piritu, de indios cumanagotos, palenques, y otros* (Madrid: Infanzón, 1690).
- Ruiz de Montoya, A. *Arte de la lengua guarani*, ed. P. Restivo (Santa Maria la Mayor, Argentina: [press of the Jesuit mission], 1724).
- Ruiz de Montoya, A. *Arte de la lengua guarani*, ed. P. Restivo, digital edn by S. M. Liuzzi. celia.cnrs.fr/FichExt/Paleographies/Guarani/Guarani_intro.htm.
- Ruiz de Montoya, A. *Arte y bocabulario de la lengua guarani* (Madrid: Sánchez, 1640).
- [Ruiz de Montoya, A.]. *Linguae guarani grammatica*, ed. P. Restivo, newly ed. and trans. C. F. Seybold (Stuttgart: Kohlhammer, 1892).
- Ruiz de Montoya, A. *Tesoro de la lengua guarani* (Madrid: Sánchez, 1639).
- Ruiz de Montoya, A. *Tesoro de la lengua guarani*, ed. B. Melià (Asunción, Paraguay: Centro de Estudios Paraguayos 'Antonio Guasch', 2011).
- Ruiz de Montoya, A. *Vocabulario de la lengua guarani* (Santa Maria la Mayor, Argentina: [press of the Jesuit mission], 1722).
- Saadia Gaon. *Alfāz al-Mishna*, ed. N. Allony, in Benayahu Meir (ed.), *Nissim – Rišonim ve'Aharonim – Sefer zikkaron lehaRav Isaac Nissim* (Jerusalem: Yad Harav Nissim, 1985), 13–26.
- Saadia Gaon. *Ha'egron*, ed. N. Allony (Jerusalem: Academy of Hebrew Language, 1969).
- Saadia Gaon. *Kitāb al-sab'īn lafẓa'*, ed. N. Allony, in S. Loewinger et al. (eds.), *Ignace Goldziher Memorial Volume, Part II* (Jerusalem: Maas, 1958), 1–48.
- Saddhammapajjotikā: The Commentary on The Mahā-niddesa*, ed. A. P. Buddhadatta, 3 vols. (London [place of publication varies]: Pali Text Society, 1980–9; 1931–40).
- Ṣaġānī, 'Ubāb (1) = Raḏī al-Dīn Abū l-Faḏā'il al-Ḥasan b. Muḥammad al-Ṣaġānī, *al-'Ubāb al-zāhir wa-l-lubāb al-fāhir*, ed. M. Ḥ. Āl Yāsīn, 5 vols. (Baghdad: Dār al-Rašīd, 1977–87).
- Ṣaġānī, 'Ubāb (2) = Raḏī al-Dīn Abū l-Faḏā'il al-Ḥasan b. Muḥammad al-Ṣaġānī, *al-'Ubāb al-zāhir wa-l-lubāb al-fāhir*, ed. V. M. Ḥasan (Baghdad: al-Maġma' al-'Ilmī al-'Irāqī, 1978).
- Sagard, G. *Le grand voyage du pays des Hurons . . . avec un dictionnaire de la langue huronne* (Paris: Moreau, 1632).
- Sagard, G. *Le grand voyage du pays des Hurons, suivi du Dictionnaire de la langue huronne*, ed. J. Warwick (Presses de l'Université de Montréal, 1998).
- Sagard Theodat, G. *Histoire du Canada, et voyages que les frères mineurs recollects y ont faits pour la conversion des infidèles depuis l'an 1615, avec un dictionnaire de la langue huronne*, ed. E. Tross, 4 vols. (Paris: Tross, 1866).

- Şahbā'ī, I. B. *Risālah-i qavā'id-i şarf o naḥv-i urdū* (Lucknow: Munshi Naval Kishor, 1880).
- Şāhib, Muḥiṭ = al-Şāhib Abū l-Qāsim Ismā'il ibn 'Abbād, *al-Muḥiṭ fī l-luġa*, ed. M. H. Āl Yāsīn, 11 vols. (Beirut: 'Ālam al-Kutub, 1994).
- Saikia, S., C. S. Deuri, L. K. Deuri, and L. P. Deuri. *Deurī-Asamīyā citra śabdāvalī* (Mysore: Central Institute of Indian Languages, 2014).
- Şăineanu, L. *Dicţionar universal al limbii române* (Craiova: Institutul de Editură Ralian şi Ignat Samitca, 1896).
- Sale, F. da. *Fundamenti principali della lingua retica o griggiona, all' uso di due delle principali valli della Rezia, cioè di Sopraselva e di Sorset, coll' aggiunta d'un vocabolario italiano, e reto di due lingue romancie* (Disentis: Binn, 1729).
- Salesbury, W. A *Dictionary in Englyshe and Welshe* (London: Waley, 1547).
- Salle de l'Etang, S. P. de la. *Dictionnaire galibi, présenté sous deux formes: I. commençant par le mot françois; II. par le mot galibi, précédé d'un essai de grammaire* (Paris: Bauche, 1763).
- Salvá, V. *Nuevo diccionario de la lengua castellana* (Paris: Salvá, 1846).
- Salvado, R. *Memorie storiche dell'Australia, particolarmente della missione Benedettina di Nuova Norcia e degli usi e costumi degli Australiani* (Rome: Congregatio de Propaganda Fide, 1851).
- Salvado, R. *The Salvado Memoirs: Historical Memoirs of Australia and Particularly of the Benedictine Mission of New Norcia and of the Habits and Customs of the Australian Natives*, trans. E. J. Stormon (Nedlands: University of Western Australia Press, 1977).
- Samekita Prasāsana Śabdāvalī, *Āngrezī-Hindī/Consolidated Glossary of Administrative Terms, English-Hindi* (New Delhi: Standing Commission for Scientific and Technical Terminology, Government of India, 1968).
- Sami, Ş. *Kâmûs-ı Türki*, 2 vols. (Istanbul: Ikdam Matbaası, 1899–1900).
- Samyuttanikāya = Samyutta-Nikāya, ed. L. Feer, 6 vols. (Oxford [place of publication varies]: Pali Text Society, 1975–2006; ¹1884–1904).
- San Antonio, F. *Vocabulario tagalo: Tagalog-Spanish Dictionary*, ed. A. Postma (Quezon: Ateneo de Manila University, 2000).
- San Buenaventura, P. de. *Vocabulario de la lengua tagala: el romance castellano puesto primero primera, y segunda parte* (Pila: Pinpin and Loag Tagalos, 1613).
- San Buenaventura, P. de. *Vocabulario de lengua tagala*, facs. edn (Valencia: Librerías Paris-Valencia, 1994).
- Sandefur, J., and J. Sandefur. *Beginnings of a Ngukurr-Bamyili Creole Dictionary* (Darwin: Summer Institute of Linguistics, 1979).
- Sanders, D. *Wörterbuch der deutschen Sprache, mit Belegen von Luther bis auf die Gegenwart*, 2 vols. in 3 (Leipzig: Wigand, 1860–5; ²1876).
- Sanders, D. *Wörterbuch der deutschen Sprache, mit Belegen von Luther bis auf die Gegenwart*, facs. edn (of 2nd edn), ed. W. Betz, 3 vols. (Hildesheim and New York: Olms, 1969).
- Sanders, W. H., W. E. Fay, et al. *Vocabulary of the Umbundu Language, Comprising Umbundu-English and English-Umbundu: Lists of Three Thousand Words Used by the Inhabitants of Bailundu and Bihe, and Other Countries of West Central Africa* (Boston: Beacon, 1885).
- Sanseidō kokugo jiten, ed. Kindaichi Kyōsuke (Tokyo: Sanseidō, 1960; ⁷2014).
- Santos, D. de los. *Vocabulario de la lengua tagala, primera, y segunda parte: en la primera, se pone primero el castellano, y despues el tagalo, y en la segunda al contrario, que son las rayzes simples con sus acentos*, new edn (Sampaloc: Convento de Nuestra Señora de Loreto, 1794).

- Santo Tomás, D. de. *Lexicon, o vocabulario de la lengua general del Peru* (Valladolid: Fernández de Córdova, 1560).
- Sanz de Galdeano, S. *A Grammar and Dictionary of the Pela' Language Used by the Kuini Tribe of Kalúmburu*, booklet and CD-ROM (New Norcia, Western Australia: by the author, 2005).
- Saraqusṭī, Af'āl = Abū 'Uṭmān Sa'id b. Muḥammad al-Saraqusṭī, *al-Af'āl*, ed. Ḥ. M. M. Šaraf, 4 vols. (Cairo: al-Hay' a al-'Āmma li-Šu'ūn al-Maṭābi' al-Amīriyya, 1975–9).
- Sarvānanda. *Ṭīkāsarvasva = The Nāmaṅgānusāsana of Amarasimha, with the commentary Ṭīkāsarvasva of Vandyaghaṭīya-Sarvānanda*, ed. T. Gaṇapatiśāstrī, 4 vols. (Trivandrum: Travancore Government Press, 1914–17).
- Śāśvata, *Anekārthasamuccaya = The Anekārthasamuchchaya of Śāśvata: A Lexicon of Sanskrit Words*, ed. N. N. Kulkarni (Pune: Oriental Book Agency, 1929).
- Saxton, D., and L. Saxton. *Dictionary: Papago & Pima to English, O'odham–Mil-gahn; English to Papago & Pima, Mil-gahn–O'odham* (Tucson: University of Arizona Press, 1969).
- Saxton, D., L. Saxton, and S. Enos. *Dictionary: Papago/Pima–English, O'othham–Mil-gahn; English–Papago/Pima, Mil-gahn–O'othham*, 2nd edn, ed. R. L. Cherry (Tucson: University of Arizona Press, 1983; reissued 1998 as *Dictionary: Tohono O'odham/Pima to English, English to Tohono O'odham/Pima*).
- Šaybānī, Ġīm = Abū 'Amr Ishāq b. Mirār al-Šaybānī, *K. al-Ġīm*, ed. I. al-Abyārī et al., 4 vols. (Cairo: Mağma' al-Luġa al-'Arabiyya, 1974–83).
- Scappaticcio, M. C. (ed.). *Artes Grammaticae in frammenti: I testi grammaticali latini e bilingui greco-latini su papiro = Sammlung griechischer und lateinischer Grammatiker*, 17. (Berlin and Boston: De Gruyter, 2015).
- Scapula, J. *Lexicon graecolatinum* (Basel: Episcopus, 1580).
- Schele de Vere, M. *Americanisms: The English of the New World* (New York: Scribner, 1872).
- Scheueren, G. de. *Vocabularius qui intitulatur Teuthonista vulgariter dicendo der Duytschlender* (Cologne: ther Hornen, 1477).
- Schmick, J. J. *Schmick's Mahican Dictionary*, ed. C. Masthay = American Philosophical Society Memoir 197 (Philadelphia: American Philosophical Society, 1991).
- Schneider, L. *Dictionnaire esquimau-français du parler de l'Ungava et contrées limitrophes*, new edn (Quebec City: Presses de l'Université Laval, 1970; 1966).
- Schneider, L. *Dictionnaire français-esquimau du parler de l'Ungava et contrées limitrophes* (Quebec City: Presses de l'Université Laval, 1970).
- Schneider, L. *Ulinaisigutiit: An Inuktitut–English Dictionary of Northern Quebec, Labrador, and Eastern Arctic Dialects (with an English–Inuktitut Index)*, trans. D. R. F. Collis (Quebec City: Presses de l'Université Laval, 1985).
- Schnoll, S. (ed.). 'A Mishnaic Hebrew glossary in Judeo-Greek from the Cairo Geniza' [in Hebrew], *Massorot* 13–14 (2007), 223–52.
- Schoknecht, C. H. M. 'Grammar of the language of the Dieri Aborigines', trans. J. C. Schoknecht, in A. C. Schoknecht and C. P. Schoknecht (eds.), *Missionary Carl Schoknecht, Killalpaninna Mission 1871–1873: Selected Correspondence* (South Oakleigh: Schoknecht, 1997).
- Scholia D in Iliadem*, ed. H. van Thiel (Universitäts- und Stadt Bibliothek Köln, 2014), kups.ub.uni-koeln.de/5586/.
- Scholia Minora in Homerum*, ed. F. Montanari et al., aristarchus.unige.net/Scholia/en/ Database.

- The Scholia Minora in Homerum: An Alphabetical List*, ed. J. Lunden, version 1.0 = Trismegistos Online Publications, 7 (Cologne and Leuven: Trismegistos, 2012), trismegistos.org.
- Schön, J. F. *Dictionary of the Hausa Language: Part I. Hausa–English, Part II. English–Hausa; with Appendices of Hausa Literature* (London: Church Missionary House, 1876).
- Schön, J. F. *Vocabulary of the Hausa Language: Part I. English and Hausa, Part II. Hausa and English; and Phrases, and Specimens of Translations, to which are prefixed, the Grammatical Elements of the Hausa Language* (London: Church Missionary Society, 1843).
- Schottelius, J. G. *Ausführliche Arbeit von der teutschen HauptSprache* (Braunschweig: Zilligern, 1663).
- Schregle, G. *Arabisch–Deutsches Wörterbuch*, 17 fascicles published (Wiesbaden: Steiner, 1981–92).
- Schregle, G. *Deutsch–Arabisches Wörterbuch* (Wiesbaden: Harrassowitz, 1974).
- Schrevelius, C. *Lexicon manuale Graeco–Latinum et Latino–Graecum* (Leiden: Hackius, 1654).
- Schroderus, E. J. *Lexicon Latino–Scondicum* (Stockholm: Käyser, 1637).
- Schroue, T. *Einn Russisch Buch*, ed. A. Fałowski and W. Witkowski (Kraków: Universitas, 1992).
- [Schultze, B., and F. Fritz]. *Orientalisch- und Occidentalischer Sprachmeister* (Leipzig: Geszner, 1748).
- Schumann, C. L. *Neger-Englisches Wörterbuch*, www.suriname-languages.sil.org/Schumann/National/SchumannGerDict.html.
- Schürmann, C. W. *A Vocabulary of the Pamkalla Language . . . to which is Prefixed a Collection of Grammatical Rules* (Adelaide: Dehane, 1844).
- Scott, J. *English–Corean Dictionary, Being a Vocabulary of Corean Colloquial Words in Common Use* (Seoul: Church of England Mission, 1891).
- Scottish National Dictionary*, ed. W. Grant, D. Murison, et al. 10 vols. (Edinburgh: Scottish National Dictionary Association, 1931–76).
- Se‘adiah ibn Danan. *Sefer ha-šorašim*, ed. M. Jiménez-Sánchez (Universidad de Granada, 1996).
- Seaman, P. D. *Hopi Dictionary: Hopi–English, English–Hopi* = Northern Arizona University Anthropological Paper No. 2 (Flagstaff: Department of Anthropology, Northern Arizona University, 1985).
- Seco, M., O. Andrès, and G. Ramos. *Diccionario del español actual*, 2 vols. (Madrid: Aguilar, 1999).
- Seelakkhandha Mahathera, *Sârârtha Candrikâ* = Puruṣottamadeva. *The Trikandāṣeṣha: A Collection of Sanskrit Nouns, With Sârârtha Candrikâ, A Commentary* (Mumbai: Shri Venkateshwara, 1916).
- Segal, M. *Milon ivri–angli shimushi/A Concise Hebrew–English Dictionary* (Tel Aviv: Dvir, 1938).
- Shād, Muḥammad Pādshāh. *Farhang-i Ānandrāj*, ed. M. Dabīr Siyāqī, 7 vols. (Tehran: Kitabkhanah-yi Khayyam, 1956–8).
- Shakespeare, J. *A Dictionary, Hindustani and English: With a Copious Index, Fitting the Work to Serve, also, as a Dictionary of English and Hindustani*, 3rd edn (London: for the author, 1834; 1817).
- Shams al-lughāt/Shums-ool-Loghat, or, a Dictionary of the Persian and Arabic Languages, the Interpretation being in Persian*, ed. J. Baretto, 2 vols. (Calcutta: Hindoostanee Press, 1806).

- Shapiro, F. L. *Ivrit-russkij slovar'* (Moscow: Gosudarstvennoe izdatel'stvo inostrannykh i nacional'nykh slovarnej, 1963).
- Shelomo ben Meborakh, *Kitāb al-Taysīr = Libro de la facilitación: Kitāb At-Taysir* (diccionario judeoárabe de hebreo bíblico), ed. J. Martínez Delgado (Universidad de Granada, 2010).
- Sheridan, T. *A General Dictionary of the English Language* (London: Dodsley et al., 1780).
- Sherwood, W. L. 'Sierra Blanca and Coyotero vocabulary and notes' (Washington: National Anthropological Archives, Smithsonian Institution, MS 137).
- Shímíng huì jiào* 释名汇校, ed. Rèn Jìfāng 任继昉 (Jinan: Qí Lǚ shūshè, 2006).
- Shin Gicheol and Shin Yongcheol. *Pyojun Gug-eosajeon* (Seoul: Eul-yumunhwasa, 1958).
- Shin Gicheol and Shin Yongcheol. *Sae Urimal Keunsajeon* (Seoul: Samseongchulpansa, 1975).
- Shin meikai kokugo jiten*, ed. Kindaichi Kyōsuke, Kenbō Hidetoshi, Kindaichi Haruhiko, Shibata Takeshi, and Yamada Tadao (Tokyo: Sanseidō, 1972; 2011).
- Shísānjīng zhùshū* 十三經注疏, 2 vols. (Beijing, Zhōnghuá shūjú, 1980).
- The Shorter Oxford English Dictionary*, ed. W. Little, H. W. Fowler, and J. S. Coulson, 2 vols. (Oxford: Clarendon, 1933).
- The Shorter Oxford English Dictionary*, ed. A. Stevenson and L. Brown, 6th edn, 2 vols. (Oxford University Press, 2007).
- Shū Xīnchéng* (ed.). *Cǐhǎi* 辭海, 2 vols. (Shanghai: Zhōnghuá Shūjú, 1936).
- Sībawayhi, *Kitāb* = Abū Bīr 'Amr b. 'Uṭmān Sībawayhi, *al-Kitāb*, ed. 'A. M. Hārūn, 5 vols. (Cairo: al-Hay 'a al-Miṣriyya al-'Āmma li-l-Kitāb, 1977).
- Silva, A. de Morais. *Dicionário da língua portuguesa* (Lisbon: Ferreira, 1789).
- Silva, A. de Morais. *Grande dicionário da língua portuguesa*, 12 vols. (Lisbon: Confluência, 1949–59).
- Silva, P., et al. *A Dictionary of South African English on Historical Principles* (Oxford University Press, 1996).
- Silvy, A. *Dictionnaire montagnais-français*, ed. L. Angers, G. E. McNulty, and D. E. Cooter (Montreal: Presses de l'Université du Québec, 1974).
- Sim Uilin. *Botonghaggyo Joseon-eosajeon* (Seoul: Imundang, 1925).
- Sirvydas, K. *Dictionarium trium linguarum* (Vilnius: Typis academicis Societatis Jesu, 1642; 'before 1620?').
- Sirvydas, K. *Dictionarium trium linguarum* (1997) = *Senasis Konstantino Sirvydo žodynas*, ed. K. Pakalka (Vilnius: Mokslo ir enciklopedijų leidybos institutas, 1997).
- Sitārāmasvāmī, Ī. V. *Trībhāṣā-mañjari: A Vocabulary in Telugu, Hindi, and Persian* (Masulipatam: n.p., 1890).
- Sitjar, B. *Vocabulary of the Language of San Antonio Mission, California/Vocabulario de la lengua de los naturales de la mission de San Antonio, Alta California* = *Shea's Library of American Linguistics*, 7 (New York: Cramoisy; London: Trübner, 1861).
- Si tu paṇ chen, *Collected Works* = *Ta'i si tu pa kun mkhyen chos kyi 'byung gnas bstan pa'i nyin byed kyi bka' 'bum*, 14 vols. (Sansal: Palpung Sungrab Nyamso Khang, 1990).
- Sivan, R., and E. Levenston. *Milon Megido hekhadish/The Megiddo Modern Dictionary Hebrew-English* (Tel Aviv: Megiddo, 1965).
- Skinner, S. *Etymologicon linguae anglicanae* (London: H. Brome [et al.], 1671).
- Slovar' Akademii rossijskoj*, 6 vols. (St Petersburg: Imperatorskoi Akademii Nauk, 1789–94; 41867–8; 61897–1929; 1st edn reprint Moscow, 2001–6).

- Slovar slovenskega knjižnega jezika*, 5 vols. (Ljubljana: Slovenska Akademija Znanosti in Umetnosti, Institut za slovenski jezik, 1970–91; ²2014).
- Slovar' sovremennogo russkogo literaturnogo jazyka*, 17 vols. (Moscow and St Petersburg: Akademija Nauk SSSR, 1948–65).
- Slovník ukrajins'koji movy*, ed. I. K. Bilodid, 11 vols. (Kiev: Prosvita, 1970–80).
- Słownik staropolski*, ed. S. Urbańczyk et al., 11 vols. in 76 fascicles (Warsaw and Kraków: Polska Akademia Nauk [imprint varies], 1953–2002).
- Smart, C. 'An imperfect vocabulary of Tonto or Coyotero Apache' (Washington, DC: National Anthropological Archives, Smithsonian Institution, MS 181b).
- Solomon ben Abraham. *Ohel Mo'ed* (Venice: Giustiniani, 1548).
- Solomon ben Abraham. *Ohel Mo'ed: hebräische Sinonima*, ed. J. Willheimer (Vienna: Willheimer, 1881).
- Sommer, W. *Dictionarium Saxonico–Latino–Anglicum* (Oxford: for the author, 1659).
- Sousa e Oliveira, S. de, and Francina, M. A. de C. *Elementos grammaticaes da lingua Nbandu* (Luanda: Imprensa do Governo, 1864).
- South African Concise Oxford Dictionary*, ed. K. Kavanagh et al. (Cape Town: Oxford University Press, 2002).
- Śrīdharaśena. *Abhidhānaviśvalocana*, ed. L. Jamsal with A. Wayman (Tokyo: Naritasan Shinshoji, 1992).
- Stahl, H. *Anführung zu der Esthnischen Sprach* (Tallinn: Reusner, 1637).
- Steere, E. *A Handbook of the Swahili Language as Spoken at Zanzibar* (London: Bell & Daldy, 1870).
- Steere, E. *A Handbook of the Swahili Language as Spoken at Zanzibar*, 3rd edn, ed. A. C. Madan (London: Society for Promoting Christian Knowledge, 1884).
- Steffel, M. 'Tarahumarisches Wörterbuch nebst einigen Nachrichten von den Sitten und Gebräuchen der Tarahumaren' in C. G. von Murr (ed.), *Nachrichten von verschiedenen Ländern des Spanischen Amerika, aus eigenhändigen Aufsätzen einiger Missionare der Gesellschaft Jesu*, 2 vols. (Halle: Hendel, 1809–11), I.293–374.
- Steinbach, C. E. *Vollständiges Deutsches Wörter-Buch*, 2 vols. (Wrocław: Korn, 1734).
- Steinbach, C. E. *Vollständiges Deutsches Wörter-Buch*, facs. edn, ed. W. Schröter, 2 vols. (Hildesheim and New York: Olms, 1973).
- Steingass, F. J. (ed.). *A Comprehensive Persian–English Dictionary, Including the Arabic Words and Phrases to be Met With in Persian Literature, being Johnson and Richardson's Persian, Arabic, and English Dictionary Revised and Enlarged and Entirely Reconstructed* (London: Kegan Paul, 1892).
- Stender, G. F. *Lettisches Lexikon*, 2 vols. (Mitau: Steffenhagen, 1789).
- Stender, G. F. *Neue vollständigere Lettische Grammatik nebst einem hinlänglichen Lexico, wie auch einigen Gedichten* (Braunschweig: im Fürstlichen großen Waisenhause, 1761).
- Stephanus of Byzantium. *Ethnica*, ed. M. Billerbeck et al., 5 vols. (Berlin and New York: De Gruyter, 2006–17).
- Steuerwald, K. *Türkisch–Deutsches Wörterbuch/Türkçe–Almanca Sözlük* (Wiesbaden: Harrassowitz, 1974; ²1988).
- Stewart, G. *American Place-Names: A Concise and Selective Dictionary for the Continental United States of America* (New York: Oxford University Press, 1970).
- Stieler, K. *Der Deutschen Sprache Stammbaum und Fortwachs, oder Teutscher Sprachschatz* (Nürnberg: Hofmann, 1691).

- Stocchi, A. *Nuovo elenco di voci e maniere di dire biasimate* (Parma: Carmingnani, 1839–41).
- Story, G. L., and C. M. Naish. *Tlingit Noun Dictionary*, 2nd edn, ed. H. Davis and J. Leer (Sitka: Sheldon Jackson College, 1976; 11963).
- Story, G. L., and C. M. Naish. *Tlingit Verb Dictionary* (College: University of Alaska, 1973).
- Strehlow, C. 'Aranda–Loritja–English dictionary' (Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies, MS 377; a second copy is MS 394).
- Subhūticandra. *Kavikāmadhenu on Amarakośa 1.1.1–1.4.8, Together with Si tu Paṇ chen's Tibetan Translation*, ed. L. M. Deokar = *Indica et Tibetica*, 55 (Marburg: Indica et Tibetica, 2014).
- Subhūticandra. *Kavikāmadhenu on Darvikara (AK 1.9.5a)*, ed. L. M. Deokar = *Indica et Tibetica*, 56 (Marburg, Indica et Tibetica, 2018).
- Sūda, ed. A. Adler, 5 vols. (Leipzig: Teubner, 1928–38).
- Sūda On Line*, www.stoa.org/sol/.
- Suetonius. *De grammaticis et rhetoribus*, ed. R. A. Kaster (Oxford: Clarendon, 1995).
- Suetonius. *Peri blasphemion; peri paidion/Des termes injurieux; Des jeux grecs (extraits byzantins)*, ed. and trans. J. Taillardat (Paris: Les Belles Lettres, 1967).
- Suyūṭī, Muzhir = ʿĠalāl al-Dīn Abū l-Faḍl 'Abdalraḥmān b. Abī Bakr al-Suyūṭī, *al-Muzhir fī 'ulūm al-luġa wa-anwā 'ihā*, ed. M. A. ʿĠād al-Mawlā, 'A. M. al-Biġāwī, and M. A. Ibrāhīm, 2 vols. (Cairo: Dār Iḥyā' al-Kutub al-'Arabīyya, n.d.).
- Swabo, J. C. *Dictionarium Færoense*, ed. C. Matras, 2 vols. (Copenhagen: Munksgaard, 1966–70).
- Swētlīk, J. H. *Vocabularium Latino–Serbicum* (Bautzen: Richter, 1721).
- Synagoge*, ed. I. C. Cunningham (Berlin and New York: De Gruyter, 2003).
- Taberd, J. L. (ed.). *Dictionarium Anamitico–Latinum*, from the materials of P. Pigneau de Behaine (Serampore: Marshman, 1838).
- Taberd, J. L. (ed.). *Dictionarium Anamitico–Latinum*, from the materials of P. Pigneau de Behaine, revised by J. S. Theurel (Ninh Phú: ex typis missionis Tunquini Occidentalis, 1877).
- Taberd, J. L. (ed.). *Dictionarium Anamitico–Latinum*, from the materials of P. Pigneau de Behaine, facs. edn, with introduction ('Lời nói đầu') by Mai Quốc Liên (Hanoi: NXB Văn Học, 2004).
- Taittirīya Brāhmaṇa* = *Taittirīya Brāhmaṇa, With the Commentary of Bhaṭṭabhasakaramiśra: Aṣṭaka III, Part II*, ed. A. Mahadeva Sastri and L. Srinivasacharya (Mysore: Government of His Highness the Maharaja of Mysore, 1913).
- Tamil Lexicon*, ed. S. Vaiyapuri Pillai, 6 vols. plus supplement (Chennai: University of Madras, 1924–39; reprint in 7 vols. 1982).
- Tamiḷ Nikaṇṭukaḷ*, ed. Ca. Vē. Cuppiramaṇiyaṇ, 2 vols. (Citamparam: Meyyappaṇ Patippakam, 2008).
- Tanḥum Yerushalmi. *Al-Muršid al-Kāfi: The Lexicon of Tanḥum ben Yosef Hayrushalmi to Mishne tora of Maimonides*, ed. and trans. H. Shy (Jerusalem: National Academy of Sciences, 2005).
- Tanikawa Kotosuga. *Hanpon Wakun no shiori*, reprint of original woodblock print edn of *Wakun no shiori*, 7 vols. (Tokyo: Ōzorasha, 1998).
- Tanikawa Kotosuga. *Wakun no shiori*, 93 fascicles (Gifu: Seibidō, 1898–9; reprint Tokyo: Sumiya shobō, 1969; 1777–1886).
- Tanikawa Kotosuga. *Wakun no shiori*, facs. edn of manuscript (Tokyo: Bensei shuppan, 2008).

- Ṭāpish Dihlavī, M. J. *Shams al-bayān fī muṣṭilahāt al-hindūstān*, rev. edn (Patna: Khuda Bakhsh Oriental Public Library, 1979).
- Taplin, G. 'Notes on a comparative table of Australian languages', *Journal of the Anthropological Institute of Great Britain and Ireland*, 1 (1872), 84–8.
- Taplin, G. 'Vocabulary and grammar of the language of the Aborigines who inhabit the shores of the lakes and lower Murray' (manuscript, Barr Smith Library, University of Adelaide).
- Tauste, F. de. *Arte y vocabulario de la lengua de los indios chaymas, cumanagotos, cores, parias, y otros diversos de la Provincia de Cumana, o Nueva Andalucia* (Madrid: de Villa-Diego, 1680).
- Taylor, J., and W. Hunter. *A Dictionary, Hindoostanee and English* (Kolkata: Hindoostanee Press, 1808).
- Teichelmann, C. G., and C. W. Schürmann. *Outlines of a Grammar, Vocabulary, and Phraseology, of the Aboriginal Language of South Australia, Spoken by the Natives in and for some Distance around Adelaide* (Adelaide: Thomas, 1840).
- Teng Shou-hsin (ed.). *Far East 3000 Chinese Character Dictionary* (Taipei: Far East Book, 2006).
- Ten Raa, W. F. E., and S. T. Woenne. 'Research dictionary of the Western Desert Language of Australia', 2 vols. of unpublished printout (Perth: Department of Anthropology, University of Western Australia, 1970–3).
- Tercero cathecismo y exposicion de la Doctrina Christiana, por sermones* (Lima: Ricardo, 1585).
- Terreros y Pando, E. *Diccionario castellano con las voces de ciencias y artes*, 4 vols. (Madrid: Ibarra [imprint varies], 1786–93).
- Tesoro della lingua italiana delle origini*, ed. P. G. Beltrami (Florence: CNR, 1998–), tlio.ovi.cnr.it/TLIO.
- Thesaurus linguae Aegyptiae* (Berlin-Brandenburgische Akademie der Wissenschaften, 2004–), aew.bbaw.de/tla.
- Thesaurus linguae Latinae*, 10 vols., with supplements, and some fascicles of vol. 11 published to date (Leipzig: Teubner, 1900–).
- Thesaurus Theutonicae linguae*, ed. C. Kilian et al. (Antwerp: Plantin, 1573).
- Thibault, A. *Dictionnaire suisse romand* (Geneva and Carouge: Zoé, 1997; ²2004).
- Thibert, A. *English–Eskimo, Eskimo–English Dictionary* (Research Centre for Amerindian Anthropology, University of Ottawa, 1954; ²1958).
- Thieberger, N. (ed.). *Digital Daisy Bates*, bates.org.au.
- Thomas Magistros. *Ecloga vocum Atticarum*, ed. F. W. Ritschl (Halle: Libreria Orphanotrophei, 1832).
- Thornton, R. H. *An American Glossary*, 2 vols. (Philadelphia: Lippincott, 1912).
- Threlkeld, L. E. *An Australian Grammar, Comprehending the Principles and Natural Rules of the Language, as Spoken by the Aborigines in the Vicinity of Hunter's River, Lake Macquarie, &c., New South Wales* (Sydney: Stephens and Stokes, 1834).
- Threlkeld, L. E. *An Australian Language as Spoken by the Awabakal, the People of Awaba or Lake Macquarie near Newcastle: Being an account of their Language, Traditions, and Customs*, ed. J. Fraser (Sydney: Potter, 1892).
- Threlkeld, L. E. *Specimens of a Dialect of the Aborigines of New South Wales: Being the First Attempt to Form their Speech into a Written Language* (Sydney: Monitor Office, 1827).
- Ṭibrīzī, Kanz = Abū Zakariyyā' Yaḥyā b. 'Alī al-Ḥaṭīb al-Ṭibrīzī, *Kanz al-ḥuffāz fī Kitāb taḥdīb al-ʿAlfāz*, ed. L. Cheikho (Beirut: al-Maṭba'a al-Kāṭūlīkiyya, 1895).

- Tietze, A. *Tarihi ve Etimolojik Türkiye Türkçesi Lugatı/Sprachgeschichtliches und etymologisches Wörterbuch des Türkei-Türkischen* (varying places and imprints, 2002–).
- Tietze, A. *Tarihi ve Etimolojik Türkiye Türkçesi Lugatı*, ed. S. Tezcan, 4 vols. [A–L] (Ankara: Türkiye Bilimleri Akademisi, 2016).
- Tiktin, H. *Rumänisch–deutsches Wörterbuch*, 3 vols. (Bucharest: Staatsdruckerei, 1903–25; ²1986–9; ³2001–5).
- Tindale, N. B. 'Compiled vocabularies and grammatical sketches' (Adelaide: South Australian Museum, records series AA/38/o8, catalogued at archives.samuseum.sa.gov.au/aa338/AA338-o8.htm).
- Tivākaram* (1839) = *Cēntaṇ Rīvākaram*, ed. Tāṇṭavarāyamutaliyār (sections 1–8) and Pu. Nayanappa mutaliyār (sections 9–10) ([Chennai]: n.p., 1839).
- Tivākaram* (1840) = *Cēntaṇ Tivākaram*, ed. Tāṇṭavarāyamutaliyār ([Chennai?]: American Mission, 1840).
- Tivākaram*, ed. Mu. Caṇmukam Piḷḷai and I. Cuntaramūrtti, 2 vols. (Chennai: Cēṇṇaip Palkalaik Kaḷakam, 1990–3).
- Tobler, A., and E. Lommatzsch. *Altfranzösisches Wörterbuch*, 11 vols. (Berlin: Weidmann; Wiesbaden: Steiner, 1925–2002).
- Tolkāppiyam Collatikāram*, with commentary by Cēṇāvaiyār, ed. Piramaśrī Ci. Kaṇṇēaiyār (Chunnakam: Tirumakaḷ Aḷuttakam, 1938).
- Tolkāppiyam Collatikāram*, with commentary by Teyvaccilaiyār, ed. Ti. Vē. Kōpālaiyār (Chennai: Tamiḷmaṇ Patippakam, 2003).
- Tolkāppiyam Eluttatikāram*, with commentary by Naccinārkkaiyār, ed. Piramaśrī Ci. Kaṇṇēaiyār (Chunnakam: Tirumakaḷ Aḷuttakam, 1937).
- Tolkāppiyam Muṟṟōṭal* [recitation of entire *Tolkāppiyam*], 5 CDs (Mysore: Centre of Excellence for Classical Tamil, Central Institute of Indian Languages, c. 2007).
- Tolkāppiyam Poruḷatikāram*, with the commentary by ḷampūraṇar, ed. Ti. Vē. Kōpālaiyār (Chennai: Tamiḷmaṇ Patippakam, 2003).
- Tommaseo, N., and B. Bellini. *Dizionario della lingua italiana*, 7 vols. (Turin: Unione tipografico-editrice, 1863–79), www.tommaseobellini.it.
- 'Tönnies Fenne's Low German manual of spoken Russian, Pskov 1607: an electronic text edition', ed. P. Hendriks and J. Schaeken, version 1.1 (July 2008), schaeken.nl/lu/research/online/editions/fenne11.pdf.
- Torres Rubio, D. de. *Arte de la lengua aymara* (Lima: del Canto, 1616).
- Torres Rubio, D. de. *Arte de la lengua quichua* (Lima: Lasso, 1619).
- Torres Rubio, D. de. *Arte de la lengua quichua, nuevamente van añadidos los Romances, el Cathecismo pequeño, todas las Oraciones, los días de fiesta, y ayunos de los Indios, el Vocabulario añadido, y otro Vocabulario de la lengua chinchaisuyo*, ed. J. de Figueredo (Lima: de Contreras, y Alvarado, 1700).
- Trávníček, F., and P. Váša. *Slovník jazyka českého*, 2 vols. (Prague: Borový, 1937).
- Trésor de la langue française: Dictionnaire de la langue du XIXe et du XXe siècle (1789–1960)*, ed. P. Imbs, 16 vols. (Paris: Éditions du CNRS/Gallimard, 1971–94).
- Trésor de la langue française informatisé*, CD-ROM (Paris: CNRS Éditions, 2004), stella.atilf.fr.
- Trévoux = Dictionnaire universel françois et latin*, 3 vols. (Trévoux: Ganeau, 1704; ²1721, ³1732–40, ⁴1743, ⁵1752–4, ⁶1771).
- Trigault, N. *Xīrú Ērmùzī* 西儒耳目資 (Hangzhou: n.p., 1626).

- Trumbull, J. H. *Natick Dictionary* = Smithsonian Institution, Bureau of American Ethnology, Bulletin 25 (Washington, DC: Government Printing Office, 1903; reprint 2009).
- Tshabe, S. L., Z. Guzana, and A. Nokele. *Isichazi-magama SesiXhosa* (Pietermaritzburg: Nutrend, 2008).
- Türkçe Sözlük*, ed. M. A. Ağakay et al. (Ankara: Türk Dil Kurumu, 1945; ¹¹2011).
- Türkiyede halk ağzından söz derleme dergisi*, 6 vols. (Istanbul: Maarif Matbaası/Cumhuriyet Matbaası; Ankara: Türk Tarih Kurumu Basımevi, 1939–57).
- Türkiye’de halk ağzından söz derleme sözlüğü*, 12 vols. (Ankara: Türk Dil Kurumu, 1963–82).
- Ueda Kazutoshi [et al.]. *Ueda’s Daijiten: A Japanese Dictionary of Chinese Characters and Compounds* (Cambridge, MA: Harvard University Press, 1942).
- Uguccione. *Derivationes*, ed. E. Cecchini and G. Arbizzoni, 2 vols. (Tavarnuzze: SISMELE edizioni del Galuzzo, 2004).
- Ullrich, J. *New Lakota Dictionary: Lakhōtiyapi–English, English–Lakhōtiyapi*, 2nd edn (Bloomington, IN: Lakota Language Consortium, 2011).
- Ulmann, K. C. *Lettisches Wörterbuch*, part 1, *Lettisch–deutsches Wörterbuch* (Riga: Brutzer, 1872).
- Ulmann, K. C., and G. S. Brasche. *Lettisches Wörterbuch*, part 2, *Deutsch–lettisches Wörterbuch* (Riga and Leipzig: Brutzer, 1880).
- Underwood, H. G., with H. B. Hulbert and J. S. Gale. *A Concise Dictionary of the Korean Language in Two Parts: Korean–English & English–Korean* (Yokohama: Kelly & Walsh, 1890).
- Unified Dictionary of Geographical Terms* (Rabat: ALECSO, 1994).
- Unified Dictionary of Human Sciences Terms* (Rabat: ALECSO, 1997).
- Uplegger, F. J. ‘San Carlos Apache dictionary’ (San Marino, CA: Huntington Library, MS HM 26144; copy at Arizona State University Library Special Collections).
- Uplegger, F. J. ‘Thesaurus of Apache language forms’ (manuscript, San Carlos, Arizona; carbon copy at Peridot Lutheran Mission, Peridot, Arizona).
- Urbano, A. *Arte breve de la lengua otomí y vocabulario trilingüe español–náhuatl–otomí*, ed. R. Acuña (Mexico City: Universidad Nacional Autónoma de México, 1990).
- Urdū lughāt: tārikhī usūl par* (2017), www.udb.gov.pk/.
- Urimal Keunsajeon*, ed. Hangeul Hakhoe, 4 vols. (Seoul: Eomungag, 1991–2).
- Ušakov, D. N. *Tolkovyj slovar’ russkogo jazyka*, 4 vols. (Moscow: Sovetskaja Ėnciklopedija, 1935–40).
- Valdemārs, K. *Kreewu–latveeschu–vahzu wardnize* (Moscow: n.p., 1872).
- Valdivia, L. de. *Arte y gramatica general de la lengua que corre en todo el Reyno de Chile, con un vocabulario, y confessorio* (Lima: del Canto, 1606).
- Valdivia, L. de. *Arte y gramatica general de la lengua que corre en todo el Reyno de Chile, con un vocabulario, y confessorio*, 2 vols. [facs. reprint, issued with E. Ridruejo, *El Arte de la lengua de Chile de Luis de Valdivia*] (Madrid: Agencia Española de Cooperación Internacional, 2007).
- Valdivia, L. de. *Doctrina Christiana y cathecismo en la lengua allentiac, que corre en la ciudad de S. Juan de la Frontera, con un confessorio, arte, y vocabulario breues* (Lima: del Canto, 1607).
- Valdivia, L. de. *Vocabulario breve en lengua millcayac, en que estan los vocablos necesarios y suficientes para catequizar y confessar* (Lima: del Canto, 1607).

- Valentino, P. *Opera nuova che insegna a parlare la lingua schiavonesca alli grandi, alli piccoli, et alle donne* ([Ancona]: n.p., 1527).
- Valiquette, H. *A Basic Kukatja to English Dictionary* (Wirrimanu (Balgo): Luurnpa Catholic School, 1993).
- Vallot, P. G. *Petit dictionnaire annamite-français, composé sur le plan des dictionnaires de l'Évêque d'Adran (éditions de Mgr. Tabert et de Mgr. Theurel) et à l'aide du dictionnaire franco-tonkinois*, 2nd edn = *Cours complet de langue annamite*, vol. VI (Hanoi: Schneider, 1904).
- Vāmanabhāṭṭa Bāṇa. *Śabdaratnākara*, ed. B. R. Sharma (Dharbhanga: Mithila Institute of Post Graduate Studies and Research in Sanskrit Learning, 1965).
- Van Barlement, N. *Vocabulare van nieus geordineert* (Antwerp: van Liesvelt, 1527).
- Van Mol, M., with K. Berghman. *Leerwoordenboek Arabisch-Nederlands* (Amsterdam: Bulaaq, 2001).
- Van Mol, M., with K. Berghman. *Leerwoordenboek Nederlands-Arabisch* (Amsterdam: Bulaaq, 2001).
- Varo, F. 'Vocabulario da lingoa mandarina' (Rome: Biblioteca Apostolica Vaticana, Ms Borg. cin. 420).
- Varo, F. *Vocabulario de la lengua mandarina* (2006) = *Francisco Varo's Glossary of the Mandarin Language*, ed. W. S. Coblin, 2 vols. (Nettetal: Sankt Augustin, 2006).
- Varalāṛru Murai Tamiḷ Ilakkiyap Pērakarāṭi*, 5 vols. (Chennai: Cānti Cāṭaṇā, 2001–2).
- Vefik Paşa, A. *Lehce-i Osmānî* (Istanbul: Tabhane-i Âmire, 1876).
- Vefik Paşa, A. *Lehce-i Osmānî*, ed. in Latin script by R. Toparlı (Ankara: Türk Dil Kurumu, 2000).
- Velasco, J. de. *Vocabulario de la lengua índica*, ed. P. Peñaherrera de Costales and A. Costales Samaniego (Quito: Instituto Ecuatoriano de Antropología y Geografía, and Biblioteca Ecuatoriana Aurelio Espinosa Pólit, 1964).
- Verborum in institutione grammatica contentorum in Ungaricum et Slavonicum translatio* (Trnava: typis academicis, 1648).
- Viguer, P. F. *Éléments de la langue turque* (Constantinople: Imprimerie du Palais de France, 1790).
- Viśvanātha. *Kośakalpataru*, ed. M. M. Patkar and K. V. Krishnamurthy Sarma, 2 fascicles (Pune: Deccan College Postgraduate and Research Institute, 1957–66).
- Vivar, P. 'Calepino ylocano' (2012) = Fernández Rodríguez, 'Lexicografía de la lengua ilocana', 486–868.
- Vocabulario da lingoa Canarina com versam Portuguese* (Lisbon: Junta de investigações do Ultramar, 1973).
- Vocabulário da lingua de Iapam com a declaração em Português* (Nagasaki: Collegio de Iapam da Companhia de Iesus, 1603–4).
- Vocabulário da lingua de Iapam* (1973) = *Nippo jisho/Vocabulário da lingoa de Iapam com a declaração em Português, facs. edn*, ed. T. Kamei (Tokyo: Bensheisha, 1973).
- Vocabulario degli Accademici della Crusca* (Venice: Alberti, 1612; ⁴1729–38).
- Vocabulario de Iapon declarado primero en portugues por los padres de la compania de Iesus de aquel reyno, y agora en castellano en el Colegio de Santo Thomas de Manila*, facs. edn, ed. H. Izui (Tokyo: Tenri Central Library, 1972).
- Vocabulario de la lengua névome: the Pima Bajo of Central Sonora*, Vol. II, ed. C. W. Pennington (Salt Lake City: University of Utah Press, 1979).

- El vocabulario de la lengua xebera: una doctrina Cristiana en xebero y quechua, y la gramática xebera* (Siglo XVIII), ed. A. Alexander-Bakkerus (Madrid: Iberoamericana; Frankfurt am Main: Vervuert, 2016).
- Vocabulário na língua brasílica: Manuscrito português-tupi do século XVII*, ed. P. Ayrosa (São Paulo: Departamento de Cultura, 1938).
- Vocabulário na língua brasílica*, 2nd edn, ed. C. Drumond, 2 vols. (São Paulo: n.p., 1952–3).
- Vocabulário ortográfico da língua portuguesa* (Lisbon: Imprensa Nacional, 1940).
- Vocabularium sanscreticum, seu linguae brachmanum dictum Namalingam noesjasjanam* (Vatican City: Bibliotheca Apostolica, Ms. Vat. Ind. 8).
- Vocabularius ex quo*, ed. K. Grubmüller et al., 6 vols. (Tübingen: Niemeyer, 1988–2001).
- Vocabularius incipiens Teutonicum* ([Speyer: Drach, c. 1483–4]).
- A Vocabulary of the Navaho Language*, 2 vols. (St Michaels, AZ: Franciscan Fathers, 1912).
- Vodňanský, J. *Vocabularius cuius nomen Lactifer* (Plzeň: Bakalarz, 1511).
- Vossius, G. J. *Etymologicon linguae Latinae* (Amsterdam: Elzevir, 1662).
- Vrančić, F. *Dictionarium quinque linguarum nobilissimarum Europae linguarum Latinae, Italicae, Germanicae, Dalmaticae et Ungaricae* (Venice: Morettus, 1595).
- Wagner, M. L. *Dizionario etimologico sardo*, 3 vols. (Heidelberg: Winter, 1960–4; ²2008).
- Wahrig, G. *Das grosse deutsche Wörterbuch* (Gütersloh: Bertelsmann, 1966).
- Wahrig Deutsches Wörterbuch, mit einem 'Lexikon der deutschen Sprachlehre'*, 8th edn, ed. R. Wahrig-Burfeind (Gütersloh: Bertelsmann, 2010).
- Walker, J. *A Critical Pronouncing Dictionary and Expositor of the English Language* (London: Robinson et al., 1791).
- Wāng Yǐ (ed.). *Guóyǔ cídiǎn 國語辭典* (Shanghai: Shāngwù Yīnshūguǎn, 1937–45).
- Webster, D. H., and W. Zibell. *Iñupiat Eskimo Dictionary* (Fairbanks, AK: Summer Institute of Linguistics for the Alaska Rural School Project, 1970).
- Webster, N. *An American Dictionary of the English Language*, 2 vols. (New York: Converse, 1828).
- Webster, N. *An American Dictionary of the English Language*, ed. N. Porter and C. Goodrich (Springfield, MA: Merriam, 1864).
- Webster's International Dictionary*, ed. N. Porter (Springfield, MA: Merriam, 1890).
- Webster's Third New International Dictionary of the English Language*, ed. P. B. Gove, 2 vols. (London: Bell; Springfield, MA: Merriam, 1961).
- Wehr, H. *Arabisches Wörterbuch für die Schriftsprache der Gegenwart* (Leipzig: Harrassowitz, 1952; ⁵1985).
- Wehr, H. *A Dictionary of Modern Written Arabic*, trans. J. Milton Cowan (Wiesbaden: Harrassowitz; Ithaca, NY: Cornell University Press, 1961; ⁴1994).
- Wehr, H. *Supplement zum arabischen Wörterbuch für die Schriftsprache der Gegenwart* (Wiesbaden: Harrassowitz, 1959).
- Wei Jiao. *Liushu jingyun 六書精蘊 = Siku quanshu cunmu congshu jingbu*, vol. 189 (Jinan: Qi Lu shushe chubanshe, 1997).
- Weinbach, L., and E. Lauden. *Rav-Milon: milon didakti du-leshoni ivri-ivri-angli, angli-ivri* (Tel Aviv: Ad, 1993).
- Wellig, A. *Beiträge zur lettischen Sprachkunde* (Mitau: Steffenhagen und Sohn, 1828).
- Wells Williams, S. *A Tonic Dictionary of the Chinese Language in the Canton Dialect* (Guangzhou: Chinese Repository, 1856).
- Wemmers, J. *Lexicon Aethiopicum* (Rome: Congregatio de Propaganda Fide, 1638).

- Whipple, A. W. 'Vocabularies of North American languages' in A. W. Whipple, T. Ewbanks, and W. W. Turner, *Report upon the Indian Tribes* = Reports of explorations and surveys, to ascertain the most practicable and economical route for a railroad from the Mississippi River to the Pacific Ocean, vol. III, part 3 (Washington, DC: [War Department], 1855), 54–103.
- White, J. B. 'Corrected wordlists in the Apache language, San Carlos Indian Reservation, A. T.' (Washington, DC: National Anthropological Archives, Smithsonian Institution, MS 178-a-4).
- White, J. B. 'Vocabulary of the Apache and Tontoe language, San Carlos Indian Reservation, A. T., October 30, 1873' (Washington, DC: National Anthropological Archives, Smithsonian Institution, MS 178-b-1).
- Whitworth, G. C. *Anglo-Indian Dictionary* (London: Kegan Paul, Trench, 1885).
- Whomersley, D. *Sheffieldish: A Beginner's Phrase-Book* (Sheffield: Publicity Department, Sheffield City Council, 1981).
- Williams, R. *A Key into the Language of America* (London: Dexter, 1643).
- Williamson, J. P. *An English–Dakota Dictionary/Wasícun ȝa dakota ieska wowapi* (New York: American Tract Society, 1902; reprint 1992).
- Wilson, H. H. *A Dictionary, Sanscrit and English: Translated, Amended and Enlarged from an Original Compilation Prepared by Learned Natives for the College of Fort William* (Kolkata: Hindoostanee Press, 1819; ²1831).
- Wiltens, C. *Vocabularium ofte Woort-boeck naer ordre vanden Alphabet in't Duytsch–Maleysch ende Maleysch–Duytsch* (The Hague: van Wouw, 1623).
- Winer, L. *Dictionary of the English/Creole of Trinidad and Tobago on Historical Principles* (Montreal and Kingston: McGill-Queen's University Press, 2009).
- Witsen, N. *Noord en Oost Tartarye*, 2 vols. (Amsterdam: [for the author], 1692).
- Wittenberg, M. *Milon ivri–angli/Comprehensive Modern Hebrew–English Dictionary*, 2 vols. (Jerusalem: Yeda, 1977).
- Wokabularz trydencki*, ed. S. Urbańczyk, *Język Polski*, 52 (1962), 15–29.
- Woordeboek van die Afrikaanse taal*, ed. P. C. Schoonees et al., 14 vols. to date (Pretoria: Die Staatsdrukker [and other imprints], 1950–).
- Woordenboek der Nederlandsche Taal*, ed. M. de Vries et al., 29 vols. in 39, plus supplements (The Hague: Nijhoff, 1864–2001), ivdnt.org/onderzoek-a-onderwijs/lexicologie-a-lexicografie/wnt.
- Wordsworth, C. *Rutland Words* (London: English Dialect Society, 1891).
- Wörterbuch der deutschen Gegenwartssprache*, ed. R. Klappenbach and W. Steinitz, 6 vols. (Berlin: Akademie-Verlag, 1961–77; final edn 1980–5).
- Wright, J. *English Dialect Dictionary*, 6 vols. (London: for the editor, 1898–1905).
- Wright, T. *Dictionary of Obsolete and Provincial English*, 2 vols. (London: Bell and Daldy, 1852).
- Wú Jǐngróng et al. (eds.). *Hàn–Yīng cídiǎn* 汉英词典 (Beijing: Shāngwù Yīnshūguǎn, 1978).
- Wurdboek fan de Fryske Taal/Woordenboek der Friese Taal*, 25 vols. (Leeuwarden: Fryske Akademy, 1984–2011), www.inl.nl/onderzoek-a-onderwijs/lexicologie-a-lexicografie/wft.
- Xiàndài Hànyǔ cídiǎn* 現代漢語詞典 (Beijing: Shāngwù Yīnshūguǎn, 1978; ²2016).
- XIII. *asırdan günümüze kadar kitaplardan toplanmış tanıklarıyla tarama sözlüğü. Türkiye Türkçesinin tarihi sözlüğü hazırlıklarında*, ed. Ö. A. Aksoy and D. Delçin, 4 vols. (Istanbul: Cumhuriyet Basımevi, 1943–57).

- XIII. yüzyıldan beri Türkiye Türkçesiyle yazılmış kitaplardan toplanmış tanıklarıyla tarama sözlüğü, ed. Ö. A. Aksoy and D. Delçin, 8 vols. (Ankara: Türk Dil Kurumu, 1963–77).
- Xīnhuá zìdiǎn* 新華字典 (Beijing: Rénmín Jiàoyù Chūbǎnshè, 1953).
- Xǔ Shèn 許慎. *Shuōwén jiězì* 說文解字, ed. Xú Xuàn 徐鉉 (Changsha: Yuè lù shūshè, 2006).
- Yādavaprakāśa. *Vaijayantī*, ed. G. Oppert (Chennai: Madras Sanskrit and Vernacular Text Publication Society, 1893).
- Yazīdī, Ġarīb = Abū ‘Abdalrahmān ‘Abdallāh b. Yaḥyā b. al-Mubārak al-Yazīdī, Ġarīb al-Qur‘ān wa-tafsīruhu, ed. M. S. al-Ḥāḡḡ (Beirut: ‘Ālam al-Kutub, 1985).
- Yehuda ibn Bal‘am, *Kitāb al-Tajnīs = šəlosha səfarim šel Rav Yehudah ben Bal‘am: maqor və-tirgum lə-‘ivrit yots ‘im la‘or ‘al pi kitve yad və-sifre dəfus bə-tseruf məvo‘ot və-he‘arot*, ed. S. Abramson (Jerusalem: Qiryat-sefer, 1975).
- Yi Heeseung. *Gug-eo Daejaejon* (Seoul: Minjungseogwan, 1961).
- Yongrong, Ji Yun, et al. *Qinding Siku quanshu zongmu* 欽定四庫全書總目 (Taipei: Taiwan shangwu yinshuguan, 1983).
- Yonsei Hangug-eosajeon* (Seoul: Dusandong-a, 1998).
- Young, R. W. ‘Comparative English–Athabaskan dictionary’ (typescript in Alaska Native Language Archives, University of Alaska, Fairbanks), www.uaf.edu/anla/.
- Young, R. W., and W. Morgan, Sr. *The Navajo Language: A Grammar and Colloquial Dictionary*, rev edn (Albuquerque: University of New Mexico Press, 1987; 1980).
- Young, R. W., and W. Morgan, Sr., with S. Midgette. *Analytical Lexicon of Navajo* (Albuquerque: University of New Mexico Press, 1992).
- Yule, H., and A. C. Burnell. *Hobson-Jobson, Being a Glossary of Colloquial Anglo-Indian Words and Phrases* (London: Murray, 1886; 1903, ed. W. Crooke).
- Yule, H., and A. C. Burnell. ‘Hobson-Jobson’: *The Definitive Glossary of British India*, ed. K. Teltscher (Oxford University Press, 2013).
- Zabīdī, Tāḡ, = Abū l-Faḍl Muḥammad Murtaḍā b. Muḥammad al-Zabīdī, *Tāḡ al-‘arūs min ḡawāhir al-Qāmūs*, 10 vols. (Cairo: al-Maṭba‘a al-Ḥayriyya, 1306 AH).
- Zaḡḡāḡī, Lāmāt = Abū l-Qāsim ‘Abdalrahmān b. Ishāq al-Zaḡḡāḡī, *al-Lāmāt*, ed. M. al-Mubārak (Damascus: Maḡma‘ al-Luḡa al-‘Arabiyya bi-Dimašq, 1969).
- Zamaḥṣarī, Asās = Abū l-Qāsim Maḥmūd b. ‘Umar al-Zamaḥṣarī, *Asās al-balāḡa* (Beirut: Dār Šādir and Dār Bayrūt, 1965).
- Zamaḥṣarī, *Muqaddima* = Abū l-Qāsim Maḥmūd b. ‘Umar al-Zamaḥṣarī, *Muqaddimat al-adab*, ed. J. G. Wetzstein (Leipzig: Barth, 1843–50).
- Zamboni, A., et al. *Dizionario etimologico storico friulano*, 2 vols. (Udine: Casamassima, 1984–7).
- Zdanowicz, A., et al. (eds.). *Słownik języka polskiego*, 2 vols. (Vilnius: Orgelbrand, 1861).
- Zeisberger, D. *Zeisberger’s Indian Dictionary: English, German, Iroquois (the Onondaga) and Algonquian (the Delaware)*, ed. E. N. Horsford (Cambridge, MA: Wilson, 1887).
- Zemánek, P., A. Moustafa, and F. Ondráš. *Arabsko-český slovník* (Prague: Set Out, 2006).
- Zenker, J. T. *Dictionnaire turc-arabe-persan/Türkisch-arabisch-persisches Handwörterbuch*, 2 vols. (Leipzig: Engelmann, 1866–76).
- Zheng Xuan, et al. (eds.). *Liji zhushu* 禮記註疏 (Taipei: Taiwan shangwu yinshuguan, 1983).
- Zhongguo minsu fangyan yaoyan congkan chubian* 中国民俗方言谣言丛刊初编 (Yangzhou: Jiangsu Guangling guji keyinshe, 1989).

- Ziegenbalg, B. *Bibliotheca Malabarica: Bartholomäus Ziegenbalg's Tamil Library*, ed. and trans. W. Sweetman with R. Ilakkuvan = Collection Indologie, 119 (Pondicherry: Institut français de Pondichéry, 2012).
- Ziegenbalg, *Genealogie* = Bartholomäus Ziegenbalgs 'Genealogie der malabarischen Götter', ed. D. Jeyaraj (Halle: Franckesche Stiftung zu Halle, 2003).
- Zingarelli, N. *Vocabolario della lingua italiana*, 2nd edn (Milano: Bietti e Reggiani, 1922; ¹1917 A-ferro only).
- Zubaydī, *Ṭabaqāt* = Abū Bakr Muḥammad b. al-Ḥasan al-Zubaydī al-Andalusī, *Ṭabaqāt al-naḥwīyyīn wa-l-luġawīyyīn*, ed. M. A. Ibrāhīm, 2nd edn (Cairo: Dār al-Ma'ārif, 1973).
- Zyzanij, L. *Leksis sirěč' rečenija, v"krat"čě s"branny i iz slovenskago jazyka na prosty russkij dijalekt"* istol"kovany (Vilnius: Drukarnia Bractwa Św. Ducha, 1596).

Secondary Sources

- Abou El Aazm, A. 'al-Luġa al-'arabiyya al-ḥadīta fī mu'jam as-sabīl mu'jam tunā'iy al-luġa', *Études lexicographiques*, 2 (2003), 5–26.
- 'About Norsk Ordbok 2014', no2014.uio.no/eNo/tekst/framside_english.html.
- Adams, J. N. *Bilingualism and the Latin Language* (Cambridge University Press, 2003).
- Adams, M. 'The apprentice: the origins of Mitford Mathews' lexicography', paper presented at the 18th Biennial Meeting of the Dictionary Society of North America, Montreal, 9 June 2011.
- Adams, M. 'Before DARE', paper presented at the 19th Biennial Meeting of the Dictionary Society of North America, University of Georgia, 24 May 2013.
- Adams, M. 'Credit where it's due: authority and recognition at the *Dictionary of American English*', *Dictionaries*, 19 (1998), 1–20.
- Adams, M. 'The lexical ride of a lifetime', *American Speech*, 88 (2013), 168–95.
- Adams, M. 'Phantom dictionaries: the *Middle English Dictionary* before Kurath', *Dictionaries*, 23 (2002), 95–114.
- Adams, M. 'Reading between the lines: language ideology in the *Dictionary of American English*', paper presented at HEL-LEX 5, University of Zürich, 18 February 2017.
- Adams, M. 'The Special Relationship: Anglo-American lexicography on the eve of war', paper presented at the International Conference on Historical Lexicography and Lexicology, St Anne's College, University of Oxford, 18 June 2010.
- Adams, M. 'Words of America: a field guide', *Humanities*, 32.5 (2011), 14–19 and 53.
- Adelaar, W., with P. C. Muysken. *The Languages of the Andes* (Cambridge University Press, 2004).
- Adler, A. 'Suidas [I]' in *Paulys Realencyclopädie A* 4.1, 675–717.
- Aerts, W. J. 'The Monza vocabulary' in W. F. Bakker, A. F. Van Gemert, and W. J. Aerts (eds.), *Studia Byzantina et Neohellenica Neerlandica* (Leiden: Brill, 1972), 36–73.
- Afane-Otsaga, T. 'Lexicography of Fang' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 962–3.
- 'Affī, R. 'Anjomanārā, Farhang-e' in *Encyclopædia Iranica*.
- Ahlqvist, A. (ed.). *The Early Irish Linguist: An Edition of the Canonical Part of the Auraicept na nÉces with Introduction, Commentary, and Indices* (Helsinki: Societas Scientiarum Fennica, 1982).
- Ahn Jaewon and Kim Hyun-Kwon. 'Parvum vocabularium Latino-Coreanum gwa Nahansajeon-ui hageog gachiwa jung-yoseong-e daehayeo', *Journal of Korealex*, 7 (2006), 7–33.

- Aitken, A. J. 'The period dictionaries', in Burchfield (ed.), *Studies in Lexicography*, 94–116.
- Akhtar, M. S. 'Dhārṡāl, Qāzī Khan Badr Moḥammad Dehlavī' in *Encyclopædia Iranica*.
- Alam, M. 'The pursuit of Persian: language in Mughal politics', *Modern Asian Studies*, 32.2 (1998), 317–49.
- ALECSO, www.arabization.org.ma/.
- Algaze, G. 'Initial social complexity in southwestern Asia: the Mesopotamian advantage', *Current Anthropology*, 42.2 (2001), 199–233.
- Allon, M. *Style and Function: A Study of the Dominant Stylistic Features of the Prose Sections of Pāli Canonical Sutta Texts and Their Mnemonic Function* (Tokyo: International Institute for Buddhist Studies, 1997).
- Allsopp, J. 'Dictionaries of Caribbean English', in Cowie (ed.), *Oxford History of English Lexicography*, I.353–77.
- Allsopp, R. 'A critical commentary on the Dictionary of Jamaican English', *Caribbean Studies*, 10.2 (1970), 90–117.
- Alpers, K. *Das attizistische Lexikon des Oros* (Berlin and New York: De Gruyter, 1981).
- Alpers, K. *Bericht über Stand und Methode der Ausgabe des Etymologicum genuinum* (Copenhagen: Munksgaard, 1969).
- Alpers, K. 'Eine byzantinische Enzyklopädie des 9. Jahrhunderts: Zu Hintergrund, Entstehung und Geschichte des griechischen Etymologikons in Konstantinopel und im italogriechischen Bereich' in G. Cavallo, G. De Gregorio, and M. Maniaci (eds.), *Scrittura, libri e testi nelle aree provinciali di Bisanzio: Atti del seminario di Erice (18–25 settembre 1988)* (Spoleto: Centro Italiano di Studi sull'Alto Medioevo, 1991), 235–69.
- Alpers, K. 'Difficult problems in the transmission and interrelation of the Greek etymologica' in G. A. Xenis (ed.), *Literature, Scholarship, Philosophy, and History: Classical Studies in Memory of Ioannis Taifacos* (Stuttgart: Steiner, 2015), 293–314.
- Alpers, K. 'Griechische Lexikographie in Antike und Mittelalter, dargestellt an ausgewählten Beispielen' in H.-A. Koch (ed.), *Welt der Information: Wissen und Wissensvermittlung in Geschichte und Gegenwart* (Stuttgart: Metzler, 1990), 14–38.
- Alpers, K. 'Ein Handschriftenfund zum Cyrill-Glossar in der Staats- und Universitätsbibliothek Bremen', in W. Hörandner and E. Trapp (eds.), *Lexicographica Byzantina* (Vienna: Österreichische Akademie der Wissenschaften, 1991), 11–52.
- Alpers, K. 'Marginalien zur Überlieferung der griechischen Etymologika' in D. Harlfinger and G. Prato (eds.), *Paleografia e codicologia greca*, 2 vols. (Alessandria, Italy: Edizioni dell'Orso, 1991), I.523–41.
- Alpers, K. 'Review of Daly, *Contributions to a History of Alphabetization*', *Gnomon*, 47 (1975), 113–17.
- Alpers, K. '“Zonarae” Lexicon' in *Paulys Realencyclopädie A.10*, 732–63.
- Alsford, N., and B. Fuehrer. 'Carstairs Douglas (1830–1877) and his Chinese–English Dictionary of the Vernacular or Spoken Language of Amoy (1873)', *Journal of Translation Studies*, 1.1 (2017), 137–82.
- Alster, B. *Proverbs of Ancient Sumer: The World's Earliest Proverb Collections* (Bethesda: CDL, 1997).
- Alster, B. 'Some new Sumerian proverbs' in G. Barjamovic et al. (eds.), *Akkade Is King: A Collection of Papers by Friends and Colleagues Presented to Aage Westenholz on the Occasion of His 70th Birthday, 15th of May 2009* = Publications de l'Institut Historique et

- Archéologique Néerlandais de Stamboul 118 (Leiden: Nederlands Instituut voor het Nabije Oosten, 2011), 9–27.
- Alston, R. C. *A Bibliography of the English Language 1500–1800*, 21 vols. in 39 parts (Leeds etc.: for the author, [1965]–2011).
- Alvar Ezquerro, M. ‘Los diccionarios del español en su historia’, *International Journal of Lexicography*, 8.3 (1995), 173–201.
- Āl Yāsīn, M. H. *al-Dirāsāt al-luġawiyya ‘inda l-‘Arab ilā nihāyat al-qarn al-tālīt* (Beirut: Dār Maktabat al-Ḥayāt, 1980).
- Amīn, A. *Ḍuḥā l-Islām*, 2nd edn, 3 vols. (Cairo: Laġnat al-Ta’līf wa-l-Tarġama wa-l-Našr, 1938).
- Anderson, B. R. O’G. *Imagined Communities: Reflections on the Origin and Spread of Nationalism* (London: Verso, 1983).
- Andrews, J. R. *Introduction to Classical Nahuatl*, rev. edn (Norman: University of Oklahoma Press, 2003).
- Aoki Takashi 青木孝. ‘Jisho, sakuin sakusei no rekishi’ 辞書・索引作成の歴史. in Saeki Umetomo 佐伯梅友, Nakada Norio 中田祝夫, and Hayashi Ōki 林大 (eds.), *Kokugogaku* 国語学 = Kokugo kokubungaku kenkyūshi taisei 15 (Tokyo: Sanseidō, 1961), 226–79.
- Arabic Gigaword Corpus*, 2nd edn, ed. D. Graff (Philadelphia: Linguistic Data Consortium, 2006).
- ArabiCorpus*, ed. D. Parkinson, arabiccorpus.byu.edu.
- Aranha, P. ‘*Vulgaris seu universalis*: early modern missionary representations of an Indian cosmopolitan space’ in J. Flores, C. Lefèvre, and I. G. Županov (eds.), *Cosmopolitismes en Asie du Sud: sources, itinéraires, langues (XVIe–XVIIIe siècle)* ([Paris]: Éditions de l’École des hautes études en sciences sociales, 2015), 331–60.
- Århammer, N. ‘Friesische Lexikographie’ in Hausmann et al., *Wörterbücher*, II.2022–36.
- Arlotto, A. T. ‘Jones and Coeurdoux: correction to a footnote’, *Journal of the American Oriental Society*, 89.2 (1969), 416–17.
- Arts, T. ‘The making of a large English–Arabic/Arabic–English dictionary: the Oxford Arabic Dictionary’ in A. Abel, C. Vettori, and N. Ralli (eds.), *Proceedings of the XVI EURALEX International Congress: The User in Focus* (Bolzano: Institute for Specialised Communication and Multilingualism, 2014), 109–23, euralex.org.
- Arts, T., and K. McNeil. ‘Corpus-based lexicography in a language with a long lexicographical tradition: The case of Arabic’ in E. Atwell and A. Hardie (eds.), *Proceedings of WACL’2: Second Workshop on Arabic Corpus Linguistics* (2013), [14–16], ww.w.academia.edu/7758383.
- Aruṇācalam, Mu. *Tamiḷ Ilakkīya Varalāru*, new edn, 14 vols. (Chennai: Ti Pārkkar, 2005).
- Ashgate Critical Essays on Early English Lexicographers*, 5 vols. (Farnham, UK: Ashgate, 2012).
- Ashtor, E. *Qorot haYhudim bi-Sfarad haMuslamit*, 2 vols. (Jerusalem: Kiryat Sefer, 1977).
- Aslanov, C. *Le provençal des Juifs et l’hébreu en Provence: le dictionnaire ‘Šaršot ha-Kesef’ de Joseph Caspi* (Paris: Peeters, 2001).
- Assunção, C. ‘Portuguese missionary work and inter-linguistic contact in the east’, in *Metamorfoses: 25 anos do Departamento de Letras – Artes e Comunicação* (Vila Real: Universidade de Trás-os-Montes e Alto Douro, Centro de Estudos em Letras, 2011), 91–117.
- Atsujī Tetsujī. ‘Kanwa jiten no rekishi’ in Jiten kyōkai, *Nihon no jisho no ayumi*, 69–94.

- Aune, M. G. 'Review of T. Herbert, *Travels in Africa, Persia, and Asia the Great*, ed. J. A. Butler (2012)', *Seventeenth-Century News*, 72.3/4 (2014), 230–3.
- Auroux, S. (ed.). *Histoire des idées linguistiques*, 3 vols. (Liège: Mardaga, 1989–2000).
- Auroux, S., E. F. K. Koerner, H. J. Niederehe, and K. Versteegh (eds.). *History of the Language Sciences/Geschichte der Sprachwissenschaften/Histoire des sciences du langage*, 2 vols. (Berlin and New York: De Gruyter, 2000–1).
- Austin, P. K. 'The Gamilaraay (Kamilaroi) language, northern New South Wales: a brief history of research' in McGregor (ed.), *Encountering Aboriginal Languages*, 37–58.
- Azar, M. 'Milon Even-Shoshan, revisited', *Kernerman Dictionary News*, 12 (2004), 32–3.
- Baalbaki, R. M. *The Arabic Lexicographical Tradition from the 2nd/8th to the 12th/18th Century* (Leiden and Boston: Brill, 2014).
- Baalbaki, R. M. 'Coinage in modern English–Arabic lexicography', *Zeitschrift für Arabische Linguistik*, 43 (2004), 67–71.
- Baalbaki, R. M. 'The influence of poetry on the rhyme system in lexicography: evidence from Bandanīgī's (d. 284/897) *Kitāb al-Taḥfiya*', *Journal of Arabic Linguistics Tradition*, 10 (2012), 20–30.
- Baalbaki, R. M. 'Introduction: the early Islamic grammatical tradition', in R. Baalbaki (ed.), *The Early Islamic Grammatical Tradition = The Formation of the Classical Islamic World*, 36 (Aldershot: Ashgate, 2007), xiii–1.
- Baalbaki, R. M. *The Legacy of the 'Kitāb': Sibawayhi's Analytical Methods Within the Context of the Arabic Grammatical Theory* (Leiden: Brill, 2008).
- Baalbaki, R. M. 'A precursor to the rhyme system in Arabic lexicography: Ibn al-Sikkīt's (d. 244/858) chapter on *fa' l* and *fa' al* in *Iṣlāḥ al-Manṭiq*', *al-Abḥāth*, 60–1 (2012–13), 5–14.
- Baas, P. 'De VOC in Flora's luthoven' in Blussé and Ooms (eds.), *Kennis en compagnie*, 124–37.
- Bacher, W. *Die hebräisch-neuhebräische und hebräisch-aramäische Sprachvergleichung des Abulwalīd Merwān ibn Ganāḥ* (Vienna: Gerold, 1885).
- Baevskii, S. I. *Early Persian Lexicography: Farhangs of the Eleventh to the Fifteenth Centuries*, trans. N. Killian (Honolulu: University of Hawai'i Press, 2007).
- Baevskii, S. I. 'Farhang-e Ānandrāj' in *Encyclopædia Iranica*.
- Baevskij, S. I. *Rannjaja persidskaja leksikografija XI–XV vv.* (Moscow: Nauka, 1989).
- Bagnall, R. S. (ed.). *The Oxford Handbook of Papyrology* (Oxford University Press, 2009).
- Bai, Q. *Fu Shan's World: The Transformation of Chinese Calligraphy in the Seventeenth Century* (Cambridge, MA: Harvard University Asia Center, 2003).
- Bailey, D. C. 'Early Japanese lexicography', *Monumenta Nipponica*, 16.1/2 (1960), 1–52.
- Bailey, D. C. 'The *Rakuyōshū*', *Monumenta Nipponica*, 16.3/4 (1960–1), 289–376.
- Bailey, D. C. 'The *Rakuyōshū* II: the *Irohajishū* and the *Shōgokuhen*', *Monumenta Nipponica*, 17 (1962), 214–64.
- Bailey, T. G. 'Jami' *ul-lugāt*: a new Urdu dictionary' [review of Majid, *Jāmi' al-lughāt*], *Bulletin of the School of Oriental Studies*, 9 (1938), 440–1.
- Bajema, J. 'Jewish grammarians in the lexicography of Wilhelm Gesenius, specifically the comparisons to Arabic' (in Hebrew), unpublished PhD thesis, Hebrew University of Jerusalem (2017).
- Balasubramanian, K. 'Bilingual lexicography of the Indian subcontinent' in Hausmann et al. (eds.), *Wörterbücher*, III.3096–3107.
- Baldanza, K. *Ming China and Vietnam* (Cambridge University Press, 2016).

- Baldauf, I. 'Missionshilfe oder Kuriositätensammlung? Das komanisch-deutsche Glossar im Codex Cumanicus (14. Jh.)' in K. Kreiser, *Germano-Turcica. Zur Geschichte des Türkisch-Lernens in den deutschsprachigen Ländern* (Bamberg: Universitätsbibliothek, 1987), 15–18.
- Balode, I. *Deutsch-lettische Lexikographie: eine Untersuchung zu ihrer Tradition und Regionalität im 18. Jahrhundert* (Tübingen: Niemeyer, 2002).
- Bandle, O., et al. (eds.). *The Nordic Languages: An International Handbook of the History of the North Germanic Languages*, 2 vols. (Berlin: De Gruyter, 2002–2005).
- Barnes, M. P. 'Language cultivation and language planning V: Faroe Islands' in Bandle et al. (eds.), *Nordic Languages*, II.2007–12.
- Barreto, L. F. 'Reseña de "Dicionário Português-Chinês" de John W. Witek (ed.)', *Bulletin of Portuguese/Japanese Studies*, 5 (2002), 117–26.
- Bartelmus, A. *Fragmente einer grossen Sprache: Sumerisch im Kontext der Schreiberausbildung des kassitenzeitlichen Babylonien* = Untersuchungen zur Assyriologie und vorderasiatischen Archäologie, 12 (Berlin and Boston: De Gruyter, 2016).
- Bartholomew, D. A., and L. C. Schoenhals. *Bilingual Dictionaries for Indigenous Languages* (Mexico City: Instituto Lingüístico de Verano, 1983).
- Baxter, W. *A Handbook of Old Chinese Phonology* (Berlin: Mouton de Gruyter, 1992).
- Beal, J. C. 'The contribution of the Rev. Joseph Hunter's Hallamshire Glossary (1829) to Wright's English Dialect Dictionary', in Markus et al. (eds.), *Joseph Wright's 'English Dialect Dictionary' and Beyond*, 39–48.
- Beal, J. C. 'Enregisterment, commodification, and historical context: "Geordie" versus "Sheffieldish"', *American Speech*, 84 (2009), 138–56.
- Bečka, J. 'Afghan dictionaries', *Archív Orientální*, 53 (1985), 69–70.
- Becker, D. *Məqorot 'arviyim lə-diḡduḡ šel Rabbi Yona ibn Janāḥ* (Tel Aviv University Press, 1998).
- Beckerath, J. von. *Chronologie des pharaonischen Ägypten* = Münchner Ägyptologische Studien, 46 (Mainz: Zabern, 1997).
- Bedell, G. 'Kokugaku grammatical theory', unpublished PhD thesis, Massachusetts Institute of Technology (1968).
- Beekman, E. M. *Fugitive Dreams: An Anthology of Dutch Colonial Literature* (Amherst: University of Massachusetts Press, 1988).
- Béjoint, H. *The Lexicography of English: From Origins to Present* (Oxford University Press, 2010).
- Benarroch, M. 'L'apport des dictionnaires de Jerónimo Cardoso (16e siècle) à la datation du Dicionário Houaiss (2001)' in M. Iliescu et al. (eds.), *Actes du 25e Congrès International de Linguistique et de Philologie Romanes: Innsbruck, 3–8 septembre 2007* (Berlin: De Gruyter, 2010), 623–32.
- Bender, E. 'Review of Deshpande and Deshpande, *Universal English-Gujarati Dictionary*', *Journal of the American Oriental Society*, 112.1 (1992), 169–70.
- Bendor-Samuel, J. 'Adaptations of the Roman alphabet: African languages' in Daniels and Bright (eds.), *The World's Writing Systems*, 689–91.
- Beni, K. N. 'Historique des dictionnaires bilingues (dictionnaires persan-français et français-persan)', *La revue de Téhéran*, no. 95 (October 2013), www.teheran.ir/.
- Benson, T. G. 'A century of Bantu lexicography', *African Language Studies*, 5 (1964), 64–91.
- Benzehra, R. 'Issues and challenges for a modern English-Arabic dictionary', *Dictionaries*, 33 (2012), 83–102.

- Benzie, W. Dr. F. J. Furnivall: *A Victorian Scholar Adventurer* (Norman, OK: Pilgrim, 1983).
- Benzing, J. *Bibliographie der Schriften Johannes Reuchlins im 15. und 16. Jahrhundert* (Bad Bocklet etc.: W. Krieg, 1955).
- Bergenholtz, H., and L.-L. Pálfi, 'Videnskabernes Selskabs Ordbog: das erste dänische Nationalwörterbuch' in Haß (ed.), *Große Lexika*, 137–47.
- Bergmann, R. *Verzeichnis der althochdeutschen und altsächsischen Glossenhandschriften* (Berlin and New York: Walter de Gruyter, 1973).
- Bergmann, R., E. Glaser, and C. Moulin-Fankhänel (eds.). *Mittelalterliche volkssprachige Glossen* (Heidelberg: Winter, 2001).
- Bex, T., and R. J. Watts, (eds.). *Standard English: The Widening Debate* (London: Routledge, 1999).
- Bhattacharya, S. 'Experts question roots of "Modern" Assamese: reprint of Nathan Brown's original 1848 grammar creates flutter in the state's literary circles', *Telegraph* (India), 19 July 2011.
- Birwé, R. 'The *Amarakośa* and the lexicographical chapters of the *Agnipurāṇa*', *Journal of the American Oriental Society*, 96.3 (1976), 383–403.
- Birwé, R. 'Fragments from Three Lost *Kośas*: Vyādi's *Utpalinī* II', *Journal of the American Oriental Society*, 87.1 (1967), 39–52.
- Black, R. *Humanism and Education in Medieval and Renaissance Italy: Tradition and Innovation in Latin Schools from the Twelfth to the Fifteenth Century* (Cambridge University Press, 2001).
- Blackburn, S. *Print, Folklore, and Nationalism in Colonial South India* (Delhi: Permanent Black, 2003).
- Blair, A. *Too Much to Know: Managing Scholarly Information Before the Modern Age* (New Haven: Yale University Press, 2010).
- Blau, J. 'Millonim Miqrā'yyim.' *Encyclopaedia Biblica* (Jerusalem: Mossad Bialik, 1962), vol. IV, cols. 901–9.
- Blekastad, M. *Comenius: Versuch eines Umrisses von Leben, Werk und Schicksal des Jan Amos Komenský* (Oslo: Universitetsforlaget, 1969).
- Blochmann, H. 'Contributions to Persian lexicography', *Journal of the Royal Asiatic Society of Bengal*, 37.1 (1868), 1–72.
- Blom, A. H. 'Endlicher's glossary', *Études Celtiques*, 37 (2011), 159–81.
- Blum, R. *Kallimachos: The Alexandrian Library and the Origins of Bibliography* (Madison: University of Wisconsin Press, 1991).
- Blussé, L., and I. Ooms (eds.). *Kennis en compagnie: de Verenigde Oost-Indische Compagnie en de moderne wetenschap* (Amsterdam: Balans, 2002).
- Boisson, C., P. Kirtschuk, and H. Béjoint, 'Aux origines de la lexicographie: les premiers dictionnaires monolingues et bilingues', *International Journal of Lexicography*, 4.4 (1991), 261–315.
- Bol, P. K. 'Wang Anshi and the *Zhouli*' in B. A. Elman and M. Kern (eds.), *Statecraft and Classical Learning: The 'Rituals of Zhou' in East Asian History* (Leiden: Brill, 2010), 229–51.
- Bollée, A. 'Lexicographie créole: problèmes et perspectives', *Revue française de linguistique appliquée*, 10 (2005), 53–63.
- Bombaci, A. 'Una lettera turca in caratteri latini del dragomanno ottomano Ibrāhīm al veneziano Michele Membre (1567)', *Rocznik Orientalistyczny*, 15 (1939–49), 129–44.
- Bonvini, E. 'Les deux premières grammaires françaises du Wolof (Sénégal): une systématisation contrastée', *Histoire Épistémologie Langage*, 23 (2001), 101–16.

- Bonvini, E. 'Repères pour une histoire des connaissances linguistiques des langues africaines 1. Du XVIe siècle au XVIIIe siècle: dans le sillage des explorations', *Histoire Épistémologie Langage*, 18 (1996), 127–48.
- Borjian, H. 'Tajikistan V: dictionaries' in *Encyclopædia Iranica*.
- Bortolami, G. 'I Bakongo: società, tradizioni, e cambiamento in Angola', unpublished PhD thesis, Università degli Studi di Sassari (2012).
- Bostock, J. K. *A Handbook on Old High German Literature*, 2nd edn, revised by K. C. King and D. R. McLintock (Oxford: Clarendon Press, 1976).
- Botley, P. *Learning Greek in Western Europe, 1396–1529: Grammars, Lexica, and Classroom Texts* (Philadelphia: American Philosophical Society, 2010).
- Botta, A. F. *The Aramaic and Egyptian Legal Traditions at Elephantine* (London: Clark, 2009).
- Bottéro, F. 'Le développement des livres de rimes en dictionnaires: Qièyùn et Kānmǐù bǔquē Qièyùn' in Cao Guangshun et al. (eds.), *Breaking Down the Barriers: Interdisciplinary Studies in Chinese Linguistics and Beyond*, 2 vols. (Taipei: Institute of Linguistics, Academia Sinica, 2013), II.1031–50.
- Bottéro, F. 'The Qièyùn manuscripts from Dūnhuáng' in I. Galambos (ed.), *Studies in Chinese Manuscripts: From the Warring States Period to the 20th Century* (Budapest: Institute of East Asian Studies, 2013), 1–34.
- Bottéro, F. *Sémantisme et classification dans l'écriture chinoise: les systèmes de classement des caractères par clés du Shuowen jiezi au Kangxi zidian* (Paris: Institut des Hautes Études Chinoises, Collège de France, 1996).
- Bottéro, F., and C. Harbsmeier. *Chinese Lexicography on Matters of the Heart: An Exploratory Commentary on the Heart Radical in Shuō wén jiě zì 說文解字* (Paris: École des Hautes Études en Sciences Sociales, Centre des Recherches Linguistiques sur l'Asie Orientale, 2016).
- Boulhol, P. *Grec languaige n'est pas doulz au François: Étude et enseignement du grec dans la France ancienne* (Aix-en-Provence: Presses Universitaires de Provence, 2014).
- Branner, D. P. (ed.). *The Chinese Rime Tables: Linguistic Philosophy and Historical-Comparative Phonology* (Amsterdam: Benjamins, 2006).
- Branner, D. P. 'Notes on the beginnings of systematic dialect description and comparison in Chinese', *Historiographia Linguistica*, 24.3 (1997), 235–66.
- Branner, D. P. 'The Suí-Táng tradition of Fānqiè phonology' in Aurox et al. (eds.), *History of the Language Sciences*, I.36–46.
- Branner, D. P., and Y.-Y. Meng, 'Curious lexicographic relic of the Cultural Revolution', *Journal of the Royal Asiatic Society*, 3rd ser., 23.4 (2013), 551–82.
- Branner, D. P. and W. Yi. *Yīntōng: Chinese Phonological Database*, americanorientalsociety.org/yintong/.
- Brásio, A. (ed.). *Monumenta Missionaria Africana*, 15 vols. (Lisbon: [imprint varies], 1952–88).
- Bray, L. 'La lexicographie française des origines à Littré' in Hausmann et al. (eds.), *Wörterbücher*, II.1788–1818.
- Breatnach, L. *A Companion to the Corpus iuris Hibernici* (Dublin: School of Celtic Studies, 2005).
- Breen, Q. 'The *Observationes in M. T. Ciceronem* of Marius Nizolius', *Studies in the Renaissance*, 1 (1954), 49–58.
- Brewer, C. 'OED Online re-launched: distinguishing old scholarship from new', *Dictionaries*, 28 (2013), 101–26.

- Brewer, C. 'Pronouncing the P: prescription or description in English 19th- and 20th-century dictionaries?', *Historiographia Linguistica*, 34 (2007), 257–80.
- Brewer, C. 'Shakespeare, word-coining, and the OED', *Shakespeare Survey*, 65 (2013), 345–57.
- Brewer, C. "'That reliance on the ordinary": Jane Austen and the *Oxford English Dictionary*', *Review of English Studies*, 66 (2015), 744–65.
- Brewer, C. *Treasure-House of the Language: The Living OED* (New Haven and London: Yale University Press, 2007).
- Brisman, S. *History and Guide to Judaic Dictionaries and Concordances* (Hoboken, NJ: KTAV, 2000).
- Brody, R. *The Geonim of Babylonia and the Shaping of Medieval Jewish Culture* (New Haven: Yale University Press, 1998).
- Bronner, D. 'Dry-point glosses in Irish manuscripts' (2015), halshs.archives-ouvertes.fr.
- Bronner, Y. *Extreme Poetry: The South Asian Movement of Simultaneous Narration* (New York: Columbia University Press, 2010).
- Brose, M. C. 'Uyghur technologists of writing and literacy in Mongol China', *T'oung Pao*, 91.4/5 (2005), 396–435.
- Brovarski, E. 'Two Old Kingdom writing boards from Giza', *Annales du Service des Antiquités de l'Égypte*, 71 (1987), 27–52.
- Brown, C. P. *Essay on the Language and Literature of the Telugus* (Chennai: Pharos, 1839).
- Browning, R. 'Teachers', in G. Cavallo (ed.), *The Byzantines* (University of Chicago Press, 1997), 95–116.
- Buchi, E., and P. Renders. 'Gallo-Romance I: historical and etymological lexicography' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 653–62.
- Buckwalter, T., and D. Parkinson. 'Modern lexicography' in J. Owens (ed.), *The Oxford Handbook of Arabic Linguistics* (Oxford University Press, 2013), 539–60.
- Buenaventura de Carrocera, A. 'Los capuchinos españoles en el Congo y el primer Diccionario congolés', *Missionalia Hispanica*, 5 (1945), 209–30.
- Bukhārī, S. Y. 'Maulvī Sayyid Aḥmad Dihlavī' in S. A. Dihlavī, *Rusūm-i dihlī* (Rampur: Kitābkār, 1965), 12–35.
- Burchfield, R. W. (ed.). *Studies in Lexicography* (Oxford: Clarendon Press, 1987).
- Burdett-Jones, M. T. 'Early Welsh dictionaries', in P. H. Jones and E. Rees (eds.), *A Nation and Its Books: A History of the Book in Wales* (Aberystwyth: National Library of Wales, 1998), 75–81.
- Burgess, M. A. 'The Kunmunya years: an evaluation of the contribution of Rev. J. R. B. Love to Australian Aboriginal studies', unpublished B.A. (Hons) thesis, Australian National University (1986).
- Buringh, E. *Medieval Manuscript Production in the Latin West* (Leiden: Brill, 2011).
- Burjiyān, H. 'Farhangnivīsī dar Tājīkistān', *Irānshināsī*, 11.1 (1999), 125–45.
- Burkard, G. 'Bibliotheken im alten Ägypten', *Bibliothek: Forschung und Praxis*, 4 (1980), 79–115.
- Burke, P. *Languages and Communities in Early Modern Europe* (Cambridge University Press, 2004).
- Burnett, S. G. *From Christian Hebraism to Jewish Studies: Johannes Buxtorf (1564–1629) and Hebrew Learning in the Seventeenth Century* (Leiden: Brill, 1996).
- Burns, S. L. 'The politics of philology in Japan: ancient texts, language, and Japanese identity' in S. Pollock, B. A. Elman, and K. K. Chang (eds.), *World Philology* (Cambridge, MA: Harvard University Press, 2015), 245–63.

- Bursill-Hall, G. L. *A Census of Medieval Latin Grammatical Manuscripts* (Stuttgart: Frommann-Holzboog, 1981).
- Busane, M. 'Lexicography in central Africa: the user perspective, with special reference to Zaïre' in R. R. K. Hartmann (ed.), *Lexicography in Africa* (University of Exeter Press, 1990), 19–35.
- Calhoun, D. 'Reading paratexts in missionary linguistic works: an analysis of the preface to the Holy Ghost Fathers' (1855), *Dictionnaire français-wolof et wolof-français*, *Language & History*, 60.1 (2017), 53–72.
- Callanan, C. K. *Die Sprachbeschreibung bei Aristophanes von Byzanz* (Göttingen: Vandenhoeck & Ruprecht, 1987).
- Calvo Pérez, J. *El Arte y Vocabulario en la Lengua General del Perú, llamada Quichua y en la lengua española (1586): estudio* (Madrid: Agencia Española de Cooperación Internacional para el Desarrollo, 2009).
- Calvo Pérez, J. 'El concepto de diccionario bilingüe y su aplicación a las relaciones entre el español y las lenguas amerindias' in J. Calvo Pérez and J. Carlos Godenzzi (eds.), *Multilingüismo y educación bilingüe en América y España* (Cuzco: Centro de Estudios Regionales Andinos 'Bartolomé de las Casas', 1997), 257–82.
- Calvo Pérez, J. 'El vocabulario de Velasco 1787: finalidad de su confección', *Lexis*, 25.1/2 (2001), 33–49.
- Cancino Cabello, N. 'Los tratados millcayac y allentiac (1607) de Luis de Valdivia: noticia de un hallazgo bibliográfico', *Onomázein*, 37 (2017), 112–43.
- Capell, A. 'History of research in Australian and Tasmanian languages' in Sebeok (ed.), *Current Trends in Linguistics*, vol. VIII, *Linguistics in Oceania*, 661–720.
- Capell, A., and H. H. J. Coate. *Comparative Studies in Northern Kimberley Languages* (Canberra: Pacific Linguistics, 1984).
- Carey, H. M. 'Lancelot Threlkeld and missionary linguistics in Australia to 1850' in Zwartjes and Hovdhaugen (eds.), *Missionary Linguistics*, 253–75.
- Carter, M. G. 'Arabic lexicography' in M. J. L. Young, J. D. Latham, and R. B. Serjeant (eds.), *The Cambridge History of Arabic Literature*, vol. III, *Religion, Learning, and Science in the 'Abbasid Period* (Cambridge University Press, 1990), 106–17.
- Cassidy, F. G. 'On collecting American dialect', *American Speech*, 23 (1948), 185–93.
- Cassidy, F. G. 'The meaning of "regional" in DARE', *American Speech*, 48 (1973), 282–9.
- Cassidy, F. G., with A. R. Duckert. *A Method for Collecting Dialect* = Publication of the American Dialect Society, 20 (Gainesville, FL: American Dialect Society, 1953).
- Cavigneaux, A. 'Lexikalische Listen', in *Reallexikon der Assyriologie und vorderasiatischen Archäologie*, ed. E. Ebeling et al. (Berlin: De Gruyter, 1932–), VI.609–41.
- Cellerini, A. *Introduzione all'Etymologicum Gudianum* (Rome: Accademia Nazionale dei Lincei, 1988).
- Ceresa, M. 'Favorino, Guarino' in *Dizionario Biografico degli Italiani*.
- Chabata, E. 'An advanced dictionary? Similarities and differences between Duramazwi ReChiShona and Duramazwi Guru ReChiShona', *Lexikos*, 12 (2002), 105–16.
- Chabata, E. 'Lexicography of Shona', in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 947–52.
- Chalmers, A. *The General Biographical Dictionary*, 32 vols. (London: Nichols et al., 1812–17).
- Chambon, J. P. 'Brèves remarques sur le *Tresor dou Felibrige* de Frédéric Mistral', *Revue de linguistique romane*, 74 (2010), 199–214.

- Chamorro, G. 'Un diccionario manuscrito castellano-guaraní anónimo', *Corpus: Archivos virtuales de la alteridad americana*, 4.2 (2014), 1–23.
- Chatelain, H. *Folk-Tales of Angola: Fifty Tales, with Ki-mbundu Text, Literal English Translation, Introduction, and Notes* (Boston and New York: American Folk-Lore Society, 1894).
- Chen, P. *Modern Chinese: History and Sociolinguistics* (Cambridge University Press, 2008).
- Chen Xinxiong 陳新雄. *Guangyun yanjiu 廣韻研究* (Taipei: Taiwan xuesheng shuju, 2004).
- Chevillard, J.-L. 'The challenge of bi-directional translation as experienced by the first European missionary grammarians and lexicographers of Tamil' in É. Aussant (ed.), *La traduction dans l'histoire des idées linguistiques: Représentations et pratiques* (Paris: Geuthner, 2015), III–30.
- Chevillard, J.-L. 'How Tamil was described once again: towards an XML-encoding of the Grammatici Tamulici', *Histoire Epistémologie Langage*, 39.2 (2017), 103–27.
- Chevillard, J.-L. 'On a 1968 incarnation of the *Piṅkalam*, a traditional Tamil thesaurus' in O. Vecherina, N. Gordiyshuk, and T. Dubyanskaya, *Tamiḷ tanta paricu: The Collection of Articles in Honor of Alexander M. Dubyanskiy* (Moscow: Russian State University for the Humanities, 2016), 15–28.
- Chevillard, J.-L. "'Rare words" in classical Tamil literature: from the *Uriyiyal* to the *Tivākaram*', *Acta Orientalia Academiae Scientiarum Hungaricae*, 63.3 (2010), 301–17.
- Cho Namho. 'Seoyang-in pyeonchan-ui gaehwagi daeyeog sajeon-e daehan jonghabeog geomto', *Gug-eosa yeongu*, 22 (2016), 47–80.
- Choi Kyeongbong. 'Wongolo nam-eun choechoi ulimal sajeon, *Malmoi*', *Saegug-eosaenghwal*, 24.3 (2014), 20–37.
- Choueka, Y. 'Rav-Milim: a modern dictionary for an ancient yet thriving language', *Kernerman Dictionary News*, 12 (2004), 27–31.
- Çınar, B. *Miftâh-ı Lisân. Manzum Fransızca–Türkçe sözlük* (Ankara: Altınpost, 2012).
- Civil, M. 'Ancient Mesopotamian lexicography' in J. M. Sasson (ed.), *Civilizations of the Ancient Near East* (New York: Scribner's, 1995), 2305–14.
- Civil, M. 'Lexicography' in S. J. Lieberman (ed.), *Sumerological Studies in Honor of Thorkild Jacobsen on His Seventieth Birthday, June 7, 1974* = *Assyriological Studies*, 20 (University of Chicago Press, 1976), 123–57.
- Civil, M. 'Remarks on AD-GI₄ (a.k.a. "Archaic Word List C" or "Tribute")', *Journal of Cuneiform Studies*, 65 (2013), 13–67.
- Clackson, J. *Language and Society in the Greek and Roman Worlds* (Cambridge University Press, 2015).
- Claes, F. *Bibliographisches Verzeichnis der deutschen Vokabulare und Wörterbücher, gedruckt bis 1600* (Hildesheim: Olms, 1977).
- Claes, F. *Lijst van Nederlandse woordenlijsten en woordenboeken gedrukt tot 1600* (Nieuwkoop: De Graaf, 1974).
- Clanchy, M. T. *From Memory to Written Record: England 1066–1307*, 2nd edn (Oxford, UK, and Cambridge, MA: Blackwell, 1993).
- Clark, P. *The Kokugo Revolution: Education, Identity, and Language Policy in Imperial Japan* (Berkeley: Institute of East Asian Studies, 2009).
- Clauson, G. *An Etymological Dictionary of Pre-Thirteenth-Century Turkish* (Oxford University Press, 1972).

- Clayton, M. L., and R. J. Campbell. 'Alonso de Molina as lexicographer' in Frawley et al. (eds.), *Making Dictionaries*, 336–90.
- Clements, R. *A Cultural History of Translation in Early Modern Japan* (Cambridge University Press, 2015).
- Coates, A., et al. (eds.). *A Catalogue of Books Printed in the Fifteenth Century Now in the Bodleian Library, Oxford*, 6 vols. (Oxford University Press, 2005).
- Coblin, W. S. *A Handbook of 'Phags-pa Chinese* (Honolulu: University of Hawai'i Press, 2007).
- Coblin, W. S. 'Robert Morrison and the phonology of mid-Qīng Mandarin', *Journal of the Royal Asiatic Society*, 3rd ser., 13.3 (2003), 339–55.
- Coeurdoux, G.-L. 'Réponse au Mémoire de M. l'Abbé Barthélemy', in *Histoire et mémoires de l'Académie des Inscriptions et Belles-Lettres, de 1701 à 1793* (Paris: Imprimerie Impériale, 1808), 647–67.
- Cohen, A. 'Making memories in a medieval miscellany', *Gesta*, 48.2 (2009), 135–52.
- Cohen, C., and Y. Choueika. 'Hebrew lexicography' in *Encyclopedia of Language and Linguistics*, V.268–72.
- Colas, G. 'A cultural encounter in the early 18th century: the collection of South Indian manuscripts by the French Jesuit Fathers of the Carnatic mission' in R. Saraju (ed.), *Aspects of Manuscript Culture in South India* (Leiden: Brill, 2012), 69–80.
- Colas, G., and U. Colas-Chauhan (eds.). *Manuscripts telugu: catalogue raisonné* (Paris: Bibliothèque Nationale de France, 1995).
- Cole, D. T. 'The history of African linguistics to 1945' in T. A. Sebeok (ed.), *Current Trends in Linguistics*, vol. 7: *Linguistics in Sub-Saharan Africa* (The Hague: Mouton, 1971), 1–29.
- Coleridge, D. 'Observations on the plan of the Society's proposed New English Dictionary', *Transactions of the Philological Society*, 7 (1860), 152–68.
- Collins, J. T. 'Language death in Maluku: the impact of the VOC', *Bijdragen tot de taal-, land- en volkenkunde/Journal of the Humanities and Social Sciences of Southeast Asia*, 159.2 (2003), 247–89.
- Considine, J. P. *Academy Dictionaries 1600–1800* (Cambridge University Press, 2014).
- Considine, J. P. *Dictionaries in Early Modern Europe: Lexicography and the Making of Heritage* (Cambridge University Press, 2008).
- Considine, J. P. 'The history of the concept of lexicography' in C. Assunção, G. Fernandes, and R. Kemmler (eds.), *History of Linguistics 2014* (Amsterdam and Philadelphia: Benjamins, 2016), 31–42.
- Considine, J. P. 'John Jamieson, Franz Passow, and the double invention of lexicography on historical principles', *Journal of the History of Ideas*, 75 (2014), 261–81.
- Considine, J. P. 'The lexicography of the learned languages in seventeenth-century England' in *Ashgate Critical Essays*, IV.343–63.
- Considine, J. P. 'Parkade: one Canadianism or two Americanisms?', *American Speech*, 92 (2017), 281–97.
- Considine, J. P. '“Si hoc saeculo natus fuisset”: refurbishing the *Catholicon* for the sixteenth century', *Historiographia Linguistica*, 44.2/3 (2017), 412–29.
- Considine, J. P. *Small Dictionaries and Curiosity: Lexicography and Fieldwork in Post-Medieval Europe* (Oxford University Press, 2017).
- Considine, J. P. 'Wordlists of exotic languages in seventeenth-century England' in *Ashgate Critical Essays*, IV.365–75.

- Cooper, M. 'The *Nippo Jisho*' [review of *Vocabulario da lingua de Iapam* (1973)], *Monumenta Nipponica*, 31.4 (1976), 417–30.
- Copp, P. F. *The Body Incantatory: Spells and the Ritual Imagination in Medieval Chinese Buddhism* (New York: Columbia University Press, 2014).
- Corff, O., et al. (eds.). *Auf kaiserlichen Befehl erstelltes Wörterbuch des Manjurischen in fünf Sprachen 'Fünfsprachenspiegel': Systematisch angeordnet Wortschatz auf Manjurisch, Tibetisch, Mongolisch, Turki, und Chinesisch*, 7 vols. (Wiesbaden: Harrassowitz, 2013).
- Cormier, M., and A. Francoeur. 'Un siècle de lexicographie au Québec: morceaux choisis', *International Journal of Lexicography*, 15.1 (2002), 55–73.
- Cornillac, G. 'De la nécessité de concevoir pour l'esquimaux un dictionnaire sans mots', *Amerindia*, 19/20 (1995), 203–9.
- Corpus de textes linguistiques fondamentaux* (2007–), ctlf.ens-lyon.fr.
- Costa, D. J. 'The St-Jérôme dictionary of Miami-Illinois' in H. C. Wolfart (ed.), *Papers of the 36th Algonquian Conference* (Winnipeg: University of Manitoba Press, 2005), 107–33.
- Couto-Ferreira, M. É. 'Etnoanatomía y patonomía del cuerpo humano en Sumerio y Acadio: el léxico Ugu-Mu', unpublished doctoral thesis, Universitat Pompeu Fabra, 2009.
- Cowie, A. P. (ed.). *The Oxford History of English Lexicography*, 2 vols. (Oxford University Press, 2009).
- Cowley, A. E. (ed. and trans.). *Aramaic Papyri of the Fifth Century BC* (Oxford: Clarendon Press, 1923).
- Craigie, W. A. 'New dictionary schemes', *Transactions of the Philological Society 1925–1930* (1931), 6–9.
- Craigie, W. A. 'Response', *The Periodical*, 15 March 1934, 24–6.
- Crisostomo, J. *Translation as Scholarship: Language, Writing, and Bilingual Education in Ancient Babylonia* = *Studies in Ancient Near Eastern Records*, 22 (Berlin and Boston: De Gruyter, 2018).
- Crowley, T. *Proper English? Readings in Language, History, and Cultural Identity* (London: Routledge, 1991).
- Dabīrsiāqī, M. 'Borhān-e jāmi' in *Encyclopædia Iranica*.
- Dabīrsiāqī, M. 'Borhān-e qāte' in *Encyclopædia Iranica*.
- Dabīr Siyāqī, M. *Farhang-hā-yi Fārsī va Farhang-gūnah-hā*. (Tehran: Ispark, 1989).
- Dahan, G. 'Interpretationes nominum Hebraicorum' in *Encyclopaedia of the Middle Ages*.
- Dalby, A. *A Guide to World Language Dictionaries* (London: Library Association, 1998).
- D'Alençon, É. 'Essai de bibliographie capucino-congolaise', *Neerlandia Franciscana*, 1 (1914), 33–42.
- Dallet, C. *Hangug cheonjugyohoesa*, trans. An Eunglyeol and Choi Seog-u (Seoul: Hanguggyohoesayeonguso, 1987).
- Daly, L. W. *Contributions to a History of Alphabetization in Antiquity and the Middle Ages* (Brussels: Latomus, 1967).
- Daly, L. W., and B. A. Daly. 'Some techniques in mediaeval Latin lexicography', *Speculum*, 39.2 (1964), 229–39.
- Daniels, P. T., and W. Bright (eds.). *The World's Writing Systems* (Oxford University Press, 1996).
- Dareau, M. 'DOST: its history and completion', *Dictionaries*, 23 (2002), 208–31.

- Dareau, M., and I. Macleod. 'Dictionaries of Scots', in Cowie (ed.), *Oxford History of English Lexicography*, 1.302–25.
- Datla, K. S. *The Language of Secular Islam: Urdu Nationalism and Colonial India* (Honolulu: University of Hawai'i Press, 2013).
- Datla, K. S. 'Making a worldly vernacular', unpublished PhD dissertation, University of California, Berkeley (2006).
- Debergh, M. 'Les débuts des contacts linguistiques entre l'occident et le Japon', *Langages*, 68 (1982), 27–44.
- De Cesare, R. 'Su di un gruppo di glosse antico-francesi in un lessico latino dell'xi secolo', *Aevum*, 27 (1953), 430–48.
- DeFrancis, J. *The Chinese Language: Fact and Fantasy*, paperback edn (Honolulu: University of Hawai'i Press, 1986).
- DeFrancis, J. *Nationalism and Language Reform in China* (Princeton University Press, 1950).
- Dehkhoda Lexicon Institute*, <https://icps.ut.ac.ir/about-lexicon-e.html>.
- De Kind, J., G.-M. de Schryver, and K. Bostoen. 'Pushing back the origin of Bantu lexicography: the *Vocabularium Congense* of 1652, 1928, 2012', *Lexikos*, 22 (2012), 159–94.
- Delling, G. 'Das erste griechisch-lateinisch Wörterbuch zum Neuen Testament', *Novum Testamentum*, 18.3 (1976), 213–40.
- Delsaerd, P. 'De kennismachines van Cornelis Kiliaan: een typografische analyse van drie Nederlandse woordenboeken uit de 16de eeuw', in T. Van Hal, L. Isebaert, and P. Swiggers (eds.), *De tuin der talen: Taalstudie en taalcultuur in de Lage Landen, 1450–1750* (Leuven etc.: Peeters, 2013), 93–121.
- Den Besten, H. 'A badly harvested field: the growth of linguistic knowledge and the Dutch Cape Colony until 1796', in Huigen et al. (eds.), *The Dutch Trading Companies as Knowledge Networks*, 267–94.
- Denis, D., and A. D'Arcy. 'Settler colonial Englishes are distinct from postcolonial Englishes', *American Speech*, 93 (2018), 3–31.
- Department of Bantu Education. *SiSwati Orthography: Terminology and Spelling Rules* (Pretoria: Government Printer, 1980).
- de Reuse, W. J. 'Apache names in Spanish and early Mexican documents: what they can tell us about the early contact Apache dialect situation' in Seymour (ed.), *From the Land of Ever Winter to the American Southwest*, 271–85.
- de Reuse, W. J. 'A new classification of Western Apache dialects', paper presented at the VIII Encuentro de Lingüística en el Noroeste, Universidad de Sonora, Hermosillo, Mexico, November 2004.
- de Reuse, W. J. 'Tonto Apache and its position within Apachean' in G. Holton (ed.), *Proceedings of the 2002 Athabaskan Languages Conference* = Alaska Native Language Center Working Papers, 2 (Fairbanks: Alaska Native Language Center, University of Alaska, 2002), 78–90.
- de Reuse, W. J. 'Yuman phonetic or phonological influences on Apachean', paper presented at the XI Encuentro Internacional de Lingüística en el Noroeste, Universidad de Sonora, Hermosillo, Mexico, November 19, 2010.
- de Reuse, W. J., with P. Goode. *A practical grammar of the San Carlos Apache language* = LINCOM Studies in Native American Linguistics, 51 (Munich: LINCOM, 2006).
- Derolez, R. (ed.). *Anglo-Saxon Glossography: Papers Read at the International Conference Held in the Koninklijke Academie voor Wetenschappen, Letteren, en Schone Kunsten van België*,

- Brussels, 8 and 9 September 1986 (Brussels: Koninklijke Academie voor Wetenschappen, Letteren, en Schone Kunsten van België, 1992).
- Dettori, A. 'Sardisch: Grammatikographie und Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, IV.913–35.
- Deufert, M. *Textgeschichte und Rezeption der plautinischen Komödien im Altertum* (Berlin and New York: De Gruyter, 2002).
- Die deutsche Literatur des Mittelalters: Verfasserlexikon*, 2nd edn, ed. Kurt Ruh et al., 14 vols. (Berlin and New York: De Gruyter, 1978–2008).
- DeVotta, N. 'Ethnolinguistic nationalism and ethnic conflict in Sri Lanka' in M. E. Brown and Š. Ganguli (eds.), *Fighting Words: Language Policy and Ethnic Relations in Asia* (Cambridge, MA: MIT Press, 2002), 105–39.
- Dhadphale, M. G. *Synonymic Collocations in the Tipiṭaka: A Study* (Poona: Bhandarkar Oriental Research Institute, 1980).
- Dickey, E. *Ancient Greek Scholarship: A Guide to Finding, Reading, and Understanding Scholia, Commentaries, Lexica, and Grammatical Treatises, from their Beginnings to the Byzantine Period* (Oxford University Press, 2007).
- Dickey, E. 'Columnar translation: an ancient interpretive tool that the Romans gave the Greeks', *Classical Quarterly*, 65 (2015), 807–21.
- Dickey, E. 'The creation of Latin teaching materials in antiquity: a re-interpretation of P. Sorb. inv. 2069', *Zeitschrift für Papyrologie und Epigraphik*, 175 (2010), 188–208.
- Dietrich, W. 'Die Lexikographie des Tupi-Guaraní', in Hausmann et al. (eds.), *Wörterbücher*, III.2670–6.
- Dietrich, W. 'O conceito de "Língua Geral" à luz dos dicionários de língua geral existentes', *Documentação de Estudos em Linguística Teórica e Aplicada (DELTA)*, 30 (2014), 591–622.
- Dillery, J. 'Literary interaction between Greece and Egypt: Manetho and synchronism' in Rutherford (ed.), *Greco-Egyptian Interactions*, 107–38.
- Di Marco, A. 'Per un riesame della tradizione umanistica di Festo', *Res publica litterarum*, 38 (2015), 35–61.
- Dionisotti, A. C. 'From Ausonius' schooldays? A schoolbook and its relatives', *Journal of Roman Studies*, 72 (1982), 83–125.
- Dionisotti, A. C. 'Greek grammars and dictionaries in Carolingian Europe' in M. W. Herren with S. A. Brown (eds.), *The Sacred Nectar of the Greeks: The Study of Greek in the West in the Early Middle Ages* (London: King's College, 1988), 1–56.
- Dionisotti, A. C. 'On the Greek studies of Robert Grosseteste' in A. C. Dionisotti, A. Grafton, and J. Kraye (eds.), *The Uses of Greek and Latin: Historical Essays* (London: Warburg Institute, 1988), 19–39.
- Dionisotti, A. C. 'On the nature and transmission of Latin glossaries', in Hamesse (ed.), *Les manuscrits des lexiques et glossaires*, 205–52; reprint *Ashgate Critical Essays* I.13–60.
- Dirks, N. B. *Castes of Mind: Colonialism and the Making of Modern India* (Princeton University Press, 2001).
- Dixon, R. M. W. *Australian Languages: Their Nature and Development* (Cambridge University Press, 2002).
- Dixon, R. M. W. *The Languages of Australia* (Cambridge University Press, 1980).
- Dizionario Biografico degli Italiani*, ed. A. M. Ghisalberti et al., 100 vols. projected (Rome: Istituto della Enciclopedia italiana, 1960–), www.treccani.it/biografico/.

- Doi, T. *The Study of Language in Japan: A Historical Survey* (Tokyo: Shinozaki shorin, 1976).
- Doke, C. M. 'Bantu language pioneers of the nineteenth century', *African Studies*, 18 (1959), 1–27.
- Doke, C. M. *Bantu: Modern Grammatical, Phonetical, and Lexicographical Studies Since 1860* (London: International African Institute, 1945).
- Doke, C. M. 'Early Bantu literature: the age of Brusciotto', *Bantu Studies*, 9 (1935), 87–114.
- Dollinger, S., and L. von Schneidmeyer. 'Canadianism, Americanism, North Americanism? A comparison of DARE and DCHP', *American Speech*, 86 (2011), 115–51.
- Donadoni, S. 'Gli Egiziani e le lingue degli altri', *Vicino Oriente*, 3 (1980), 1–14.
- Dormels, R. 'Hunmin-chôngum and the transcription of Chinese pronunciation in the King Sejong Era', in A. Fedotoff and S. Karteva-Dancheva (eds.), *Outlining the Korean Society* (proceedings volume, International Jubilee Conference on Korean Studies Dedicated to the 20th Anniversary of the Korean Studies Program) (Sofia University St Kliment Ohridski, 2015), 17–29.
- Dotan, A. *Nitsanim rishonim bə-ḥoxmat ha-millim: min ha-masora el rešit ha-milonaʿut ha-ivrit* (Jerusalem: Academy of Hebrew Language, 2005).
- Douglas, W. H. 'Alphabetising Bardi' in McGregor (ed.), *Studies in Kimberley Languages*, 189–204.
- Drachmann, A. B. *Die Überlieferung des Cyrillglossars* (Copenhagen: Levin & Munksgaard, 1935).
- Drimba, V. *Codex Comanicus. Édition diplomatique avec fac-similés* (Bucharest: Editura Enciclopedică, 2000).
- Duncan, J. B. 'Examinations and orthodoxy in Chosŏn Dynasty Korea' in B. A. Elman, J. B. Duncan, and H. Ooms (eds.), *Rethinking Confucianism: Past and Present in China, Japan, Korea, and Vietnam* (Los Angeles: UCLA Asian Pacific Monograph Series, 2002), 65–94.
- Duro, A. 'La lexicographie italienne du XXe siècle' in Hausmann et al. (eds.), *Wörterbücher*, II.1863–78.
- Duval, M. 'Korean lexicography', in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 900–10.
- Dwyer, R. *Gujarati: A Complete Course for Beginners* (Lincolnwood, IL: NTC, 1995).
- Dykstra, A. 'Lexicography of Modern West Frisian' in Munske (ed.), *Handbook of Frisian Studies*, 148–55.
- Echols, J. M. 'Presidential address: dictionaries and dictionary making, Malay and Indonesian', *Journal of Asian Studies*, 38.1 (1978), 11–24.
- Edel, E. *Die ägyptisch-hethitische Korrespondenz aus Boghazköi in babylonischer und hethitischer Sprache*, 2 vols. (Opladen: Westdeutscher, 1994).
- Edzard, L. 'Die SIG₇-ALAN = Nabnitu-Liste und das Konzept der semitischen Wurzel', *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 161 (2011), 17–37.
- Eickmans, H. 'Woordenboek der Nederlandsche Taal (WNT)' in Haß (ed.), *Große Lexika*, 271–91.
- Ellis, F. W. 'Note to the Introduction', in A. D. Campbell, *A Grammar of the Teloogoo Language* (Chennai: College, 1816), 1–31.
- Encyclopædia Iranica*, ed. E. Yarshater, 15 vols. and 2 fascicles to date, plus entries only available online (London etc.: [imprints vary], 1982–), www.iranicaonline.org.

- Encyclopedia of Ancient Greek Language and Linguistics*, ed. G. K. Giannakis et al., 3 vols. (Leiden and Boston: Brill 2014), referenceworks.brillonline.com/browse/encyclopedia-of-ancient-greek-language-and-linguistics.
- Encyclopedia of Arabic Language and Linguistics*, ed. K. Versteegh et al., 5 vols. (Leiden: Brill, 2006–9), referenceworks.brillonline.com/browse/encyclopedia-of-arabic-language-and-linguistics.
- Encyclopedia of Hebrew Language and Linguistics*, ed. G. Khan, 4 vols. (Leiden: Brill, 2012–13).
- Encyclopaedia of Islam*, 2nd edn, ed. P. J. Bearman et al., 12 vols. plus supplementary vols. (Leiden: Brill, 1954–2006), referenceworks.brillonline.com/browse/encyclopaedia-of-islam-2.
- Encyclopaedia of Islam* 3, ed. K. Fleet et al., in progress. (Leiden: Brill, 2007–), referenceworks.brillonline.com/browse/encyclopaedia-of-islam-3.
- Encyclopedia of Language and Linguistics*, 2nd edn, ed. K. Brown, 14 vols. (Amsterdam: Elsevier, 2006).
- Encyclopaedia of the Middle Ages*, ed. A. Vauchez, trans. A. Walford, 2002, www.oxfordreference.com.
- English Dialect Society. *Complete List of Publications 1873–1896 with Indexes* (London: Trübner, 1896).
- Englund, R. K. 'Texts from the late Uruk period' in P. Attinger and M. Wäflern (eds.), *Mesopotamien: Späturuk-Zeit und Frühdynastische Zeit: Annäherungen 1 = Orbis Biblicus et Orientalis*, 160/1 (Fribourg: Universitätsverlag, 1998), 15–233.
- Erbse, H. *Untersuchungen zu den attizistischen Lexika = Abhandlungen der Deutschen Akademie der Wissenschaften zu Berlin, Phil.-hist. Klasse, for 1949*, no. 2 (Berlin: Deutsche Akademie der Wissenschaften zu Berlin, 1950).
- Ernst, G. 'Romanian' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 687–701.
- Ersoy, A. 'Şemseddin Sami: Turkish Lexicon' in A. Ersoy, M. Górný and V. Kehriotis (eds.), *Modernism: Representation of National Culture. Discourses of Collective Identity in Central and Southeast Europe 1770–1945. Texts and Commentaries*, vol. III/2 (Budapest and New York: Central European University Press, 2010), 33–9.
- Eschbach-Szabo, V. 'La réflexion linguistique au Japon' in Auroux (ed.), *Histoire des idées linguistiques*, I.459–64.
- Esposito, E. 'Fragments of Greek lexicography in the papyri', *Trends in Classics*, 1 (2009), 251–93.
- Ethé, H. *Catalogue of Persian Manuscripts in the Library of the India Office*, 2 vols. (Oxford University Press, 1903).
- Ethnologue: Languages of the World*, 21st edn, ed. G. F. Simons and C. D. Fennig (Dallas: SIL International, 2018), www.ethnologue.com.
- Evans, J. G. *Report on Manuscripts in the Welsh Language*, vol. I part 3, *Peniarth* (London: HMSO, 1905).
- Eytan, E., et al. 'Hebrew language: modern period' in F. Skolnik (ed.), *Encyclopaedia Judaica* (2007), VIII.671–80.
- Ezquerria, M. A. 'Spanisch: Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, VI.1.636–51.
- Fabian, J. 'The other revisited: critical afterthoughts', *Anthropological Theory*, 6.2 (2006), 139–52.

- Fang, C. *The Asami Library: A Descriptive Catalog*, ed. E. Huff (Berkeley: University of California Press, 1969).
- Farber, W. 'Neues aus Uruk: Zur "Bibliothek des Iqīša"', *Die Welt des Orients*, 18 (January 1987), 26–42.
- Farina, D. M. T. Cr. 'Marrism and Soviet lexicography' in B. B. Kachru and H. Kahane (eds.), *Cultures, Ideologies, and the Dictionary: Studies in honor of Ladislav Zgusta* (Berlin and New York: Mouton de Gruyter, 1995), 153–70.
- Feder, F. 'Coptic translations' in A. Lange and E. Tov (eds.), *Textual History of the Bible: The Hebrew Bible*, vol. 1A, *Overview Articles* (Leiden and Boston: Brill, 2016), 331a–345b.
- Feder, F. 'Die koptische Übersetzung des Alten und Neuen Testaments im 4. Jahrhundert' in J. Tubach and S. Vashalomidze (eds.), *Stabilisierung und Profilierung der koptischen Kirche im 4. Jahrhundert* = Hallesche Beiträge zur Orientwissenschaft, 44 (Halle: Martin-Luther-Universität, 2007), 65–93.
- Fellman, J. *The Revival of a Classical Tongue: Eliezer Ben Yehuda and the Modern Hebrew Language* (The Hague: Mouton, 1973).
- Fenton, P.-B. 'Sa 'adyā ben Yōsēf' in *Encyclopaedia of Islam*, 2nd edn, VIII.661–2.
- Ferg, A., (ed.). *Western Apache Material Culture: The Goodwin and Guenther Collections* (Tucson: University of Arizona Press, 1987).
- Ferguson, C. A. 'Diglossia', *Word*, 15 (1959), 325–40.
- Ferguson, D. W. 'Robert Knox's Sinhalese vocabulary', *Journal of the Ceylon Branch of the Royal Asiatic Society*, 14 (1869), 155–200.
- Fernández Rodríguez, R. 'El Calepino ilocano (c. 1797) del P. Vivar: innovaciones lexicográficas y política traductora' in Zwartjes et al. (eds.), *Missionary Linguistics IV*, 249–72.
- Fernández Rodríguez, R. 'Lexicografía de la lengua ilocana: estudio de una obra manuscrita del siglo XVIII: el calepino ilocano', unpublished PhD thesis, Universidad de Valladolid (2012).
- Fernández Rodríguez, R. 'Lexicography in the Philippines (1600–1800)', *Historiographia Linguistica*, 41.1 (2014), 1–32.
- Filipović, R. 'The beginnings of lexicography in Croatia' in Hartmann (ed.), *History of Lexicography*, 65–73.
- Filliozat, J. 'Deux cents ans d'indianisme: critique des méthodes et des résultats', *Bulletin de l'Ecole française d'Extrême-Orient*, 76.1 (1987), 83–116.
- Findley, C.V. 'Sir James W. Redhouse (1811–1892): the making of a perfect Orientalist?', *Journal of the American Oriental Society*, 99.4 (1979), 573–600.
- Fishman, J. A., 'The lexicography of Yiddish' in Hausmann et al. (eds.), *Wörterbücher*, II.2246–54.
- Flajšhans, V. *Klaret i jeho družina*, 2 vols. (Prague: Nákladem České akademie věd a umění, 1926–8).
- Fonkič, B. L. *Grečeskie rukopisi evropejskich sobranij: paleografičeskie i kodikologičeskie issledovanija 1988–1998 gg./Manuscripts grecs dans les collections européennes: Études paléographiques et codicologiques 1988–1998* (Moscow: Indrik, 1999).
- Fonseca, M. do C. 'Fontes manuscritas e impressas da Central Library of Pangim: memória linguística de uma região Lusófona', in *Actas do 6º Colóquio anual da lusofonia, Bragança, Portugal 3–6 Outubro 2007* (Câmara Municipal de Bragança, 2007), 89–95, 231–4.

- Fonseca, M. do C. *Historiografia linguística portuguesa e missionária: preposições e posposições no séc. XVII* (Lisbon: Colibri, 2006).
- Fonseca, M. do C. 'Historiographie linguistique portugaise: le processus de grammatisation des langues extra-européennes' in G. Staudacher-Valliamée (ed.), *L'écriture et la construction des langues dans le sud-ouest de l'océan Indien* (Saint-Denis: Université de la Réunion; Paris: L'Harmattan, 2007), 181–90.
- Fortune, G. G. '75 years of writing in Shona', *Zambezia*, 1 (1969), 55–68.
- Fountain, C. 'Transculturation, assimilation, and appropriation in the missionary presentation of Nahuatl', in K. Zimmermann and B. Kellermeier-Rehbein (eds.), *Colonialism and Missionary Linguistics* (Berlin: De Gruyter, 2015), 177–97.
- Fournet, J.-L. 'The multilingual environment of Late Antique Egypt: Greek, Latin, Coptic, and Persian documentation' in Bagnall (ed.), *Oxford Handbook of Papyrology*, 418–51.
- Frahm, E. *Babylonian and Assyrian Text Commentaries: Origins of Interpretation* = Guides to the Mesopotamian Textual Record, 5 (Münster: Ugarit-Verlag, 2011)
- Franzen, *Middle English* = *Ashgate Critical Essays*, vol. II.
- Franzen, *Old English* = *Ashgate Critical Essays*, vol. I.
- Frawley, W., K. C. Hill, and P. Munro (eds.). *Making Dictionaries: Preserving Indigenous Languages of the Americas* (Berkeley, etc.: University of California Press, 2002).
- Freidkin, U., and M. Freidkin. 'Rav-Milim, Hamilon hashalem: Milon ivri khadash', *Leshonenu Laam*, 49 (1977), 3–25.
- Fressura, M. *Vergilius Latinograecus: corpus dei manoscritti bilingui dell'Eneide, parte prima* (1–8) (Rome and Pisa: Fabrizio Serra, 2017).
- Fumagalli, G. *Bibliografia etiopica* (Milan: Hoepli, 1893).
- Furuya Akihiro 古屋昭弘. "'Zhengzi tong" banben ji zuozhe kao' 《正字通》版本及作者考, *Zhongguo yuwen*, 247.4 (1995), 306–11.
- Gaeffke, P. 'A rock in the tides of time: Oriental Studies then and now', *Academic Questions*, 3.2 (1990), 67–74.
- Galambos, I. *Translating Chinese Tradition and Teaching Tangut Culture: Manuscripts and Printed Books from Khara-khoto* (Berlin: De Gruyter, 2015).
- A galáxia das línguas na época da expansão* (Lisbon: Comissão nacional para as comemorações dos descobrimentos portugueses, 1992).
- Gandolfo, M. D. 'Roman Slavdom' in G. Lepschy (ed.), *History of Linguistics*, vol. III, *Renaissance and Early Modern Linguistics* (London: Longman, 1998), 108–23.
- Ganz, D. 'The "Liber glossarum": a Carolingian encyclopedia', in P. L. Butzer and D. Lohrmann (eds.), *Science in Western and Eastern Civilization in Carolingian Times* (Basel: Birkhäuser, 1993), 127–35.
- García-Medall, J. *Vocabulários hispano-asiáticos: traducción y contacto intercultural* = Monográficos de la revista Hermeneus, 11 (Soria: Vertere, 2009).
- Gardiner, A. H. *Egyptian Grammar*, 3rd edn (Oxford: Griffith Institute, Ashmolean Museum, 1957).
- Garrison, M. B. 'Antiquarianism, copying, collecting' in D. T. Potts (ed.), *A Companion to the Archaeology of the Ancient Near East*, 2 vols. (Oxford: Wiley-Blackwell, 2012), I.27–47.
- Gaul, N. 'Moschopulos, Lopadiotes, Phrankopulos (?), Magistros, Staphidakes: Prosopographisches und Methodologisches zur Lexikographie des frühen 14. Jahrhunderts' in E. Trapp and S. Schönauer (eds.), *Lexicologica Byzantina: Beiträge zum*

- Kolloquium zur byzantinischen Lexikographie (Bonn, 13.–15. Juli 2007) (Göttingen: Vandenhoeck & Ruprecht, 2008), 163–96.
- Gaul, N. *Thomas Magistros und die spätbyzantische Sophistik: Studien zum Humanismus urbaner Eliten der frühen Palaiologenzeit* (Wiesbaden: Harrassowitz, 2011).
- Gaul, N. 'The twitching shroud: Collective construction of *paideia* in the circle of Thomas Magistros', *Segno e testo*, 5 (2007), 263–340.
- Gauton, R. 'Lexicography of the Nguni languages' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 912–29.
- Geller, M. J., and L. Vacin (eds.). *Healing Magic and Evil Demons: Canonical Udug-Hul Incantations* = *Die babylonisch-assyrische Medizin in Texten und Untersuchungen*, 8 (Berlin and Boston: De Gruyter, 2015).
- Geng Zhensheng 耿振生. *Ming Qing dengyunxue tonglun 明清等韵学通论* (Beijing: Yuwen chubanshe, 1992).
- Gesche, P. D. *Schulunterricht in Babylonien im ersten Jahrtausend v. Chr.* = *Alter Orient und Altes Testament*, 275 (Münster: Ugarit-Verlag, 2001).
- Ghatage, A. M. 'Traditional lexica and the Sanskrit Historical Dictionary', in Ghatage, R. N. Dandekar, and M. A. Mehendale (eds.), *Studies in Historical Sanskrit Lexicography* (Pune: Deccan College, 1973), 25–39.
- Gilles, H. 'Glosses, legal' in *Encyclopaedia of the Middle Ages*.
- Gilliver, P. *The Making of the Oxford English Dictionary* (Oxford University Press, 2016).
- Glassner, J.-J. *Writing in Sumer: The Invention of Cuneiform* (Baltimore: Johns Hopkins University Press, 2003).
- Glinert, L. (ed.). *Hebrew in Ashkenaz: A Language in Exile* (Oxford University Press, 1993).
- Glinert, L. 'Leshitot hamilonaut bemilon Even-Shoshan: normativiyut vedeskrptiviyut', *Mekhhkarim Belashon*, 2–3 (1987), 399–409.
- Glinert, L. 'Modern Hebrew lexicography: the last hundred years', *Jewish Book Annual*, 52 (1994), 37–58.
- Gnanapragasam, V. M. 'Contributions of Fr Beschi to Tamil', unpublished doctoral dissertation, University of Madras, 1965, shodhganga.inflibnet.ac.in/handle/10603/97704.
- Goddard, C., and N. Thieberger. 'Lexicographic research on Australian Aboriginal languages 1968–1993' in D. Tryon and M. Walsh (eds.), *Boundary Rider: Essays in Honour of Geoffrey O'Grady* (Canberra: Pacific Linguistics, 1997), 175–208.
- Godfrey, J. J. 'Sir William Jones and Père Coeurdoux: a philological footnote', *Journal of the American Oriental Society*, 87.1 (1967), 57–9.
- Goetz, G. 'Glossographie', in *Paulys Realencyclopädie*, VII.1.1433–66.
- Golden, P. B. 'Codex Cumanicus' in H. B. Paksoy (ed.), *Central Asian Monuments*, vlib.iue.it/carrie/texts/carrie_books/paksoy-2/cam2.html.
- Golden, P. B. 'Rasulid Hexaglot' in *Encyclopædia Iranica*.
- Golla, V. *California Indian Languages* (Berkeley: University of California Press, 2011).
- Gong, Y. *Die Namen der Keilschriftzeichen* = *Alter Orient und Altes Testament*, 268 (Münster: Ugarit-Verlag, 2000).
- Goodman, G. K. *Japan and the Dutch 1600–1853* (Richmond, UK: Curzon, 2000).
- Goodwin, G. *The Social Organization of the Western Apache* (University of Chicago Press, 1942; reprint 1969).
- Goshen-Gottstein, M. 'The lexicography of Hebrew', *Lexicographica*, 7 (1991), 72–93.
- Gottlieb, N. *Kanji Politics: Language Policy and Japanese Script* (London: Kegan Paul, 1995).

- Gouws, R. H. 'Aspects of Afrikaans lexicography' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 827–35.
- Gouws, R. H., U. Heid, W. Schweickard, and H. E. Wiegand (eds.). *Dictionaries. An International Encyclopedia of Lexicography, Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography* (Berlin and Boston: De Gruyter Mouton, 2013).
- Grafton, A. *Joseph Scaliger: A Study in the History of Classical Scholarship*, 2 vols. (Oxford: Clarendon Press, 1983–93).
- Gray, E. G. 'Missionary linguistics and the description of "exotic languages"' in Auroux et al. (eds.), *History of the Language Sciences*, 1.929–37.
- Green, N. 'The uses of books in a late Mughal *takiyya*: Persianate knowledge between person and paper', *Modern Asian Studies*, 44.2 (2010), 241–65.
- Greenbaum, A. 'The status of Hebrew in Soviet Russia from the revolution to the Gorbachev thaw' in Glinert (ed.), *Hebrew in Ashkenaz*, 242–8.
- Greenberg, M. 'The Illyrian movement: a Croatian vision of South Slavic unity', in J. Fishman and O. García (eds.), *Handbook of Language and Ethnic Identity*, 2 vols. (Oxford University Press, 2010–11), II.364–80.
- Grønbech, K. *Komanisches Wörterbuch: Türkischer Wortindex zu Codex Cumanicus* (Copenhagen: Munksgaard, 1942).
- Grønvik, O. 'The Saga of *Norsk Ordbok*: A scholarly dictionary for the Norwegian vernacular and the Nynorsk written language', *Kernerman Dictionary News*, 25 (July 2017), 19–24.
- Gros, F. 'Review of A. Dhamotharan, *Tamil Dictionaries: A Bibliography*', *Bulletin de l'École française d'Extrême-Orient*, 67 (1980), 346–58.
- Gross, A., Y. V. Kumaradoss, and H. Liebau (eds.). *Halle and the Beginning of Protestant Christianity in India* (Halle: Franckesche Stiftung, 2006).
- Groß, C. 'Les dictionnaires étymologiques et historiques des langues romanes' in Klump et al. (eds.), *Manuel des langues romanes*, 91–117.
- Grossman, A. *Haxme Tsarfat ha-rishonim: qorotehem, darkam bə-hanhagat ha-tsibur, yətsiratam ha-ruhanit* (Jerusalem: Magnes Press, 1995).
- Grubmüller, K. *Vocabularius ex quo: Untersuchungen zu lateinisch-deutschen Vokabularen des Spätmittelalters* (Munich: Beck, 1967).
- Grunhaus, N. *The Challenge of Received Tradition: Dilemmas of Interpretation in Radak's Biblical Commentaries* (Oxford University Press, 2013).
- Gruszczyński, W., and Z. Saloni. 'From multilingual to monolingual dictionaries: a historical overview of Polish lexicography', *Studies in Polish Linguistics*, 8.4 (2013), 205–27.
- 'Guinness World Records Announces Xinhua Dictionary as the "Most Popular Dictionary" and the "Best-Selling book (regularly updated)"'. 12 April 2016, www.guinnessworldrecords.com.
- Gül, B. 'İbni Mühenna Lügati'nin Türk ve Moğol dil araştırmalarındaki yeri ve önemi', *Türkbilig*, 19 (2010), 87–95.
- Gundersen, D. 'Norwegian lexicography' in Hausmann et al. (eds.), *Wörterbücher*, II.1923–8.
- Gunn, E. *Rewriting Chinese: Style and Innovation in Twentieth-Century Chinese Prose* (Stanford University Press, 1991).
- Gunnlaugur Ingólfsson. 'De første trykte islandske ordbøger', *Nordiske studier i leksikografi*, II (2012), 309–17.

- Guptan Nair, S. 'Introduction to Portuguese Malayalam Dictionary' in Pallath (ed.), *Arnos Padiri*, 34–45.
- Ġurāb, 'I. H. *al-Ma 'aġim al- 'Arabiyya: Riġla fi l-ġudūr wa-l-taṭawwur wa-l-huwiyya* (Dimyāṭ: Maktabat wa-Maṭba 'at Nānsī, 2005).
- Gurdon, P. R. T. 'Review of *Chandra-kanta abhidhan*', *Bulletin of the School of Oriental Studies, University of London*, 8.1 (1935), 235–7.
- Gurney, O. R. 'A list of copper objects', *Iraq*, 31 (1969), 3–7.
- Ha Gang-Jin. 'Pyojeeo daeyeog hanja-eoui tansaeng-gwa *hanbuljajeon-ui gachi*', *Cogito*, 80 (2016), 107–43.
- Haakh, N. M. 'Adolf Hitlers "Mein Kampf" und seine arabischen Übersetzungen', unpublished paper, Freie Universität Berlin, 2007, fu-berlin.academia.edu/NoraHaakh.
- Häberlein, M., and A. Keese (eds.). *Sprachgrenzen, Sprachkontakte, Kulturelle Vermittler: Kommunikation zwischen Europäern und Aussereuropäern (16.–20. Jahrhundert)* (Stuttgart: Steiner, 2010).
- Haensch, G. 'Katalanische Lexikographie' in Hausmann et al. (eds.), *Wörterbücher*, II.1770–88.
- Haensch, G. 'Spanische Lexikographie' in Hausmann et al. (eds.), *Wörterbücher*, II.1738–67.
- Haensch, G., and C. Omeñaca. *Los diccionarios del español en siglo XXI* (Salamanca: Ediciones Universidad, 1997).
- Hagemann, E. R. 'Surgeon Smart and the Indians: an 1866 Apache word-list', *Journal of Arizona History*, 11.2 (1970), 126–40.
- Hakala, W. N. 'Ajay čand nāma' in C. W. Ernst and F. Speziale (eds.), *Perso-Indica: The Sultanate Period and the Early Mughal Empire (Until 1556)* (Leiden: Brill, forthcoming).
- Hakala, W. N. 'The authorial problem in the *Khāliq bārī* of "Kḥusrau"', *Indian Economic & Social History Review*, 51.4 (2014), 481–96.
- Hakala, W. N. *Negotiating Languages: Urdu, Hindi, and the Definition of Modern South Asia* (New York: Columbia University Press, 2016).
- Hakala, W. N. 'Qaṣīda dar luġāt-i hindi' in C. W. Ernst and F. Speziale (eds.), *Perso-Indica: The Sultanate Period and the Early Mughal Empire (Until 1556)* (Leiden: Brill, forthcoming).
- Hālī, Ā. H. *Yādgar-i Ghālīb* (Kanpur: Nami Press, 1897).
- Hall, C. *Macaulay and Son: Architects of Imperial Britain* (New Haven: Yale University Press, 2012).
- Hamann, B. E. *The Translations of Nebrija: Language, Culture, and Circulation in the Early Modern World* (Amherst and Boston: University of Massachusetts Press, 2015).
- Hamesse, J. (ed.). *Les manuscrits des lexiques et glossaires de l'antiquité tardive à la fin du moyen âge* (Louvain-la-Neuve: Fédération Internationale des Instituts d'Études Médiévales, 1996).
- Hamilton, A. *The Copts and the West, 1439–1822: The European Discovery of the Egyptian Church* (Oxford University Press, 2006).
- Han Yong-un. 'Features and procedures of a unified and unabridged Korean dictionary, *Gyeoremal Keunsajeon*', *Lexicography: Journal of ASIALEX*, 3 (2016), 5–18.
- Handschriftencensus: Eine Bestandsaufnahme der handschriftlichen Überlieferung deutschsprachiger Texte des Mittelalters*, www.handschriftencensus.de.
- Hanka, V. *Sbirka nejdávnejších slovníků* (Prague: n.p., 1833).
- Hanks, P. 'Definitions and explanations' in Sinclair (ed.), *Looking Up*, 113–36.

- Hanks, P., and G.-M. de Schryver (eds.). *International Handbook of Modern Lexis and Lexicography* (Berlin and Heidelberg: Springer, print edition forthcoming, some chapters published online, 2015–).
- Hanzeli, V. E. *Missionary Linguistics in New France: A Study of Seventeenth- and Eighteenth-Century Descriptions of American Indian Languages* = *Janua Linguarum*, series maior, 29 (The Hague: Mouton, 1969).
- Harbsmeier, C. *Language and Logic* = J. Needham et al. (eds.), *Science and Civilisation in China*, vol. VII.1 (Cambridge University Press, 1998).
- Hardman-de-Bautista, M. 'Aymara lexicography' in Hausmann et al. (eds.), *Wörterbücher*, III.2684–90.
- Harlfinger, D. (ed.). *Griechische Handschriften und Aldinen: eine Ausstellung anlässlich der XV. Tagung der Mommsen-Gesellschaft in der Herzog August Bibliothek Wolfenbüttel* (Wolfenbüttel: Herzog August Bibliothek, 1978).
- Harshav, B. *Language in Time of Revolution* (Berkeley: University of California Press, 1993).
- Hartmann, R. R. K. (ed.). *The History of Lexicography* (Amsterdam and Philadelphia: Benjamins, 1986).
- Hartmann, R. R. K., and G. James. *Dictionary of Lexicography* (London and New York: Routledge, 2002).
- Hary, B., and M. Ángeles-Gallego. 'Lexicography and dialectology in Spanish *Magre Dardege*', in H. Ben-Shammai and B. Hary (eds.), *Esoteric and Exoteric Aspects in Judeo-Arabic Culture* (Leiden: Brill, 2006), 227–56.
- Hasitzka, M. *Neue Texte und Dokumentation zum Koptischunterricht* = *Mitteilungen aus der Papyrussammlung der Österreichischen Nationalbibliothek*, 18 (Vienna: Hollinek, 1990).
- Haslam, M. J. 'The Homer lexicon of Apollonius Sophista, I: composition and constituents', *Classical Philology*, 89 (1994), 1–45.
- Haß, U. 'Daniel Sanders: Wörterbuch der deutschen Sprache (1859–1865)' in Haß (ed.), *Große Lexika*, 253–69.
- Haß, U. (ed.). *Große Lexika und Wörterbücher Europas: Europäische Enzyklopädien und Wörterbücher in historischen Porträts* (Berlin and Boston: De Gruyter, 2012).
- Haß, U. 'In search of the European dimension of lexicography', unpublished lecture, Fifth International Conference for Historical Lexicography and Lexicology, Oxford, June 2010, www.linse.uni-due.de/linse/tagungen/Hass_European-dimension-of-lexicography_SHORT_FINAL-VERSION_fer_LINSE.pdf.
- Hassanein, A. T. 'Lexicography: monolingual dictionaries' in *Encyclopedia of Arabic Language and Linguistics*, III.37–45.
- Haßler, G. 'Las partículas entre la gramática y la lexicografía' in M. Bierbach, B. von Gemmingen, and Y. Stork (eds.), *Das gefesselte Wort: Beiträge zur Entwicklung von Wörterbüchern und Grammatiken des Spanischen/La palabra atada: Contribuciones sobre la evolución de diccionarios y gramáticas del español* (Bonn: Romanistischer, 2007), 86–101.
- Haß-Zumkehr, U. "'Alle welt erwartet hier eine erklärung von mir": Jacob Grimms Vorrede zum Deutschen Wörterbuch zwischen Apologie und Programm', *Zeitschrift für germanistische Linguistik*, 25 (1997), 1–23.
- Haß-Zumkehr, U. *Deutsche Wörterbücher: Brennpunkt von Sprach- und Kulturgeschichte* (Berlin: De Gruyter, 2001).
- Haß-Zumkehr, U. 'Propagandainstrument Wörterbuch: zur lexikografischen Methodik im Nationalsozialismus' in H. E. Wiegand (ed.), *Wörterbücher in der Diskussion IV*:

- Vorträge aus dem Heidelberger Lexikographischen Kolloquium (Tübingen: Niemeyer, 2000), 135–53.
- Haugen, E. I. 'Introduction' in E. I. Haugen, *A Bibliography of Scandinavian Dictionaries* (White Plains, NY: Kraus, 1984), 1–61.
- Hausmann, F. J. 'Wörterbuch und Wahrheit: zur Rezeption des Wörterbuchs der deutschen Gegenwartssprache in der Bundesrepublik' in H. Malige-Klappenbach, *Das Wörterbuch der deutschen Gegenwartssprache* (Tübingen: Niemeyer, 1986), 175–192.
- Hausmann, F. J., O. Reichmann, H. E. Wiegand, and L. Zgusta (eds.). *Wörterbücher/ Dictionaries/ Dictionnaires: Ein internationales Handbuch zur Lexicographie/ An International Encyclopedia of Lexicography/ Encyclopédie internationale de lexicographie*, 3 vols. (Berlin: De Gruyter, 1989–91).
- Haviland, J. B. 'A last look at Cook's Guugu Yimidhirr word list', *Oceania*, 44.3 (1974), 216–32.
- Haywood, J. A. *Arabic Lexicography: Its History and Its Place in the General History of Lexicography*, 2nd edn (Leiden: Brill, 1965).
- Haywood, J. A. 'Arabic lexicography' in Hausmann et al. (eds.), *Wörterbücher*, III.2438–48.
- Haywood, J. A. 'Bilingual lexicography with Arabic' in Hausmann et al. (eds.), *Wörterbücher*, III.3086–96.
- Hazai, G. 'Al-Kāshgharī' in *Encyclopaedia of Islam*, 2nd edn, IV.699–701.
- Heestermans, H. 'Niederländische Lexikographie und Lexikographie des Afrikaans' in Hausmann et al. (eds.), *Wörterbücher*, II.2010–22.
- Heo Jae-Young. 'Gug-eosajeon pyeonchan jeongchaeg mich geu yeogsa', *Gug-eogyoyug Yeongu*, 15 (2005), 57–83.
- Heo Jae-Young. 'Gug-eosajeon pyeonchan yeongusa', *Journal of Korealex*, 13 (2009), 286–312.
- Hercus, L. 'Looking at some details of Reuther's work' in Peterson and Kenny (eds.), *German Ethnography in Australia*, 115–35.
- Hermanson, E. A. 'Colenso's first attempt at Bible translation in Zulu', in J. A. Draper (ed.), *The Eye of the Storm: Bishop John William Colenso and the Crisis of Biblical Interpretation* (London and New York: Clark, 2003), 5–28.
- Hernández, E. 'En torno al vocabulario hispano-maya conservado en la biblioteca John Carter Brown (*Codex Indicus* 8)', *Revista de lexicografía*, 14 (2008), 111–22.
- Hernández, E. 'Indigenismos en el Vocabulario de la lengua Cakchiquel atribuido a fray Domingo de Vico, MS. BNF R. 7507', *Revista de filología española*, 88.1 (2008), 67–88.
- Hernández, E. *Lexicografía hispano-amerindia 1550–1800: catálogo descriptivo de los vocabularios con las lenguas indígenas americanas* (Madrid: Iberoamericana; Frankfurt am Main: Vervuert, 2018).
- Hernández, E. 'La marca de uso metafórico para el náhuatl: una innovación de la lexicografía hispano-amerindia', *Colonial Latin American Review*, 24.3 (2015), 318–35.
- Hernández, E. 'Los vocabularios Hispano-Mayas del siglo XVI' in Zwartjes et al. (eds.), *Missionary Linguistics* IV, 129–49.
- Hewson, J. 'The study of the native languages of North America: the French tradition' in Auroux et al. (eds.), *History of the Language Sciences*, 1.966–73.
- Hida Yoshifumi, Matsui Shigekazu, and Sakaida Toshinobu (eds.). *Meiji-ki kokugo jisho taiki bekkann: Shoshi to kenkyū* (Tokyo: Ōzorasha, 2003).

- Hildebrand, P. *Le martyr Georges de Geel et les débuts de la mission du Congo (1645–1652)* (Antwerp: Archives des Capucins, 1940).
- Hilgert, M. 'Von "Listenwissenschaft" und "epistemischen Dingen": Konzeptuelle Annäherungen an altorientalische Wissenspraktiken', *Journal for General Philosophy of Science*, 40 (2009), 277–309.
- Hill, B. S. 'Judeo-Slavic' in Kahn and Rubin (eds.), *Handbook of Jewish Languages*, 599–617.
- Hill, K. C. 'On publishing the Hopi Dictionary' in Frawley et al. (eds.), *Making Dictionaries*, 299–311.
- Hjorth, P. L. 'Danish lexicography' in Hausmann et al. (eds.), *Wörterbücher*, II.1913–22.
- Hodgson, M. G. S. *The Venture of Islam: Conscience and History in a World Civilization*, 3 vols. (University of Chicago Press, 1974).
- Hoffmann, F. *Ägypten: Kultur und Lebenswelt in griechisch-römischer Zeit: eine Darstellung nach den demotischen Quellen* (Berlin: Akademie Verlag, 2000).
- Holm, L., and H. Jonsson, 'Swedish lexicography' in Hausmann et al. (eds.), *Wörterbücher*, II.1933–43.
- Holtus, G., M. Metzeltin, and C. Schmitt (eds.). *Lexikon der romanistischen Linguistik*, 8 vols. in 12 parts (Tübingen: Niemeyer, 1988–2005).
- Hong Kisoon. 'Ppuijillo nohansajeon-e gwanhan ilgo', *Korean Journal of Russian Language and Literature*, 23.3 (2011), 325–51.
- Hoogland, J. C. 'Lexical gaps in Arabic lexicography with evidence from Arabic dictionaries' in E. Ditters and H. Motzki (eds.), *Approaches to Arabic Linguistics, Presented to Kees Versteegh on the Occasion of His Sixtieth Birthday* (Leiden: Brill, 2007), 455–73.
- Hoogland, J. C. 'Lexicography: bilingual dictionaries' in *Encyclopedia of Arabic Language and Linguistics*, III.21–30.
- Hruša, I. *Die akkadische Synonymenliste malku = šarru: Eine Textedition mit Übersetzung und Kommentar = Alter Orient und Altes Testament*, 50 (Münster: Ugarit-Verlag, 2010).
- Huang Fushan 黃復山. Wang Anshi 'Zishuo' zhi yanjiu 王安石《字說》之研究 (Taipei: Hua Mulan wenhua chubanshe, 2008).
- Hucker, C. O. *A Dictionary of Official Titles in Imperial China* (Stanford University Press, 1985).
- Huigen, S. 'Introduction' in Huigen et al. (eds.), *The Dutch Trading Companies as Knowledge Networks*, 1–17.
- Huigen, S., J. L. de Jong, and E. Kolfin (eds.). *The Dutch Trading Companies as Knowledge Networks* (Leiden and Boston: Brill, 2010).
- Hüllen, Werner. *English Dictionaries, 800–1700: The Topical Tradition*, revised edn (Oxford: Clarendon Press, 2006).
- Hunger, H. *Die hochsprachliche profane Literatur der Byzantiner*, vol. II, *Philologie, Profandichtung, Musik, Mathematik und Astronomie, Naturwissenschaften, Medizin, Kriegswissenschaft, Rechtsliteratur* (Munich: Beck, 1978).
- Hunt, T. *Teaching and Learning Latin in Thirteenth-Century England*, 3 vols. (Cambridge: Brewer, 1991).
- Ibrahim, Z. 'Lexical variation: Modern Standard Arabic' in *Encyclopedia of Arabic Language and Linguistics*, III.13–21.
- Idema, W. L. *Chinese Vernacular Fiction: The Formative Period* (Leiden: Brill, 1974).

- Iima Hiroaki. *Kokugo jiten no yukue* (Tokyo: NHK shuppan, 2017).
- Illing, K. 'Abstractum-Glossar' in *Die deutsche Literatur des Mittelalters: Verfasserlexikon*, s.v.
- Inṇāci, Cū. *Caturakarāti Ārāycci* ('Studies in Caturakarāti') (Madurai: Madurai Kāmarācar Palkalaik Kaḷakam, n.d., but 1976 or later).
- Irākavaiaṇkāṛ, Mu. *Ārāycci Tokuti*, facs. reprint (Tanjore: Tamil University, 1984; ¹1938).
- Iszchanian, R. *Książka ormiańska w latach 1512–1920* (Kraków: Ossolineum, 1994).
- Jackson, K. *Language and History in Early Britain* (Edinburgh University Press, 1953).
- Jacques, R. *Portuguese Pioneers of Vietnamese Linguistics Prior to 1650*, trans. R. F. Uson and M. Uson (Bangkok: Orchid Press, 2002).
- James, G. 'Aspects of the structure of entries in the earliest missionary dictionary of Tamil' in Zwartjes et al. (eds.), *Missionary Linguistics IV*, 273–301.
- James, G. *Col-porūḷ: A History of Tamil Dictionaries* (Chennai: Cre-A, 2000).
- James, G. 'The manuscript and printed versions of Antão de Proença's *Vocabulario tamulico* (1679)', academia.edu/10107360.
- Jankowiak, L. A. 'Słowniki Bartłomieja z Bydgoszczy (1532 i 1544): stan badań i perspektywy badawcze', *Poznańskie Studia Polonistyczne*, 24.1 (2014), 11–26.
- Janson, T. 'Languages and language names in Mozambique, 150 years ago and now' in K. Bromber and B. Smieja (eds.), *Globalisation and African Languages: Risks and Benefits* (Berlin: De Gruyter, 2004), 297–311.
- Jarvis, B. 'Welsh humanist learning' in R. G. Gruffydd (ed.), *A Guide to Welsh Literature*, vol. 3, c. 1530–1700 (Cardiff: University of Wales Press, 1997), 128–53.
- Jauk-Pinhak, M. 'Some notes on the pioneer Indologist Filip Vesdin (Paulinus a Sancto Bartholomaeo)', *Indologica Taurinensia*, 12 (1984), 129–37.
- Jerchow, S. 'From the *Makre Dardeke* to *Sefer 'Arb'ah ve-'esrim*: the tradizione della traduzione and directions for research', in F. Aspesi et al. (eds.), *Il mio cuore è a Oriente/libbi bə-mizrah: studi di linguistica storica, filologia, e cultura ebraica dedicati a Maria Luisa Mayer Modena* (Milan: Cisalpino, 2008), 493–510.
- Ji Yonghai 季永海. 'Dà Qīng quánshū yánjiū' 《大清全书》研究. *Manyu yanjiu*, 2 (1990), 42–50.
- Jiāng Qiao 江桥. *Kāngxī 'Yùzhì Qīngwén jiàn'* 康熙《御制清文鉴》研究. (Beijing: Beijing Yanshan Chubanshe, 2009).
- Jiten kyōkai (ed.). *Nihon no jisho no ayumi* (Tokyo: Jiten kyōkai, 1996).
- Johnstone, B. *Speaking Pittsburghese: The Story of a Dialect* (New York: Oxford University Press, 2013).
- Jones, W. 'The Third Anniversary Discourse, on the Hindus, delivered 2d of February, 1786' in *The Works of Sir William Jones*, 13 vols. (London: Stockdale, 1807), III.24–46.
- Jones, W. J. *German Lexicography in the European Context* (Berlin: De Gruyter, 2000).
- Jones, W. J. 'Lingua Teutonum victrix? Landmarks in German lexicography (1500–1700)', *Histoire Épistémologie Langage*, 13.2 (1991), 131–52.
- Kaçalın, M. S. 'Abuşka sözlüğü üzerine', *Türk Dilleri Araştırmaları*, 13 (2003), 23–8.
- Kahan, M. 'Aspects of medieval lexicography: between Jonah ibn Janāh's Kitāb al-'Uṣūl and Joseph Kaspi's Šarṣot Kesef', *Revue des Etudes Juives*, 177.3–4 (2018), 297–316.
- Kahan, M. *Ha-ʿvirit bi-r'i ḥoxmat ha-higgayon: mišnato ha-logit ha-filosofit ha-balšanit šel Yosef ibn Kaspi* (Ramat Gan: Bar-Ilan University Press, 2018).
- Kahn, L., and A. D. Rubin (eds.). *Handbook of Jewish Languages* (Leiden: Brill, 2016).

- Kaiser, S. 'Linguistic thought in Japan' in E. F. K. Koerner and R. E. Asher (eds.), *Concise History of the Language Sciences: From the Sumerians to the Cognitivists* (New York: Pergamon, 1995), 45–51.
- Kämper, H., A. Klosa, and O. Vietze (eds.). *Aufklärer, Sprachgelehrter, Didaktiker: Johann Christoph Adelung (1732–1806)* (Tübingen: Narr, 2008).
- Kamusella, T. *The Politics of Language and Nationalism in Modern Central Europe* (Basingstoke and New York: Palgrave Macmillan, 2009).
- Kane, D. *The Sino-Jurchen Vocabulary of the Bureau of Interpreters* (Bloomington: Indiana University Research Institute for Inner Asian Studies, 1989).
- Kang Yi-Yon. '19segi huban Joseon-e payeondoan pali oebangeongyohoe seongyosadeul-ui Bulhansajeon (Dictionnaire français-coréen) yeongu', *Gyohoesayeongu*, 22 (2004), 173–211.
- Kant, I. *Physische Geographie*, 4 vols. in 7 parts (Mainz and Hamburg: Vollmer, 1801–5).
- Kaplony-Heckel, U. 'Schüler und Schulwesen in der Ägyptischen Spätzeit', *Studien zur Altägyptischen Kultur*, 1 (1974), 227–46.
- Kapranov, V. A. 'Lexicography' in C. E. Bosworth and M. S. Asimov (eds.), *History of Civilizations*, vol. IV, *The Age of Achievement: AD 750 to the End of the Fifteenth Century*, part 2, *The Achievements* (Paris: UNESCO, 2000), 351–63.
- Kara, Gy. 'Aramaic scripts for Altaic languages' in Daniels and Bright (eds.), *The World's Writing Systems*, 536–58.
- Kara, Gy. *Books of the Mongolian Nomads: More than Eight Centuries of Writing Mongolian*, trans. J. R. Krueger (Bloomington: Indiana University Research Institute for Inner Asian Studies, 2005).
- Karow, O. 'Die Wörterbücher der Heianzeit und ihre Bedeutung für die japanische Sprachgeschichte: Teil i, Das Wamyôruijushô des Minamoto no Shitagau', *Monumenta Nipponica*, 7.1/2 (1951), 156–97.
- Karpinska, L. *English–Latvian Lexicographic Tradition: A Critical Analysis* (Berlin and Boston: De Gruyter, 2015).
- Karttunen, F. 'Nahuatl lexicography' in Hausmann et al. (eds.), *Wörterbücher*, III.2657–61.
- Karunaratne, I. M. *Teaching of English: A Sociological Study* (New Delhi: APH, 2009).
- Kaske, E. 'National Language Movement' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, III.139–44.
- Kasser, R. 'Protodialectes Coptes à systèmes alphabétiques de type Vieux Copte', in M. Immerzeel and J. Van der Vliet (eds.), *Coptic Studies on the Threshold of a New Millennium: Proceedings of the Seventh International Congress of Coptic Studies, Leiden 2000* = *Orientalia Lovanensia Analecta* 133/1 (Leuven: Peeters, 2004), 77–123.
- Kaster, R. A. *Guardians of Language: The Grammarian and Society in Late Antiquity* (Berkeley and Los Angeles: University of California Press, 1988).
- Katre, S. M. 'Amara's contribution to Indian lexicography', in R. K. Mookerji (ed.), *Vikrama Volume* (Ujjain: Scindia Oriental Institute, 1948), 199–212.
- Katre, S. M. 'Lexicography of Old Indo-Aryan: Vedic and Sanskrit' in Hausmann et al. (eds.), *Wörterbücher*, III.2487–97.
- Kawase Kazuma 川瀬何一馬. *Kojisho no kenkyû 古辞書の研究* (Tokyo: Dai Nihon Yûbenkai Kôdansha, 1955).
- Kempgen, S., et al. (eds.). *Die slavischen Sprachen/The Slavic Languages*, 2 vols. (Berlin: Mouton de Gruyter, 2009–14).

- Kennedy, G. 'The monosyllabic myth', *Journal of the American Oriental Society*, 71.3 (1951), 161–6.
- Kenny, A. *The Aranda's Pepa: An Introduction to Carl Strehlow's Masterpiece Die Aranda- und Loritja-Stämme in Zentral-Australien (1907–1920)* (Canberra: ANU Press, 2013).
- Kenny, A. 'Early ethnographic work at the Hermannsburg Mission in Central Australia, 1877–1910' in Peterson and Kenny (eds.), *German Ethnography in Australia*, 169–93.
- Kenny, A. 'From missionary to frontier scholar: an introduction to Carl Strehlow's masterpiece, *Die Aranda- und Loritja-Stämme in Zentral-Australien (1901–1909)*', unpublished PhD thesis, University of Sydney (2008).
- Keshavmurthy, P. 'Mīrzā Ḥān ibn Fahr al-Dīn, Tuḥfat al-hind' in *Perso-Indica: An Analytical Survey of Persian Works on Indian Learned Traditions*, ed. F. Speziale and C. W. Ernst (2010–), *perso-indica.net*.
- Khan, M. S. 'The early history of Bengali printing', *Library Quarterly*, 32.1 (1962), 51–61.
- Kho Songmoo. *Korean Studies in Russia (1677–1930)* = *Studia Orientalia*, 51.12 (Helsinki: Finnish Oriental Society, 1980).
- Kholkina, L. 'Lexicography, modern' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, II.594–600.
- Khoury, M. 'Dictionnaires arabes bilingues: présentation historique et étude comparative', unpublished MA thesis, University of Ottawa (1996).
- Kida Jun'ichirō (ed.). *Dai kanwa jiten o yomu* (Tokyo: Taishūkan shoten, 1986).
- Kiesler, R. A. 'Iberoromance II: synchronic lexicography' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 645–53.
- Kilian, J. 'Hermann Paul: Deutsches Wörterbuch' in Haß (ed.), *Große Lexika*, 317–33.
- Kilpatrick, H., and G. J. Toomer. 'Niqlāwus al-Ḥalabī (c. 1611–c. 1661): a Greek Orthodox Syrian copyist and his letters to Pococke and Golius', *Lias*, 43.1 (2016), 1–159.
- Kim Jonghwan and Naeree Han. 'The development of Korean online dictionaries: a case study of Naver dictionary services', *Lexicography: Journal of ASIALEX*, 3 (2016), 19–37.
- Kim-Renaud, Y.-K. (ed.). *The Korean Alphabet: Its History and Structure* (Honolulu: University of Hawai'i Press, 1997).
- Kin Bunkyo. *Literary Sinitic and East Asia: The Cultural Sphere of Reading by Gloss*, ed. R. King, trans. R. King et al. (Leiden: Brill, forthcoming).
- King, C. R. *One Language, Two Scripts: The Hindi Movement in the Nineteenth Century North India* (Mumbai: Oxford University Press, 1994).
- Kinkade, M. D. 'History of research in linguistics' in W. Suttles (ed.), *Handbook of North American Indians*, vol. VII, *Northwest Coast* (Washington, DC: Smithsonian Institution, 1990), 98–106.
- Kinra, R. 'Cultures of comparative philology in the early modern Indo-Persian world', *Philological Encounters*, 1 (2016), 225–87.
- Kirkness, A. 'Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm' in Haß (ed.), *Große Lexika*, 211–32.
- Kishimoto, E. 'The adaptation of the European polyglot dictionary of Calepino in Japan: *Dictionarium latino lusitanicum, ac iaponicum* (1595)' in Zwartjes and Altman (eds.), *Missionary Linguistics II*, 205–23.
- Kishimoto, E. 'The process of translation in *Dictionarium latino lusitanicum, ac iaponicum*', *Journal of Asian and African Studies*, 72 (2006), 17–26.

- Kiwitt, M., and S. Dörr, 'Judeo-French' in Kahn and Rubin (eds.), *Handbook of Jewish Languages*, 138–77.
- Klöter, H. 'Chinese lexicography', in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 884–93.
- Klöter, H. 'Dialect dictionaries' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, II.50–4.
- Klöter, H. *The Language of the Sangleys: A Chinese Vernacular in Missionary Sources of the Seventeenth Century* (Leiden and Boston: Brill, 2011).
- Klöter, H. "'What is correct Chinese?'" revisited in I. Tieken-Boon van Ostade and C. Percy (eds.), *Prescription and Tradition in Language: Establishing Standards Across Time and Space* (Clevedon: Multilingual Matters, 2016), 57–70.
- Klöter, H. *Written Taiwanese* (Wiesbaden: Harrassowitz, 2005).
- Klump, A., J. Kramer, and A. Willems (eds.). *Manuel des langues romanes* (Berlin and Boston: De Gruyter, 2014).
- Knappe, G. *Idioms and Fixed Expressions in English Language Study Before 1800* (Frankfurt am Main: Lang, 2004).
- Knappe, G. 'Theory meets empiricism: English lexis in John Wilkins' philosophical language and the role of William Lloyd', *Archiv für das Studium der neueren Sprachen und Literaturen*, 241 (2004), 69–89.
- Koch, H. 'R. H. Mathews' schema for the description of Australian languages' in McGregor (ed.), *Encountering Aboriginal Languages*, 179–218.
- Koerner, E. F. K., and A. J. Szwedek (eds.). *Towards a History of Linguistics in Poland: From the Early Beginnings to the End of the Twentieth Century* (Amsterdam and Philadelphia: Benjamins, 2001).
- Konno Shinji. *Genkai o yomu: Kotoba no umi to Meiji no Nihongo* (Tokyo: Kadokawa shoten, 2014).
- Kopf, L. *Studies in Arabic and Hebrew Lexicography* (partly in Hebrew), ed. M. H. Goshen-Gottstein with S. Assif (Jerusalem: Magnes, 1976).
- Kornicki, P. F. *The Book in Japan: A Cultural History From the Beginnings to the Nineteenth Century*, paperback edn (Honolulu: University of Hawai'i Press, 2001; 1998).
- Kornicki, P. F. *Languages, Scripts, and Chinese Texts in East Asia* (Oxford University Press, 2018).
- Kramer, J. 'Ladinisch: Grammatikographie und Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, III.757–63.
- Kramer, J. 'Die Lexikographie des Provençalischen, Rätoromanischen, Sardischen und Dalmatischen' in Hausmann et al. (eds.), *Wörterbücher*, II.1891–1905.
- Krauss, M. E. 'Eskimo-Aleut' in T. A. Sebeok (ed.), *Native Languages of the Americas* (New York: Plenum, 1976), 175–281.
- Krauss, S. *Griechische und lateinische Lehnwörter im Talmud, Midrasch, und Targum*, with notes by I. Löw, 2 vols. (Berlin: Calvary, 1898–9).
- Krishnamurti, B. 'Dravidian lexicography' in Hausmann et al. (eds.), *Wörterbücher*, III.2521–33.
- Krome, S. 'Das Wörterbuchprogramm WAHRIG bei BROCKHAUS', *Der Sprachdienst* (2014), part 5–6, 209–24.
- Kroskrity, P. V. 'Language renewal and the technologies of literacy and postliteracy: Reflections from Western Mono' in Frawley et al. (eds.), *Making Dictionaries*, 171–92.

- Kurashima Nagamasa. *Kokugo to Kokugo jiten no jidai*, 2 vols. (Tokyo: Shōgakkan, 1997).
- Kutscher, E. Y. *A History of the Hebrew Language*, ed. R. Kutscher (Jerusalem: Magnes, 1982).
- Kwanten, L. *The Timely Pearl: A 12th Century Tangut–Chinese Glossary*, vol. I, *The Chinese Glosses* (Bloomington: Research Institute for Inner Asian Studies, Indiana University, 1982).
- Kwok, B., B. T'sou, and A. C. Chin, 'Yuè 粵 dialects' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, IV.690–702.
- Kychanov, E. I. 'Tangut' in Daniels and Bright (eds.), *The World's Writing Systems*, 228–30.
- Labarre, A. *Bibliographie du Dictionarium d'Ambrogio Calepino (1502–1779)* (Baden-Baden: Koerner, 1975).
- Lambert, J. 'A much tortured expression: a new look at *Hobson-Jobson*', *International Journal of Lexicography*, 27 (2014), 54–88.
- Lancashire, I. 'The perils of firsts: Dating Rawlinson MS Poet. 108 and tracing the development of monolingual English lexicons' in A. Curzan and K. Emmons (eds.), *Studies in the History of the English Language II: Unfolding Conversations* (Berlin and New York: Mouton de Gruyter, 2004), 229–72.
- Landau, S. I. 'Major American dictionaries' in Cowie (ed.), *Oxford History of English Lexicography*, I.182–229.
- Langenscheidt Digital. 'Der Große Muret-Sanders jetzt online!' (2 December 2014), www.langenscheidt.de/Pressemeldungen/Der-Grosse-Muret-Sanders-jetzt-online.
- Lapidge, M. 'The career of Aldhelm', *Anglo-Saxon England*, 36 (2007), 15–69.
- Lapidge, M. 'The school of Theodore and Hadrian', *Anglo-Saxon England*, 15 (1986), 45–72 (reprint *Ashgate Critical Essays* I.63–92).
- Lapidge, M., and B. Bischoff. *Biblical Commentaries from the Canterbury School of Theodore and Hadrian* (Cambridge University Press, 1994).
- Lara, L. F. *Historia mínima de la lengua española* (Mexico City: Colegio de México, 2013).
- Lau, C., 'Hakka dialects' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, II.341–9.
- Law, V. 'Indian influence on early Arab phonetics – or coincidence?' in M. G. Carter and K. Versteegh (eds.), *Studies in the History of Arabic Grammar II: Proceedings of the 2nd Symposium on the History of Arabic Grammar, Nijmegen, 27 April–1 May 1987* (Amsterdam: Benjamins, 1990), 215–27.
- Ledgeway, A., and M. Maiden (eds.). *The Oxford Guide to the Romance Languages* (Oxford University Press, 2016).
- Ledyard, G. 'The international linguistic background of the correct sounds for the instruction of the people' in Kim-Renaud (ed.), *Korean Alphabet*, 31–87.
- Lee Eun-Ryoung. 'Palioebangjeongyohoeui *Hanbuljaejon* (1880)gwa *Bulhanjaejon* (1869) bigyoyeongu', *Cogito*, 80 (2016), 67–106.
- Lee Gimun 李基文 [Ki-Moon Lee]. 'Hunmin jeong-eum changje e daehan jaejomyeong' 訓民正音創制에 대한 再照明, *Hangugeo yeongu*, 5 (2008), 5–45.
- Lee, K.-M. *Geschichte der koreanischen Sprache*, ed. B. Lewin, trans. D. Eikemeier et al. (Wiesbaden: Reichert, 1977).
- Lee, K.-M. 'The inventor of the Korean alphabet' in Kim-Renaud (ed.), *Korean Alphabet*, 11–30.
- Lee, P. 'A brief history of Korean lexicography', *Seoul Journal of Korean Studies*, 7 (1994), 129–45.

- Lee Pyeonggeun. 'Silhak sidae ui eoneo yeongu', *Hanguksa simin gangjiwa*, 48 (2011), 113–33.
- Lee, S.-Y. 'Defining correctness: the tale of the *Contemporary Chinese Dictionary*', *Modern China*, 40.4 (2014), 426–50.
- Lee Sang-Oak. 'North Korean lexicology and lexicography', *Seoul Journal of Korean Studies*, 6 (1993), 165–86.
- Lee Yeounsuk. *The Ideology of Kokugo: Nationalizing Language in Modern Japan*, trans. M. H. Hubbard (Honolulu: University of Hawai'i Press, 2010).
- Legan Ravnika, A. 'Characteristics of the technical terminology in Vodnik's Slovene–German dictionary (1804–1806)', *Slovenski jezik/Slovene Linguistic Studies*, 8 (2011), 79–93.
- Lehner, G. *Der Druck chinesischer Zeichen in Europa: Entwicklungen im 19. Jahrhundert* (Wiesbaden: Harrassowitz, 2004).
- Lehrs, K. *De Aristarchi studiis Homerici* (Leipzig: Hirzel, 1882).
- Lendinara, P. 'Contextualized lexicography', in K. O'Brien O'Keefe and A. Orchard (eds.), *Latin Learning and English Lore: Studies in Anglo-Saxon Literature for Michael Lapidge*, 2 vols. (University of Toronto Press, 2005), 1.108–31 (reprint *Ashgate Critical Essays* 1.431–54).
- Lenzi, A. *Secrecy and the Gods: Secret Knowledge in Ancient Mesopotamia and Biblical Israel* = State Archives of Assyria Studies, 19 (Helsinki: Neo-Assyrian Text Corpus Project, 2008).
- Lepsius, K. R. *Das allgemeine linguistische Alphabet: Grundsätze der Übertragung fremder Schriftsysteme und bisher noch ungeschriebener Sprachen in europäische Buchstaben* (Berlin: Hertz, 1854).
- Lepsius, K. R. *Standard Alphabet for Reducing Unwritten Languages and Foreign Graphic Systems to a Uniform Orthography in European Letters* (London: Seelys, 1855).
- Leuker, M. T. 'Knowledge transfer and cultural appropriation: Georg Everhard Rumphius's "D'Amboinsche Rariteitkamer" (1705)' in Huigen et al. (eds.), *The Dutch Trading Companies as Knowledge Networks*, 145–70.
- Lewis, G. *The Turkish Language Reform: A Catastrophic Success* (Oxford University Press, 1999).
- Lewis, N. *Greeks in Ptolemaic Egypt: Case Studies in the Social History of the Hellenistic World* (Oxford: Clarendon Press, 1986).
- Lexicon of Greek Grammarians of Antiquity*, ed. F. Montanari, online only (Leiden and Boston: Brill, in progress), referenceworks.brillonline.com/browse/lexicon-of-greek-grammarians-of-antiquity.
- Li, J. 'Rime (rhyme)' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, III.591–4.
- Li Jiankun 李建昆. 'Lidai xuezhe dui Han Yu shi zhi pingjia' 歷代學者對韓愈詩之評價 in *Han Meng shi luncong* 韓孟詩論叢 (Taipei: Xiuwei zixun keji gufen youxian gongsi, 2005).
- Lǐ Y. 'Tuīgǎng pǔtōnghuà: pǔjí yǔ tígāo bǐngxíng' in *Zhōngguó yǔyán shēnghuó zhuàngkuàng bàogào* (2017)/*Language Situation in China: 2017* (Beijing: Shāngwù Yīnshūguǎn, 2017), 3–6.
- Liebau, H. *Cultural Encounters in India: The Local Co-Workers of the Tranquebar Mission, 18th to 19th Centuries*, trans. R. V. Rajan (New Delhi: Social Science, 2013).
- Liebau, H. *Die indischen Mitarbeiter der Tranquebarmission (1706–1845): Katecheten, Schulmeister, Übersetzer* (Tübingen: Niemeyer for the Franckesche Stiftung Halle, 2008).

- Lieberman, S. *Yəḡanit və-yaynut be-Erets-Yisra'el: meḡqarim bə-ʾorḡot-ḡayyim bə-Erets-Yisra'el bi-tqufat ha-Mishnah vəha-Talmud* (Jerusalem: Mossad Bialid and Isaac ben Zvi, 1984).
- Lien, C. 'Mín 閩 dialects' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, III.19–27.
- Lindsay, W. M. 'Note on the use of glossaries for the Dictionary of Medieval Latin', *Archivum Latinitatis Medii Aevi*, 1 (1924), 16–19.
- Lindemann, M. *Die französischen Wörterbücher von den Anfängen bis 1600: Entstehung und typologische Beschreibung* (Tübingen: Niemeyer, 1994).
- Liu Zhongfu 劉中富. *Ganlu zishu zilei yanjiu 干祿字書字類研究* (Jinan: Qi Lu shushe, 2004).
- Lobrichon, G. 'Glosses on the Bible, Latin' in *Encyclopaedia of the Middle Ages*.
- Louwrens, L. J. 'Contributions made by the Portuguese to the development of Bantu linguistics between 1500 and 1917', *Scientia Militaria: South African Journal of Military Studies*, 18 (1988), 29–33.
- Lowe, E. A. *Codices Latini Antiquiores: A Palaeographical Guide to Latin Manuscripts Prior to the Ninth Century*, 12 vols. (Oxford: Clarendon, 1934–71).
- Lucas, R. and D. Fergie. 'Pulcaracuranie: losing and finding a cosmic centre with the help of J. G. Reuther and others' in Peterson and Kenny (eds.), *German Ethnography in Australia*, 79–113.
- Lundbæk, K. 'The establishment of European sinology 1801–1815' in S. Clausen, R. Starrs, and A. Wedell-Wedellsborg (eds.), *Cultural Encounters: China, Japan, and the West: Essays Commemorating 25 Years of East Asian Studies at the University of Aarhus* (Aarhus University Press, 1995), 15–54.
- Lurie, D. B. 'The development of Japanese writing' in Stephen Houston (ed.), *The Shape of Script: How and Why Writing Systems Change* (Sante Fe, NM: SAR, 2012).
- Lurie, D. B. *Realms of Literacy: Early Japan and the History of Writing* (Cambridge, MA: Harvard University Asia Center, 2011).
- Maas, P. 'Review of Drachmann, *Die Überlieferung des Cyrillglossars*', *Byzantinische Zeitschrift*, 37.2 (1937), 380–1.
- Macaulay, T. B. *Prose and Poetry* (Cambridge, MA: Harvard University Press, 1970).
- Maciuszak, K. 'Persian lexicography in India', *Cracow Indological Studies*, 11 (2009), 183–212.
- MacKenzie, D. N. 'Frahāng ī pahlawīg' in *Encyclopædia Iranica*.
- MacKenzie, D. N. 'Persian lexicography' in *Encyclopaedia of Islam*, 2nd edn, IV.525–7.
- Mac Lean, J. 'The introduction of books and scientific instruments into Japan, 1712–1854', *Japanese Studies in the History of Science*, 13 (1974), 10–68.
- MacNeill, E. 'De origine Scotticae linguae', *Ériu*, 11 (1932), 112–29.
- Maeda Tomiyoshi 前田富祺. 'a i u e o to iroha no rekishi' アエイウエオとイロハの歴史. *Gengo seikatsu*, 187 (1967), 16–22.
- Maehler, H. 'Zweispachiger Aeneis-Codex' in J. Bingen and G. Nachtergaele (eds.), *Actes du XVème congrès international de papyrologie*, 4 vols. = *Papyrologica Bruxellensia*, 16–19 (Brussels: Fondation Égyptologique Reine Élisabeth, 1978–9), II.18–41.
- Mahmoodi-Bakhtiari, B. 'Āṣaf al-lughāt' in *Encyclopaedia of Islam* 3.
- Mahon, W. J. 'Contributions to the study of early Irish lexicography', unpublished PhD dissertation, Harvard University (1987).
- Mair, V. H. 'Buddhism and the rise of the written vernacular in east Asia: the making of national languages', *Journal of Asian Studies*, 53.3 (1994), 707–51.

- Mair, V. H. (trans.) 'Preface to the Seven Sounds', *Xin Tang*, 7 (1986), 134–5.
- Mair, V. H. 'What is a Chinese "dialect/topolect"? Reflections on some key Sino-English linguistic terms', *Sino-Platonic Papers*, 29 (1991), 1–31.
- Majda, T. *Rozwój języka tureckiego w XVII wieku (rękopis z 1611 r., ze zbiorów Biblioteki Uniwersyteckiej we Wrocławiu, sygn. M. 1529)* (Warsaw: Wydawnictwa Uniwersytetu Warszawskiego, 1985).
- Malandra, W. W. 'Frahang ī oīm' in *Encyclopædia Iranica*.
- Maleci, S. *Il codice Barberinianus Graecus 70 dell'Etymologicum Gudianum* (Rome: Accademia Nazionale dei Lincei, 1995).
- Malkiel, Y. 'A typological classification of dictionaries on the basis of distinctive features' in F. W. Householder and S. Saporta (eds.), *Problems in Lexicography: Report of the Conference on Lexicography Held at Indiana University, November 11–12, 1960* (Bloomington: Indiana University Research Center in Anthropology, 1962), 3–24.
- Malmgren, S.-G. 'Scandinavian languages' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 702–13.
- Malmgren, S.-G. 'Das schwedische Akademiewörterbuch und das Wörterbuch der dänischen Sprache' in Haß (ed.), *Große Lexika*, 301–13.
- Maman, A. *Comparative Semitic Philology in the Middle Ages from Saʿādiyah Gaon to Ibn Barūn (10th–12th C.)*, trans. D. Lyons (Leiden: Brill, 2004).
- Maman, A. 'The flourishing era of Jewish exegesis in Spain. The linguistic school: Judah Hayyūj, Jonah ibn Janāh, Moses ibn Chiquitilla and Judah ibn Balʿam', in M. Sæbø (ed.), *Hebrew Bible, Old Testament: The History of its Interpretation*, vol. I, part 2, *The Middle Ages* (Goettingen: Vandenhoeck & Ruprecht, 2000), 261–81.
- Maman, A. 'Lə-darko šel Rav Hai Gaon bə-millono Kitāb al-Hāwi' in S. Sharvit (ed.), *Studies in Ancient and Modern Hebrew in Honour of M. Z. Kaddari* (Ramat-Gan: Bar-Ilan University, 1999), 235–49.
- Maman, A. 'Lə-middat hissardutam šel ḥibburim bə-ḥoxmat ha-lašon: šəloša miqrei mivḥan', in M. Ben-Sasson et al. (eds.), *Uncovering the Canon: Studies in Canonicity and Genizah* (Jerusalem: Magnes, 2010), 217–33.
- Maman, A. 'Menahem ben Saruq's *Mahberet*: the first Hebrew–Hebrew dictionary', *Kernerman Dictionary News*, 13 (2005), 5–10.
- Maman, A. *Otsrot lashon: kitve ha-yad ye-kit'e ha-Genizah be-hokhmat ha-lashon me-osef Bet ha-midrash le-rabanim ba-Amerikah/Otsrot Lashon: The Hebrew Philology Manuscripts and Geniza Fragments in the Library of the Jewish Theological Seminary of America* (Jerusalem: Makhon Schocken; New York: Jewish Theological Seminary of America, 2006).
- Maman, A. 'Peshat and derash in medieval Hebrew lexicons' in A. Arazi, J. Sadan, and D. J. Wasserstein (eds.), *Compilation and Creation in Adab and Luḡa: Studies in Memory of Naphtali Kinberg (1948–1997)* = *Israel Oriental Studies*, 19 (1999), 343–57.
- Maman, A. 'Seder u-mašmaʿut bə-ʾotiyot ha-šoreš: ha-ḥeleq ha-šəviʿi min Kitāb al-Muštamil lə-Abū l-Faraj Harūn', *Peʿamim*, 89 (2002), 83–95.
- Maman, A. 'Sefer diqduq ʾavud mimey ha-benayim və-taʿalumat Ibn Yashush', *Akadem* (Bulletin of the Academy of Hebrew Language, Jerusalem), 43 (2011), 6–7.
- Maman, A. 'Sefer ha-šorašim leRabbi Səʿadiyah ibn Danan: mahadurat Milgrom Jimenez-Sanchez' 'Seʿadiyah ibn Danan's Sefer Ha-Shorashim', *Tarbiz*, 68 (1999), 287–301.
- Maman, A. 'Səridei millon Kitāb al-Hāwi ləRav Hai Gaʾon me-ʾosfei Adler və-Taylor-Schechter', *Tarbiz*, 69 (2000), 341–421.

- Mandlik, V. N. 'Editor's preface' in *Transactions of the Literary Society of Bombay*, Vol. I, reprinted edn (Mumbai: Bombay Education Society, 1877), vii–xx.
- Mannheim, B. 'Lexicography of Colonial Quechua' in Hausmann et al. (eds.), *Wörterbücher*, III.2676–84.
- Marçais, W. *Articles et conférences* (Paris: Adrien-Maisonneuve, 1961).
- Marcato, C. 'Friaulish: Grammatikographie und Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, III.637–45.
- Marckwardt, A. H. 'An unnoted source of English dialect vocabulary', *Journal of English and Germanic Philology*, 46 (1947), 177–82.
- Marks, S. 'Southern and central Africa, 1886–1910', in R. Oliver and G. N. Sanderson (eds.), *The Cambridge History of Africa*, vol. VI, 1870–1905 (Cambridge University Press, 1985), 422–92.
- Markus, M., 'Introduction to part 1: the genesis of Joseph Wright's English Dialect Dictionary', in Markus et al. (eds.), *Joseph Wright's 'English Dialect Dictionary' and Beyond*, 11–20.
- Markus, M., C. Upton, and R. Heuberger (eds.). *Joseph Wright's 'English Dialect Dictionary' and Beyond* (Frankfurt am Main: Lang, 2010).
- Marsden, W. *A Catalogue of Dictionaries, Vocabularies, Grammars, and Alphabets* (London: [for the compiler], 1796).
- Martin, W. 'New developments in Dutch lexicography from 1990 onwards' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 816–27.
- Martorelli, L. (ed.). *Greco antico nell'Occidente carolingio: frammenti di autori attici nell'Ars di Prisciano* (Hildesheim: Olms, 2014).
- Maruyama, T. 'Linguistic studies by Portuguese Jesuits in sixteenth and seventeenth century Japan' in Zwartjes and Hovdhaugen (eds.), *Missionary Linguistics*, 141–60.
- Masini, F. 'Chinese dictionaries prepared by Western missionaries in the seventeenth and eighteenth centuries' in X. Wu (ed.), *Encounters and Dialogues: Changing Perspectives on Chinese-Western Exchanges from the Sixteenth to Eighteenth Centuries* (Nettetal: Steyler, 2005), 179–93.
- Masini, F. *The Formation of Modern Chinese Lexicon and Its Evolution Toward a National Language: The Period from 1840 to 1898* (Berkeley: Project on Linguistic Analysis, University of California, 1993).
- Masini, F. 'Notes on the first Chinese dictionary published in Europe (1670)', *Monumenta Serica*, 51 (2003), 283–308.
- Maspero, H. 'Études sur la phonétique historique de la langue Annamite: les initiales', *Bulletin de l'École Française d'Extrême-Orient*, 12.1 (1912), 1–124.
- Mathews, M. M. *The Beginnings of American English* (University of Chicago Press, 1963).
- Mathews, M. M. 'Of matters lexicographical', *American Speech*, 31 (1956), 202–8.
- Mathews, M. M. 'Of matters lexicographical', *American Speech*, 33 (1958), 50–5.
- Mathews, M. M. 'Of matters lexicographical', *American Speech*, 34 (1959), 203–7.
- Matthaios, S. 'Greek scholarship in the imperial era and late antiquity', in Montanari et al. (eds.), *Brill's Companion to Ancient Greek Scholarship*, I.184–296.
- Mavoungou, P. A. 'Lexicography of the languages of central Africa' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 953–62.
- Mavoungou, P. A. 'Lexicography of the languages of western Africa' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 964–68.

- McArthur, T. *Worlds of Reference: Lexicography, Learning, and Language from the Clay Tablet to the Computer* (Cambridge University Press, 1986).
- McCarus, E. N. 'Modern Standard Arabic' in *Encyclopedia of Arabic Language and Linguistics*, III.238–62.
- McGeachy, J. A., Jr. 'The *Glossarium Salomonis* and its relationship to the *Liber glossarum*', *Speculum*, 13.3 (1938), 309–18.
- McGregor, W. B. 'Daisy Bates' documentations of Kimberley languages', *Language & History* 55.2 (2012), 79–101.
- McGregor, W. B. 'An early Trappist grammar of Nyulnyul (Dampier Land, Western Australia)' in P. Desmet et al. (eds.), *The History of Linguistic and Grammatical Praxis: Proceedings of the XIth International Colloquium of the Studienkreis 'Geschichte der Sprachwissenschaft'* (Leuven, 2nd–4th July, 1998) (Leuven and Paris: Peeters, 2000), 445–64.
- McGregor, W. B. (ed.). *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics* (Canberra: Pacific Linguistics, 2008).
- McGregor, W. B. 'Frs Herman Nekes and Ernest Worms' dictionary of Australian languages, Part III of Australian languages (1953)' in I. Mushin (ed.), *Proceedings of the 2004 Conference of the Australian Linguistic Society* (2005), ses.library.usyd.edu.au/handle/2123/93.
- McGregor, W. B. *Gunin/Kwini* (Munich and Newcastle: Lincom Europa, 1993).
- McGregor, W. B. 'Kukatja ethno-physiology and medicine' (review of Peile, *Body and Soul*), *Anthropos*, 94 (1999), 224–8.
- McGregor, W. B. (ed.). *Studies in Kimberley Languages in Honour of Howard Coate* (Munich and Newcastle: Lincom Europa, 1996).
- McHale, S. F. *Print and Power: Confucianism, Communism, and Buddhism in the Making of Modern Vietnam* (New Delhi: Munshiram Manoharlal, 2010; 2003).
- McKitterick, R. 'Glossaries and other innovations in Carolingian book production' in E. Kwakkel, R. McKitterick, and R. Thomson, *Turning Over a New Leaf: Change and Development in the Medieval Manuscript* (Leiden University Press, 2012), 20–76, 169–93.
- McLelland, N. J. G. *Schottelius's Ausführliche Arbeit von der Teutschen HauptSprache (1663) and its Place in Early Modern European Vernacular Language Study* (Chichester: Wiley-Blackwell, 2011).
- Meissner, B. *Babylonien und Assyrien*, 2 vols. (Heidelberg: Winter, 1920–5).
- Mencken, H. L. *The American Language* (New York: Knopf, 1936).
- Menz, A. 'Idioms and dialogues in Holdermann's *Grammaire Turque* (1730)' in É. Á. Csató, A. Menz and F. Turan (eds.), *Spoken Ottoman in Mediator Texts* (Wiesbaden: Harrassowitz, 2016), 147–59.
- Merkin, R. 'Al milon ivri shimushi bikhlah veal 'Milon hahove' bifrat', *Leshonenu Laam*, 46 (1995), 143–65.
- Merkin, R. 'Al milon ivri moderni bikhlah veal Milon Ariel Hamakif bifrat', *Balshanut Ivrit*, 61 (2008), 114–25.
- Merkin, R. 'Lexicography: modern period' in *Encyclopedia of Hebrew Language and Linguistics*, II.521–8.
- Merkin, R. 'Prakim betoldot hamilonut haivrit hakhadasha: Maamar sheni', *Leshonenu*, 54 (1991), 311–23.
- Merrilees, B. 'Deux versions de l'*Aalma* et leur importance dans la tradition lexicographique médiévale' in D. Trotter (ed.), *Actes du XXIVe Congrès International de*

- Linguistique et de Philologie Romanes*, Aberystwyth 2004 (Tübingen: Niemeyer, 2007), II.239–45.
- Messner, D. 'Ist das *Dictionarium latino lusitanicum, ac iaponicum* ein Wörterbuch der portugiesischen Sprache? Zur Rezeption Calepinos in Portugal', *Lusorama*, 38 (1999), 48–52.
- Messner, D. 'O primeiro dicionário bilingue português que utiliza uma língua estrangeira moderna', *Verba Hispanica*, 5 (1995), 57–65.
- Meuwese, M. 'Language, literacy, and education: native peoples and Dutch Protestant missions in southwestern Taiwan and northeastern Brazil, 1624, 1662', in Häberlein and Keese (eds.), *Sprachgrenzen, Sprachkontakte, Kulturelle Vermittler*, 107–30.
- Michalowski, P. 'The life and death of the Sumerian language in comparative perspective', *Acta Sumerologica*, 22 (2005), 177–202.
- Mierau, E. 'Concerning Yavapai–Apache bilingualism', *International Journal of American Linguistics*, 29.1 (1963), 1–3.
- Mikhailov, N. *Frühslawenische Sprachdenkmäler: die handschriftliche Periode der slowenischen Sprache (XIV. Jh. bis 1550)* (Amsterdam and Atlanta: Rodopi, 1998).
- Miller, R. A. 'Traditional linguistics and Western linguistics in Japan' in Aurox et al. (eds.), *History of the Language Sciences*, I.108–12.
- Mills, J. 'The *Vocabularium Cornicum*: a Cornish vocabulary?', *Zeitschrift für Celtische Philologie*, 60.1 (2013), 141–9.
- Minault, G. 'Sayyid Ahmad Dehlavi and the "Delhi Renaissance"' in R. E. Frykenberg (ed.), *Delhi Through the Ages: Essays in Urban History, Culture, and Society* (Delhi: Oxford University Press, 1986), 174–85.
- Miron, P. 'Rümanische Lexikographie' in Hausmann et al. (eds.), *Wörterbücher*, II.1880–91.
- Mirsky, H. *Torat ha-lašon šel Menaḥem ben Saruq* (Jerusalem: Ben-Zvi Institute, in press).
- Mishor, M. 'Milon ha-Hoveh and Milon sapir', *Kernerman Dictionary News*, 12 (2004), 23–6.
- Misra, B. G. (ed.). *Lexicography in India: Proceedings of the First National Conference on Dictionary Making in Indian Languages*, Mysore, 1970 (Mysore: Central Institute of Indian Languages, 1980).
- Mitchell, L. *Language, Emotion, and Politics in South India: The Making of a Mother Tongue* (Bloomington: Indiana University Press, 2009).
- Mitchell, L. 'Parallel languages, parallel cultures', *Indian Economic & Social History Review*, 42.4 (2005), 445–67.
- Mittag, A. 'Becoming acquainted with nature from the *Odes*' in H. U. Vogel and G. Dux (eds.), *Concepts of Nature: A Chinese–European Cross-Cultural Perspective* (Leiden: Brill, 2010).
- Miyoshi, K. *The First Century of English Monolingual Lexicography* (Newcastle, UK: Cambridge Scholars, 2017).
- Monaghan, P. 'Laying down the country: Norman B. Tindale and the linguistic construction of the north-west of South Australia', unpublished PhD thesis, University of Adelaide (2003).
- Monaghan, P. 'Norman B. Tindale and the Pitjantjatjara language' in McGregor (ed.), *Encountering Aboriginal Languages*, 251–72.
- Montana, F. 'Hellenistic scholarship' in Montanari et al. (eds.), *Brill's Companion to Ancient Greek Scholarship*, I.60–183.

- Montanari, F., S. Matthaios, and A. Rengakos (eds.). *Brill's Companion to Ancient Greek Scholarship*, 2 vols. (Leiden and Boston: Brill, 2015).
- Montgomery, M. B. 'The core or the periphery? The lack of a *Dictionary of Irish English on Historical Principles*', in M. Adams and A. Curzan (eds.), *Contours of English and English Language Studies* (Ann Arbor: University of Michigan Press, 2001), 213–36.
- Montgomery, M. B., and J. H. Heinmiller. 'Dictionary of Smoky Mountain and Southern Appalachian English', *Dictionaries*, 38.1 (2017), 68–81.
- Monzón, C. 'The Tarascan lexicographical tradition in the sixteenth century' in Zwartjes et al. (eds.), *Missionary Linguistics IV*, 165–95.
- Mookerji, D. N. *A Monograph on Paper and Papier-Mache in Bengal* (Kolkata: Bengal Secretariat Book Depot, 1908).
- Moore, D. C. 'A history of the 1842 Descriptive Vocabulary' in McGregor (ed.), *Encountering Aboriginal Languages*, 301–20.
- Moore, D. C. 'T. G. H. Strehlow and the linguistic landscape of Australia 1930–1960' in McGregor (ed.), *Encountering Aboriginal Languages*, 273–300.
- Morag, S. 'The emergence of Modern Hebrew: some sociolinguistic perspectives' in Glinert (ed.), *Hebrew in Ashkenaz*, 208–21.
- Moran, P. 'Greek in early medieval Ireland' in A. Mullen and P. James (eds.), *Multilingualism in the Graeco-Roman Worlds* (Cambridge University Press, 2012), 172–92.
- Moran, W. L. *The Amarna Letters* (Baltimore: Johns Hopkins University Press, 1992).
- Moreno Fernández, F. 'Antonio de Nebrija y la lexicografía americana del siglo XVI: a propósito del *Lexicon* de Fray Domingo de Santo Tomás', *Voz y letra*, 5.1 (1994), 79–104.
- Morey, S. 'Ahom and Tangsa: case studies of language maintenance and loss in North East India' in H. C. Cardoso (ed.), *Language Endangerment and Preservation in South Asia* (Manoa: University of Hawai'i Press, 2014), 46–77.
- Moropa, K., and A. Kruger, 'Mistranslation of culture-specific terms in Kropf's *Kafir-English Dictionary*', *South African Journal of African Languages*, 20 (2000), 70–9.
- Morton, H. C. *The Story of Webster's Third: Philip Gove's Controversial Dictionary and Its Critics* (Cambridge University Press, 1994).
- Moss, A. *Renaissance Truth and the Latin Language Turn* (Oxford University Press, 2003).
- Mote, F. W. *Imperial China 900–1800*, paperback edn (Cambridge, MA: Harvard University Press, 2003; 1999).
- Moussâkhâni, Z. 'Le dictionnaire encyclopédique persan de Mohammad Mo'in', *La revue de Téhéran*, no. 95 (2013), www.teheran.ir/.
- Mtuzze, P. T. 'A critical survey of Xhosa lexicography 1772–1989', *Lexikos*, 2 (1992), 165–77.
- Mugglestone, L. (ed.). *Lexicography and the OED: Pioneers in the Untrodden Forest* (Oxford University Press, 2000).
- Mugglestone, L. *Lost for Words: The Hidden History of the Oxford English Dictionary* (New Haven and London: Yale University Press, 2005).
- 'Muhammed Hevai Uskufi: Bosansko-turski rječnik iz 1631 godine'. 12 August 2013, ljubusaci.com/2013/08/12/muhammed-hevai-uskufi-bosansko-turski-rjecnik-iz-1631-godine/.
- Müller, M. *Akkadisch in Keilschrifttexten aus Ägypten: deskriptive Grammatik einer Interlanguage des späten zweiten vorchristlichen Jahrtausends anhand der Ramses-Briefe* = *Alter Orient und Altes Testament*, 373 (Münster: Ugarit-Verlag, 2010).
- Müller, P. O. *Deutsche Lexikographie des 16. Jahrhunderts* (Tübingen: Niemeyer, 2001).

- Müller-Yokota, W. 'Abriss der geschichtlichen Entwicklung der Schrift in Japan', *Bochumer Jahrbuch zur Ostasienforschung*, 10 (1987), 1–75.
- Munske, H. H. (ed., with N. Århammar) *Handbook of Frisian Studies/Handbuch des Friesischen* (Tübingen: Niemeyer, 2001).
- Munzavi, A. N. *Farhangnāmah-hā-yi 'Arabī dar Fārsī* (Tehran University Press, 1959).
- Murgia, C. E. *Prolegomena to Servius 5: The Manuscripts* (Berkeley and Los Angeles: University of California Press, 1975).
- Murray, J. A. H. *The Evolution of English Lexicography* (Oxford: Clarendon Press, 1900).
- Murray, K. M. E. *Caught in the Web of Words: James A. H. Murray and the Oxford English Dictionary* (New Haven and London: Yale University Press, 1977).
- Muru, C. 'Shaping minds and cultures: the impact of missionary translations in Southern India', in O. Zwartjes, K. Zimmermann, and M. Schrader-Kniffki (eds.), *Missionary Linguistics V: Translation Theories and Practices* (Amsterdam and Philadelphia: Benjamins, 2014), 203–30.
- Mūsā, 'A. Ḥ. 'Dirāsa tiqniyya muqārana li-ma'āğim al-Şaḥāḥ wa-Lisān al-'Arab wa-Tağ al-'arūs', *Mağallat al-Mu'ğamiyya*, 5/6 (1989–90), 147–58.
- Mūsā, 'A. Ḥ. and 'A. Şāhīn. *Dirāsa iḥsā'iyya li-ğudūr Mu'ğam Tāğ al-'arūs bi-stiḥdām al-kompyūtir* (Kuwait: Ġāmi'at al-Kuwayt, 1973).
- Muttusami Pillai, A. *Brief Sketch of the Life and Writing of Father C. J. Beschi or Vira-Mamuni, translated from the Original Tamil* (Chennai: Pharos, 1840).
- Naarden, B. 'Witsen's studies of Inner Eurasia' in Huigen et al. (eds.), *The Dutch Trading Companies as Knowledge Networks*, 211–39.
- Nagar, A. S. 'A brief survey of dictionary making in Gujarat' in Misra (ed.), *Lexicography in India*, 21–8.
- Nagle, T. C. "'There is much, very much, in the name of a book", or, the famous title of *Hobson-Jobson* and how it got that way', in M. Adams (ed.), *'Cunning Passages, Contrived Corridors': Unexpected Essays in the History of Lexicography* (Monza: Polimetrica, 2010), 111–28.
- Nagle, T. C. 'The visible and invisible influence of Yule's *Hobson-Jobson* on Murray's *Oxford English Dictionary*', *International Journal of Lexicography*, 27 (2014), 280–308.
- Nair, S., and A. Kulkarni. 'The knowledge structure in *Amarakośa*', in G. N. Jha (ed.), *Sanskrit Computational Linguistics* (Berlin and Heidelberg: Springer, 2010), 173–89.
- Naoumides, M. 'The fragments of Greek lexicography in the papyri', in *Classical Studies Presented to Ben Edwin Perry* = Illinois Studies in Language and Literature, 58 (Urbana: University of Illinois Press, 1969), 181–202.
- Naqavī, S. *Farhangnivīsi-i Fārsī dar Hind va Pākistān* (Tehran: Idarah-i Kull-i Nigarish-i Vizarat-i Farhang, 1962).
- Naregal, V. *Language Politics, Elites, and the Public Sphere* (New Delhi: Permanent Black, 2001).
- Naşşār, Ḥ. *al-Mu'ğam al-'Arabī: Naş'atuhu wa-taṭawwuruḥu*, 2nd edn, 2 vols. (Cairo: Dār Mişr li-l-Ṭibā'a, 1968).
- Nationalencyclopedia. 'Uppslagsverket', www.ne.se/info/tjanster/uppslagsverket.
- Neeman, Y., and R. Finkel. 'Rav-Milim online', *Kernerman Dictionary News*, 12 (2004), 28–31.
- Netzer, A. 'Haim, Solaymān' in *Encyclopædia Iranica*.

- Netzer, N. 'Ləşon ḥaxamim bə-kitvei ha-mdaqdəqim ha-ivriyyim bimeh ha-benayim bi-tqufat ha-ytsira ha-məqorit: Səʿadia-Ibn Balʿam', unpublished PhD thesis, Bar-Ilan University (1983).
- Neubauer, A. 'Notice sur la lexicographie hebraique avec des remarques sur quelques grammairiens posterieurs a Ibn Djana'h', *Journal asiatique*, 5th ser., 18 (1861), 441–76.
- Newman, P. 'A century and a half of Hausa language studies', in A. Rufa'i (ed.), *Nigerian Languages: Yesterday, Today, and Tomorrow* (Kano: Centre for the Study of Nigerian Languages, Bayero University, 1991), 1–18.
- Newman, R. M., and P. Newman. 'The Hausa lexicographic tradition', *Lexikos*, 11 (2001), 263–86.
- Nguyễn Đ. H. 'Vietnamese lexicography', paper presented at the 8th annual meeting of the International Association of Applied Linguistics, Sydney, 16–21 August 1987, archive .org/details/ERIC_ED302082.
- Nguyễn Đ. H. 'Vietnamese phonology and graphemic borrowings from Chinese: The Book of 3,000 Characters revisited', *Mon-Khmer Studies*, 20 (1992), 163–82.
- Nguyễn T. C. 'The Sanzijing (三字經) and primary education in premodern Vietnam', *International Journal of Chinese Character Studies (Shijie hanzi tongbao 世界漢字通報)*, 1.2 (2015), 33–54.
- Nickau, K. 'Altes und Neues zur Überlieferung der Synagoge des Symeon', *Göttinger Forum für Altertumswissenschaft*, 4 (2001), 221–3.
- Nickau, K. 'Neues zur Überlieferung der Synagoge des Symeon (Paris. gr. 2552, Paris. Suppl. gr. 1238, Vatic. gr. 2226)', *Göttinger Forum für Altertumswissenschaft*, 3 (2000), 13–31.
- Nickau, K. 'Schiffbruch in der Wüste des Sinai: zu Herennios Philon, Neilos von Ankyra, und dem Ammonioslexikon', *Hermes*, 128 (2000), 218–26.
- Nie Hongyin 聂鸿音 and Sun Bojun 孙伯君. 'Xifan yiyu' jiaolu ji huibian 《西番译语》校录及汇编. (Beijing: Shehui Kexue Wenxian Chubanshe, 2010).
- Niederehe, H. J. *Bibliografía cronológica de la lingüística, la gramática y la lexicografía del español (BICRES): Desde los comienzos hasta el año 1600* (Amsterdam and Philadelphia: Benjamins, 1994).
- Niederehe, H. J. 'Grammatiken und Wörterbücher des Quechua: Einige bibliographische Anmerkungen', in B. Staib (ed.), *Linguistica romanica et indiana: Festschrift für Wolf Dietrich zum 60. Geburtstag* (Tübingen: Narr, 2000), 533–50.
- Niederehe, H. J. 'Das "Universal vocabulario" des Alfonso Fernández de Palencia (1490) und seine Quelle', *Historiographia Linguistica*, 11 (1984), 39–54.
- Ning Jifu 甯忌浮. *Hanyu yunshu shi (Mingdai juan) 漢語韻書史：明代卷* (Shanghai: Shanghai Renmin chubanshe, 2009).
- Ning Jifu 宁继福. *Zhongyuan yinyun biaoqao 中原音韵表稿* (Changchun: Jilin wenshi chubanshe, 1985).
- Nishikawa Hideo. 'Denshi jisho no rekishi: Kokorozashi o takaku age, yume to roman o motte' in Jiten kyōkai (ed.), *Nihon no jisho no ayumi*, 167–85.
- Nkomo, D., and Z. Wababa. 'IsiXhosa lexicography: past, present, and future', *Lexikos*, 23 (2013), 348–70.
- No Yongpil 盧鏞弼. 'Hanguk godae munjahak gwa hun-gohak ui baldal' 韓國古代文字學과訓詁學의發達, *Jindan hakbo*, 110 (2010), 1–30.
- Norman, J. *Chinese* (Cambridge University Press, 1988).

- Norman, J. 'Dialect classification' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, II.42–9.
- Norman, K. R. *Pāli Literature: Including the Canonical Literature in Prakrit and Sanskrit of all the Hīnayāna Schools of Buddhism* (Wiesbaden: Harrassowitz, 1983).
- Norris, H. T. 'Edmund Castell (1606–1686) and his *Lexicon Heptaglotton* (1669)' in G. A. Russell (ed.), *The 'Arabick' Interest of the Natural Philosophers in Seventeenth-Century England* (Leiden etc.: Brill, 1994), 70–87.
- Nosco, P. *Remembering Paradise: Nativism and Nostalgia in Eighteenth-Century Japan* (Cambridge, MA: Council on East Asian Studies, Harvard University, 1990).
- Oates, L. F. 'Aboriginal recording of Aboriginal language' in P. Austin, R. M. W. Dixon, T. Dutton, and I. White (eds.), *Language and History: Essays in Honour of Luise Hercus* (Canberra: Pacific Linguistics, 1990), 221–32.
- Oates, L. F. *Against the Wind: Wycliffe Bible Translators Australia in Action* (Melbourne: Van Brummelen and Wycliffe Bible Translators Australia, 2003).
- Oates, L. F. 'The Summer Institute of Linguistics and Aboriginal-Islander research', *Aboriginal History*, 23 (1999), 43–50.
- Ó Cróinín, D. 'The earliest Old Irish glosses' in Bergmann et al. (eds.), *Mittelalterliche volkssprachige Glossen*, 7–31.
- Offord, D., G. Argent, and V. Rjéoutski. 'French and Russian in Catherine's Russia' in D. Offord et al. (eds.), *French and Russian in Imperial Russia*, 2 vols. (Edinburgh University Press, 2015), I.25–44.
- Ogilvie, S. *Words of the World: A Global History of the 'Oxford English Dictionary'* (Cambridge University Press, 2013).
- O'Grady, G. N. 'Lexicographic research in Aboriginal Australia' in Sebeok (ed.), *Current Trends in Linguistics*, vol. VIII, *Linguistics in Oceania*, 779–803.
- O'Grady, G. N., C. F. Voegelin, and F. M. Voegelin. 'Languages of the world: Indo-Pacific, fascicle 6', *Anthropological Linguistics* 8.2 (1966), 1–197.
- Ogura Shinpei 小倉進平. *Zōtei hochū Chōsen gogakushi* 増訂補注朝鮮語学史, reprinted 2nd edn, ed. Kōno Robuō 河野六郎 (Tokyo: Tōkō Shoin, 1964; ¹1920, ²1940).
- Okimori Takuya (ed.). *Zusetsu Nihon no jisho* (Tokyo: Ōfū, 2008).
- Okimori Takuya, Kurashima Tokihisa, Katō Tomomi, and Makino Takenori (eds.). *Nihon jisho jiten* (Tokyo: Ōfū, 1996).
- Okumuş, S. 'Anadolu sahasında manzum sözlükler ve Hevâ'nin *Makbûl-i Ârif*'ine dair', *Zbornik Radova*, 32.9 (Tuzla: Univerzitet u Tuzli, 2008), 147–62.
- Okumuş, S. 'Muhammed Hevâi Üsküfi ve Türkçe-Boşnakça manzum sözlüğü *Makbûl-i Ârif* (Potur Şâhidî)', *Turkish Studies*, 4.4 (2009), 823–44.
- Olivieri, O. 'I primi vocabolari Italiani fino alla prima edizione Della Crusca', *Studi di filologia Italiana*, 6 (1942), 64–192.
- Ölmez, M. 'Tarihi Türk dillerinin sözlükleri', *Kebikeç*, 6 (1998), 109–10.
- Ölmez, Z. 'Çağatayca sözlükler', *Türk Dilleri Araştırmaları*, 13 (2003), 11–21.
- Olsson, B. 'Historical and sociocultural preconditions of language in Scandinavia from the sixteenth to the end of the eighteenth century' in Bandle et al., *Nordic Languages*, II.1238–43.
- Omer, A. M. 'Early Arabic lexicons of homophonic words' in K. Dévényi, T. Iványi and A. Shvitiel (eds.), *Proceedings of the Colloquium on Arabic Lexicology and Lexicography* = Budapest Studies in Arabic, 6–7 (1993), 3–II.

- Opler, M. E. (ed.). *Grenville Goodwin Among the Western Apache: Letters from the Field* (Tucson: University of Arizona Press, 1973).
- Ornan, U. 'Matay nolda haivrit hakhadasha' in C. E. Cohen (ed.), *Matayim Vakhamishim Shmot Ivrit Khadasha* (Jerusalem: Academy of the Hebrew Language, 2009), 3–16.
- Orsini, F. 'What did they mean by "public"? Language, literature and the politics of nationalism', *Economic and Political Weekly*, 34.7 (1999), 409–16.
- Osselton, N. E. 'Bilingual lexicography with Dutch' in Hausmann et al. (eds.), *Wörterbücher*, III.3034–9.
- Osselton, N. E. 'Murray and his European counterparts', in Muggleston (ed.), *Lexicography and the OED*, 59–76.
- Osterkamp, S. 'Sinoform writing' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, IV.115–24.
- Ottósson, K. 'The data basis of a Nordic language history' in Bandle et al. (eds.), *Nordic Languages*, 1.38–48.
- Ottósson, K. 'Research in Icelandic language history 1850–1950: an overview' in Bandle et al. (eds.), *Nordic Languages*, 1.95–101.
- Pablo Núñez, L. *El arte de las palabras: Diccionarios e imprenta en el siglo d'oro*, 2 vols. (Mérida: Editora regional de Extremadura, 2010).
- Pade, M. 'Perotti's *Cornu copiae*' in P. Ford, J. Bloemendal, and C. Fantazzi (eds.), *Brill's Encyclopaedia of the Neo-Latin World* (Leiden: Brill, 2014), II.1126–8.
- Pagani, L. 'Language correctness (Hellenismos) and its criteria' in Montanari et al. (eds.), *Brill's Companion to Ancient Greek Scholarship*, II.798–849.
- Page, J. *The Black Bishop: Samuel Adjai Crowther* (New York and Chicago: Revell, 1909).
- Pak Hyong-ik. 'Sim uilin-ui Botonghaggyo Joseon-eo Sajeon (1925)ui bunseog', *Journal of Korealex*, 2 (2003), III–42.
- Pallath, J. J. (ed.). *Arnos Padiri (Joannes Ernestus Hanxleden, SJ, 1681–1732): The First Malayalam Poet, Scholar, Orientalist* (Kozhikode: Arnos Padiri, 1994).
- Palm, J. D. 'An account of the Dutch Church in Ceylon', *Journal of the Ceylon Branch of the Royal Asiatic Society*, 1.[3] (1847), 5–66.
- Pan Chonggui 潘重規. *Yingya Dunhuang yunji bielu* 瀛涯敦煌韻輯別錄 (Lanzhou: Gansu wenhua chubanshe, 1999).
- Pān Yùkūn 潘玉坤. "'Shizhòu piān" niándài kǎo' 史籀篇年代考, *Hángzhōu shifān xuéyuàn xuébào* 杭州師範學院學報 (2002), part 2, 82–5.
- Park Hyeongik. *Hanguk jajeon ui yeongu* (Seoul: Yeongnak, 2012).
- Pastreich, E. *The Observable Mundane: Vernacular Chinese and the Emergence of a Literary Discourse on Popular Narrative in Edo Japan* (Seoul University Press, 2011).
- Paranavitana, K. D. *The Dutch Sinhala Dictionary by Simon Cat* (Colombo: Department of National Archives, 2012).
- Patkar, M. M. *History of Sanskrit Lexicography* (New Delhi: Munshiram Manoharlal, 1981).
- Paul, L. 'Vullers, Johann August' in *Encyclopædia Iranica. Paulys Realencyclopädie der classischen Altertumswissenschaft*, ed. G. Wissowa et al., 51 vols. [numbered 1–24, 1A–10A, and Supplement 1–15] in 84 parts (Stuttgart: Metzler, 1894–1980).
- Pei, M. 'Ain't is in, raviolis ain't', *New York Times* (book review section), 22 October 1961, 6.
- Peiris, E. 'Sinhalese Christian literature of the XVIIth and XVIIIth centuries', *Journal of the Dutch Burgher Union*, 53.3/4 (1963), 16–35.

- Pellicer, D. 'Confesión y conversación en la *Doctrina y enseñanza de la lengua mazahua* de Diego de Nágera Yanguas', *Estudios de cultura otopame*, 5 (2006), 13–52.
- Penhallurick, R. 'Dialect dictionaries' in Cowie (ed.), *Oxford History of English Lexicography*, II.290–313.
- Pensado, J. L. 'Galician lexicography' in Hausmann et al. (eds.), *Wörterbücher*, II.1736–8.
- Perera, S. G. 'Some ancient grammars and dictionaries of the Sinhalese language', *Ceylon Literary Register*, 4.7 (1936), 327–9.
- Perez, M. 'Muva^ot mi-tox *Kitāb al-'Istighnā* šel Rabbi Šömu^el Hanagid bə-tox peruš ^oanónimi lə-sefer Təhillim', *Annual for Biblical and Ancient Near Eastern Studies*, 12 (2000), 241–87.
- Perry, J. R. 'Dictionaries: ii. Arabic/Persian' in *Encyclopædia Iranica*.
- Perry, J. R. 'Early Arabic–Persian lexicography: the *asāmī* and *mašādir* genres' in K. Dévényi et al. (eds.), *Proceedings of the Colloquium on Arabic Lexicology and Lexicography* (C.A.L.L.), part 1 = The Arabist, 6 (Budapest: Eötvös Loránd University and Csoma do Kőrös Society, 1993), 247–59.
- Perry, J. R. 'Modern Iranian lexicography' in Hausmann et al. (eds.), *Wörterbücher*, III.2477–80.
- Peters, M. *De wijze koopman: het wereldwijde onderzoek van Nicolaes Witsen (1641–1717), burgemeester en VOC-bewindhebber van Amsterdam* (Amsterdam: Bakker, 2010).
- Peterson, J. 'The personal name lists in the scribal curriculum of Old Babylonian Nippur: An overview', *Zeitschrift für Assyriologie und Vorderasiatische Archäologie* 101.2 (2011), 246–73.
- Peterson, N., and A. Kenny (eds.). *German Ethnography in Australia* (Canberra: ANU Press, 2017).
- Petrover, L. 'Darko šel rabbi David Qimḥi bə-millono ha-yaduwa^e bə-šem Sefer ha-šorašim', unpublished PhD thesis, Hebrew University of Jerusalem, 2016.
- Pétursson, M. 'Inselnordische Lexikographie' in Hausmann et al. (eds.), *Wörterbücher*, II.1928–32.
- Pfeiffer, R. *History of Classical Scholarship: From the Beginnings to the End of the Hellenistic Age* (Oxford: Clarendon, 1968).
- Pfister, M. 'Die italienische Lexikographie von den Anfängen bis 1900' in Hausmann et al. (eds.), *Wörterbücher*, II.1844–63.
- Phan, J. D. 'Chữ Nôm and the taming of the South: a bilingual defense for vernacular writing in the *Chi nam ngọc âm giải nghĩa*', *Journal of Vietnamese Studies* 8.1 (2013), 1–33.
- Phan, J. D. 'Rebooting the vernacular in seventeenth-century Vietnam' in B. A. Elman (ed.), *Rethinking East Asian Languages, Vernaculars, and Literacies, 1000–1919* (Leiden: Brill, 2014), 96–128.
- Phan, J. D. 'The secularization of the sinograph in Vietnam and its demotion from the cosmological to the aesthetic', *Journal of World Literature*, 1 (2016), 275–93.
- Phan, P. C. *Mission and Catechesis: Alexandre de Rhodes and Inculturation in Seventeenth-Century Vietnam* (Maryknoll, NY: Orbis, 1998).
- Pieroni, P. Marcus Verrius Flaccus' 'De significatu verborum' in den Auszügen von Sextus Pompeius Festus und Paulus Diaconus. Einleitung und Teilkommentar (154,19 – 186,29 Lindsay) (Frankfurt am Main: Lang, 2004).
- Pinch, W. 'Same difference in India and Europe', *History and Theory*, 38.3 (1999), 389–407.

- Piotrowski, T. 'The development of the general monolingual dictionary in Poland', in Hartmann (ed.), *History of Lexicography*, 185–95.
- Plaut, W. G. 'A Hebrew–Dakota dictionary', *Publications of the American Jewish Historical Society*, 42 (1953), 361–70.
- Plofker, K. *Mathematics in India* (Princeton University Press, 2009).
- Polański, K., and J. A. Schnert. *Polabian–English Dictionary* (The Hague and Paris: Mouton, 1967).
- Polomé, E. C. 'Lexicography of the Niger-Kordofanian languages' in Hausmann et al. (eds.), *Wörterbücher*, III.2646–9.
- Pontani, F. 'Scholarship in the Byzantine empire (529–1453)' in Montanari et al. (eds.), *Brill's Companion to Ancient Greek Scholarship*, I.297–455.
- Poppe, N. 'Eine viersprachige Zamaxšari-Handschrift', *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 101 (1951), 301–32.
- Pořízka, V. 'Hindī–Russian and Urdū–Russian dictionaries', *Archiv Orientální*, 46.1 (1978), 185–92.
- Porter, J. C. *Paper Medicine Man: John Gregory Bourke and His American West* (Norman and London: University of Oklahoma Press, 1986).
- Posener, G. 'Sur l'emploi de l'encre rouge dans les manuscrits Égyptiens', *Journal of Egyptian Archaeology*, 37 (1951), 75–80.
- Power, S. 'Aaron's rod to zurr: digitizing the card collection of the Dictionary of Newfoundland English', *Dictionaries*, 36 (2015), 150–67.
- Powitz, G. 'Le Catholicon – Esquisse de son histoire' in Hamesse (ed.), *Les manuscrits des lexiques et glossaires*, 299–336.
- Practical Orthography of African Languages* (London: International Institute of African Languages and Cultures, 1927; revised edn 1930).
- Prestina-Shapiro, L. *Jazyk zaprashchennogo jazyka* (Minsk: MET, 2005).
- Pritchett, F. W. *A Desertful of Roses: The Urdu Ghazals of Mirza Asadullah Khan 'Ghalib'*, www.columbia.edu/itc/mealc/pritchett/ooghalib/.
- Probert, P. *Ancient Greek Accentuation: Synchronic Patterns, Frequency Effects, and Prehistory* (Oxford University Press, 2006).
- Proust, J. 'De quelques dictionnaires hollandais ayant servi de relais à l'encyclopédisme européen vers le Japon', *Dix-huitième siècle*, 38 (2006), 17–38.
- Ptashnyk, S. 'Das Wörterbuch der ukrainischen Sprache von Borys Hrinčenko (1907–1909)' in Haß (ed.), *Große Lexika*, 351–66.
- Pytlowany, A., and T. Van Hal. 'Merchants, scholars, and languages: the circulation of linguistic knowledge in the context of the Dutch East India Company (VOC)', *Histoire Épistémologie Langage*, 38.1 (2016), 19–38.
- Quack, J. F. 'Die Hieratischen und Hieroglyphischen Papyri aus Tebtynis: Ein Überblick', in K. Ryholt (ed.), *Hieratic Texts from the Collection = The Carlsberg Papyri 7; Carsten Niebuhr Institute Publications 30* (Copenhagen: Museum Tusculanum, 2006), 1–7.
- Quack, J. F. 'How the Coptic script came about' in E. Grossman et al. (eds.), *Greek Influence on Egyptian-Coptic: Contact-Induced Change in an Ancient African Language = Lingua Aegyptia Studia Monographica*, 17 (Hamburg: Widmaier, 2017), 27–96.
- Quack, J. F. 'Inhomogenität von Ägyptischer Sprache und Schrift in Texten aus dem späten Ägypten' in K. Lembke, M. Minas-Nerpel, and S. Pfeiffer (eds.), *Tradition and*

- Transformation: Egypt under Roman Rule* = Culture and History of the Ancient Near East, 41 (Leiden: Brill, 2010), 313–41.
- Quack, J. F. 'Die spätägyptische Alphabetreihenfolge und das "südsemitische" Alphabet', *Lingua Aegyptia*, 11 (2003), 163–84.
- Quecke, H. 'Eine griechisch-ägyptische Wortliste vermutlich des 3. Jh. v. Chr. (P. Heid. Inv.-Nr. H 414)', *Zeitschrift für Papyrologie und Epigraphik*, 116 (1997), 67–80.
- Quemada, B. *Les dictionnaires du français moderne, 1539–1863: Etude sur leur histoire, leurs types et leurs méthodes* (Paris: Didier, 1967).
- Quemada, B. 'Französisch: Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, V.1.869–94.
- Qiú Xīguī 裘錫圭. *Wénzixué gǎiyào 文字學概要* (Beijing: Shāngwù yīnshūguǎn, 1990).
- Qiu Xigui. *Chinese Writing*, trans. G. L. Mattos and J. Norman (Berkeley: Institute of East Asian Studies, University of California, 2000).
- Quirke, S., and C. A. Andrews. *The Rosetta Stone: Facsimile Drawing with an Introduction and Translations* (London: British Museum, 1988).
- Rabault-Feuerhahn, P. *Archives of Origins: Sanskrit, Philology, Anthropology in 19th Century Germany*, trans. D. Bach and R. Willet (Wiesbaden: Harrassowitz, 2013).
- Radici Colace, P. 'Polluce nell'Onomasticon' in C. Mauduit (ed.), *L'Onomasticon de Pollux: aspects culturels, rhétoriques, et lexicographiques* (Lyon: CEROR, 2013), 25–34.
- Radloff, W. *Das türkische Sprachmaterial des Codex Comanicus* (St Petersburg: Académie Impériale des sciences, 1887).
- Raghavan, V., et al. *New Catalogus Catalogorum: An Alphabetical Register of Sanskrit and Allied Works and Authors*, 39 vols. (Chennai: University of Madras, 1949–2007; revised edn of vol. I, 1968).
- Rahman, T. 'Language policy and education in Pakistan' in N. H. Hornberger (ed.), *Encyclopedia of Language and Education*, 2nd edn (Boston: Springer, 2008), I.383–92.
- Rainey, C. F. *The El-Amarna Correspondence*, 2 vols. (Leiden: Brill, 2015).
- Raj, K. *Relocating Modern Science: Circulation and the Construction of Knowledge in South Asia and Europe, 1650–1900* (Basingstoke and New York: Palgrave Macmillan, 2007).
- Rāmamūrti, G. V. *A Memorandum on Modern Telugu* (Chennai: Guardian, 1913).
- Ramaseshan, S. 'The history of paper in India up to 1948', *Indian Journal of History of Science*, 24.2 (1989), 103–21.
- Ramaswamy, S. *Passions of the Tongue: Language Devotion in Tamil India, 1891–1970* (Berkeley: University of California Press, 1997).
- Ramsey, S. R. 'The Japanese language and the making of tradition', *Japanese Language and Literature* 38 (2004), 81–110.
- Ramsey, S. R. *The Languages of China* (Princeton University Press, 1987).
- Ramson, B. *Lexical Images: The Story of the 'Australian National Dictionary'* (South Melbourne: Oxford University Press, 2002).
- Raschellà, F. D. 'Vernacular gloss writing in medieval Scandinavia' in Bergmann et al. (eds.), *Mittelalterliche volkssprachige Glossen*, 587–99.
- Read, A. W. 'Approaches to lexicography and semantics' in T. A. Sebeok (ed.), *Current Trends in Linguistics*, 10 (1973), 145–205.
- Read, A. W. 'Boucher's linguistic pastoral of colonial Maryland', *Dialect Notes*, 6 (1933), 353–60.
- Read, A. W. 'The word blizzard', *American Speech*, 3 (1928), 191–217.

- Reed, C. A. *Gutenberg in Shanghai: Chinese Print Capitalism, 1876–1937* (Vancouver: University of British Columbia Press, 2004).
- Reichmann, O. 'Geschichte lexikographischer Programme in Deutschland' in Hausmann et al. (eds.), *Wörterbücher*, I.230–46.
- Reinhardt, J. 'Iberoromance I: historical and etymological lexicography' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 636–45.
- Reitzenstein, R. 'Etymologica' in *Paulys Realencyclopädie*, VI.1.807–17.
- Reitzenstein, R. *Geschichte der griechischen Etymologica: ein Beitrag zur Geschichte der Philologie in Alexandria und Byzanz* (Leipzig: Teubner, 1897).
- Renner, T. 'Three new Homeric', *Harvard Studies in Classical Philology*, 83 (1979), 313–21.
- Rennie, S. *Jamieson's Dictionary of Scots: The Story of the First Historical Dictionary of the Scots Language* (Oxford University Press, 2012).
- Reshef, Y. 'Modern Hebrew grammar: history of scholarship' in *Encyclopedia of Hebrew Language and Linguistics*, III.656–63.
- Reverdin, O. 'Figures d'hellénisme à Genève' in *Homère chez Calvin* (Geneva: Droz, 2000), 27–101.
- Rey, A. 'La lexicographie française depuis Littré' in Hausmann et al. (eds.), *Wörterbücher*, II.1818–43.
- Reyre, D. 'Prólogo segundo: las llaves del Tesoro de Covarrubias' in Covarrubias (ed.), *Tesoro* (2006), xlv–lxvi.
- Rhodes, R. A. 'Lexicography of the languages of the North American Indians' in Hausmann et al. (eds.), *Wörterbücher*, III.2691–6.
- Rhodes, R. A. 'Multiple assertions, grammatical constructions, lexical pragmatics, and the Eastern Ojibwa-Chippewa-Ottawa Dictionary' in Frawley et al. (eds.), *Making Dictionaries*, 108–22.
- Ribadeneira, P. de. *Bibliotheca scriptorum Societatis Jesu*, ed. P. Alegambe (Antwerp: Meursius, 1643).
- Rice, K. 'Linguistic evidence regarding the Apachean migration' in Seymour (ed.), *From the Land of Ever Winter to the American Southwest*, 249–70.
- Richter, T. S. 'Greek, Coptic, and the "language of the Hijra": the rise and decline of the Coptic language in late antique and medieval Egypt', in H. Cotton et al. (eds.), *From Hellenism to Islam: Cultural and Linguistic Change in the Roman Near East* (Cambridge University Press, 2009), 401–46.
- Richter, T. S. 'Language choice in the Qurra dossier' in A. Papaconstantinou (ed.), *The Multilingual Experience in Egypt, from the Ptolemies to the Abbasids* (Farnham: Ashgate, 2010), 189–220.
- Rico, A., and J. Solà. 'Katalanisch: Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, V.2.281–310.
- Rietbergen, P. J. A. N. *Europa's India: tussen fascinatie en cultureel imperialisme, 1760–2000* (Nijmegen: Vantilt, 2007).
- Rietbergen, P. J. A. N. 'Wie verre reizen doet ... Compagniesdienaren en andere schrijvende reizigers', in Blussé and Ooms (eds.), *Kennis en compagnie*, 164–83.
- Rieu, C. *Catalogue of the Persian Manuscripts in the British Museum*, 3 vols. (London: British Museum, 1881).
- Rieu, C. *Supplement to the Catalogue of the Persian Manuscripts in the British Museum* (London: British Museum, 1895).

- Rivara, J. H. Da C. *An Historical Essay on the Konkani Language* (Bombay: Marathi Samshodhana Mandala, 1958).
- Robins, R. H. *The Byzantine Grammarians: Their Place in History* (Berlin and New York: Mouton de Gruyter, 1993).
- Robinson, F. 'Technology and religious change: Islam and the impact of print', *Modern Asian Studies*, 27.1 (1993), 229–51.
- Roccati, A. 'Il bilinguismo interno dell'Egitto', *Vicino Oriente*, 3 (1980), 77–84.
- Rocchi, L. 'Bernardo da Parigi's *Vocabolario Italiano-Turchesco* (1665): an Ottoman-Turkish lexicographical monument still neglected', *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 132 (2015), 263–9.
- Rocchi, L. 'The importance of Bernardino Pianzola's work for Turkish lexicography' in E. Mańczak-Wohlfeld and B. Podolak (eds.), *Studies on the Turkic World: A Festschrift for Professor Stanisław Stachowski on the Occasion of His 80th Birthday* (Kraków: Jagiellonian University Press, 2010), 95–105.
- Rocchi, L. 'Gli italianismi nei testi turchi in trascrizione. Un'indagine storico-lessicografica', *Zeitschrift für Romanische Philologie*, 129.4 (2013), 888–931.
- Rocchi, L. *Il lessico turco nell'opera di Bernardino Pianzola: Materiali per la conoscenza del turco parlato di fine Settecento* (Edizioni Università di Trieste, 2009).
- Rocchi, L. *Ricerche sulla lingua osmanlı del XVI secolo. Il corpus lessicale turco del manoscritto fiorentino di Filippo Argenti (1533)* (Wiesbaden: Harrassowitz, 2007).
- Rocchi, L. 'Turkish historical lexicography', *Lexicographica*, 27 (2011), 195–220.
- Rochberg, F. *The Heavenly Writing: Divination, Horoscopy, and Astronomy in Mesopotamian Culture* (Cambridge University Press, 2004).
- Rodrigues, L. A. 'Glimpses of the Konkani language at the turn of the 16th c., III: Konkani proverbs and idioms', *Boletim do Instituto Menezes Bragança*, 123 (1979), 5–26.
- Rodrigues, L. A. 'Glimpses of the Konkani language at the turn of the 16th c., IX: religious concepts and rituals', *Boletim do Instituto Menezes Bragança*, 140 (1984), 5–19.
- Rollo, A. 'Greco medievale' e 'greco bizantino', *Annali dell'Università degli Studi di Napoli 'L'Orientale'*, 30 (2008), 429–73.
- Ronconi, F. 'Quelle grammair à Byzance? La circulation des textes grammaticaux et son reflet dans les manuscrits' in G. De Gregorio and M. Galante (eds.), *La produzione scritta tecnica e scientifica nel Medioevo: libro e documento tra scuole e professioni* (Spoleto: Fondazione Centro Italiano di Studi sull'Alto Medioevo, 2012), 63–110.
- Roquet, G. 'Savoir et pratique linguistiques dans l'Égypte pharaonique et copte', in Auroux (ed.), *Histoire des idées linguistiques*, I.119–48.
- Rosetti, A. *Brève histoire de la langue roumaine des origines à nos jours* (The Hague and Paris: Mouton, 1973).
- Rossebastiano Bart, A. *Antichi vocabolari plurilingui d'uso popolare: la tradizione del 'Solenissimo Vocabuolista'* (Alessandria: Edizioni dell'Orso, 1984).
- Rost, P. *Sprachreste der Dravāno-Polaben im Hannöverschen* (Leipzig: Hinrichs, 1907).
- Roth, R. *Introduction to the Nirukta and the Literature Related to It*, trans. D. Mackichan (Mumbai: University of Bombay, 1919).
- Rouse, M. A., and R. H. Rouse. *Authentic Witnesses: Approaches to Medieval Texts and Manuscripts* (University of Notre Dame Press, 1991).
- Rubin, A. D. 'Judeo-Italian' in Kahn and Rubin (eds.), *Handbook of Jewish Languages*, 297–364.

- Rubinstein, D. 'The Historical Dictionary of the Hebrew Language', *Kernerman Dictionary News*, 12 (2004), 34–7.
- Ruegg, D. S. 'Sanskrit–Tibetan and Tibetan–Sanskrit dictionaries, and some problems in Indo-Tibetan philosophical lexicography' in B. Oguibénine (ed.), *Lexicography in the Indian and Buddhist Cultural Field: Proceedings of the Conference at the University of Strasbourg, 25 to 27 April 1996* (Munich: Kommission für Zentralasiatische Studien, Bayerische Akademie der Wissenschaften, 1998), 115–42.
- Rusk, B. 'Old scripts, new actors: European encounters with Chinese writing, 1550–1700', *East Asian Science, Technology, and Medicine*, 26 (2007), 68–116.
- Russell, P. 'The sounds of a silence: the growth of Cormac's glossary', *Cambridge Medieval Celtic Studies*, 15 (1988), 1–30.
- Rutherford, I. (ed.). *Greco-Egyptian Interactions: Literature, Translation, and Culture, 500 BCE–300 CE* (Oxford University Press, 2016).
- Ryholt, K. 'On the contents and nature of the Tebtunis temple library: a status report' in S. Lippert and M. Schentuleit (eds.), *Tebtynis and Soknopaiu Nesos: Leben im römerzeitlichen Fajum* (Wiesbaden: Harrassowitz, 2005), 141–70.
- Ryholt, K. 'Libraries in ancient Egypt' in J. König, K. Oikonomopoulou, and G. Wolf (eds.), *Ancient Libraries* (Cambridge University Press, 2013), 23–37.
- Sabaliauskas, A. *Lietuvių kalbos tyrinėjimo istorija*, 3 vols. (Vilnius: Mokslas, 1979–2012).
- Sachau, E., H. Ethé, and A. F. L. Beeston. *Catalogue of the Persian, Turkish, Hindustani, and Pushtu Manuscripts in the Bodleian Library*, 3 vols. (Oxford: Clarendon, 1889–1953).
- Sadan, T. 'The lexicography of Hebrew' in Hanks and de Schryver (eds.), *International Handbook of Modern Lexis and Lexicography*, published online 2017.
- Sadan, T. 'Ma'agarim: Hebrew diachronic corpus and corpus query system', *Hebrew Higher Education*, 18 (2016), 209–19.
- Sadan, T. 'Tasks of the pedagogical lexicography of Modern Hebrew', *Hebrew Higher Education*, 19 (2017), 39–62.
- Said, E. *Orientalism* (London: Routledge & Kegan Paul, 1978).
- Saitō Mareshi. *Kanbunmyaku: The Literary Sinitic Context and the Birth of Modern Japanese Language and Literature*, ed. R. King and C. Laffin, trans. S. Bussell et al. (Leiden: Brill, forthcoming).
- Salazar, D. 'Towards improved coverage of Southeast Asian Englishes in the *Oxford English Dictionary*', *Lexicography*, 1 (2014), 95–108.
- Sallaberger, W. *Der babylonische Töpfer und seine Gefäße = Mesopotamian History and Environment*, ser. II: *Memoirs*, 3 (University of Ghent, 1996).
- Sallaberger, W. 'Das Ende des Sumerischen: Tod und Nachleben einer altmesopotamischen Sprache' in P. Schrijver and P.-A. Mumm (eds.), *Sprachtod und Sprachgebur*t (Bremen: Hempen, 2004), 108–40.
- Salman, P. 'Multivolume Urdu dictionary set to go online', *Dawn*, 14 May 2017.
- al-Sāmarrā'ī, I. 'al-Ma'āğim al-'Arabiyya al-qadīma' in *al-Mawsim al-īraqāfi al-awwal li-Mağma' al-Luğa al-'Arabiyya al-Urdunī*, vol. I (Amman: Mağma' al-Luğa al-'Arabiyya al-Urdunī, 1983), 183–214.
- Samuels, D. 'Review of Bray et al., *Western Apache–English Dictionary*', *American Indian Culture and Research Journal*, 24.4 (2000), 216–18.
- Sánchez Méndez, J. P. 'Arte y vocabulario de la lengua de los indios chaymas . . . : una empresa lingüística del aragonés fray Francisco de Tauste', *Archivo de Filología Aragonesa*, 68 (2012), 123–42.

- Sanders, M. *Learning Zulu: A Secret History of Language in South Africa* (Princeton University Press, 2016).
- Sankaranarayan, G. 'Preparation of administrative terminology: the role of earlier inscriptions and records', *Journal of Tamil Studies*, 24 (1983), 76–86.
- Sankara Reddy, M. *Reference Sources in Telugu: A Comprehensive Guide* (Delhi: B. R., 1996).
- Sara, S. I. 'The classical Arabic lexicographical tradition' in J. Owens (ed.), *Oxford Handbook of Arabic Linguistics* (Oxford University Press, 2013), 520–38.
- Sasaki Ken'ichi. *Jisho ni natta otoko: Kenbō sensei to Yamada sensei* (Tokyo: Bungei shunjū, 2014).
- Schaarschmidt, G. *The Historical Phonology of the Upper and Lower Sorbian Languages* (Heidelberg: Winter, 1998).
- Schaeken, J., and H. Birnbaum. *Die altkirchenslavische Schriftkultur* (Munich: Sagner, 1999).
- Schäfer, J. *Early Modern English Lexicography*, 2 vols. (Oxford: Clarendon, 1989).
- Schenker, A. M. *The Dawn of Slavic: An Introduction to Slavic Philology* (New Haven and London: Yale University Press, 1995).
- Schiffman, H. F. *Linguistic Culture and Language Policy* (London and New York: Routledge, 1996).
- Schiller, C. 'Die litauischen Seminare in Königsberg und Halle', *Nordost-Archiv*, new ser. 3 (1994), 375–92.
- Schironi, F. *From Alexandria to Babylon: Near Eastern Languages and Hellenistic Erudition in the Oxyrhynchus Glossary (P. Oxy. 1802 + 4812)* (Berlin and New York: De Gruyter, 2009).
- Schleicher, A. *Handbuch der litauischen Sprache*, vol. I, *Litauische Grammatik*, and vol. II, *Litauisches Lesebuch und Glossar* (Prague: Calve, 1856–7).
- Schlieben-Lange, B. 'Okzitanisch: Grammatikographie und Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, V.2.105–26.
- Schmid, W. *Der Atticismus in seinen Hauptvertretern von Dionysius von Halikarnass bis auf den zweiten Philostratus*, 4 vols. (Stuttgart: Kohlhammer, 1887–97).
- Schneider, E. W. *Postcolonial English: Varieties Around the World* (Cambridge University Press, 2007).
- Schoenemann, J. *De lexicographis antiquis qui rerum ordinem secuti sunt quaestiones praecursoriae* (Hannover: typis Culemannorum, 1886).
- Schön, J. F., and S. Crowther. *Journals of the Rev. James Frederick Schön and Mr. Samuel Crowther, who, with the Sanction of Her Majesty's Government, Accompanied the Expedition up the Niger in 1841, in Behalf of the Church Missionary Society* (London: Hatchard [et al.], 1842).
- Schouten, M. J. 'Malay and Portuguese as contact languages in the Southeast Asian Archipelago, 16th–18th centuries' in Häberlein and Keese (eds.), *Sprachgrenzen, Sprachkontakte, Kulturelle Vermittler*, 345–54.
- Schramm, W. *Bann, Bann! Eine sumerisch-akkadische Beschwörungserie* = Göttinger Arbeitshefte zur altorientalischen Literatur, 2 (Seminar für Keilschriftforschung der Universität Göttingen, 2001).
- Schramm, W. *Ein Compendium sumerisch-akkadischer Beschwörungen* = Göttinger Beiträge zum Alten Orient, 2 (Universitätsverlag Göttingen, 2008).
- Schreyer, R. *The European Discovery of Chinese (1550–1615), or, the Mystery of Chinese Unveiled* = Cahiers voor Taalkunde, 5 (Amsterdam: Stichting Neerlandistiek, 1992).

- Schreyer, R. 'Illustrations of authority: quotations in Samuel Johnson's *Dictionary of the English Language*', *Lexicographica*, 16 (2000), 58–103.
- Schröder, B.-J. *Titel und Text: Zur Entwicklung lateinischer Gedichtüberschriften, mit Untersuchungen zu lateinischen Buchtiteln, Inhaltsverzeichnissen, und anderen Gliederungsmitteln* (Berlin and New York: De Gruyter, 1999).
- Schuller, R. 'An unknown Matlatsinka manuscript vocabulary of 1555–1557', *Indian Notes*, 7.2 (1930), 175–94.
- Schütz, G. 'Cuervos Wörterbuch als herausragendes Werk der hispanischen Lexikographie' in Hausmann et al. (eds.), *Wörterbücher*, II.1767–70.
- Schwarzwald, O. 'Modern Hebrew dictionaries', *Kernerman Dictionary News*, 12 (2004), 18–22.
- Schweickard, W. 'Italian' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 672–87.
- Sciarra, E. 'Note sul codice Vat. Barb. gr. 70 e sulla tradizione manoscritta dell'*Etymologicum Gudianum*' in R. M. Piccione (ed.), *Selecta colligere II: Beiträge zur Technik des Sammelns und Kompilierens griechischer Texte von der Antike bis zum Humanismus* (Alessandria: Edizioni dell'Orso, 2005), 355–402.
- Scott, W. H. 'Sixteenth-century Tagalog technology from the *Vocabulario de la lengua tagalo* of Pedro de San Buenaventura, O.F.M.' in R. Carle et al. (eds.), *Gava': Studies in Austronesian Languages and Cultures Dedicated to Hans Kähler* (Berlin: Reimer, 1982), 523–37.
- Sebeok, T. A. (ed.), *Current Trends in Linguistics*, vol. VIII, *Linguistics in Oceania* (The Hague: Mouton, 1971).
- Seche, M. *Șchiță de istorie a lexicografiei române*, 2 vols. (Bucharest: Editura Științifică, 1966–9).
- Seebold, E. 'Textgeschichte des 8. Jahrhunderts (und der früheren Überlieferung)' in Seebold (ed.), *Chronologisches Wörterbuch des deutschen Wortschatzes*, vol. I, *Der Wortschatz des 8. Jahrhunderts (und frühere Quellen)* (Berlin and New York: De Gruyter, 2001), 9–69.
- Seeley, C. *A History of Writing in Japan* (Leiden: Brill, 1991).
- Segert, S. and Y. Sabar. 'Hebrew and Aramaic lexicography' in Hausmann et al. (eds.), *Wörterbücher*, III.2424–38.
- Segovia Gordillo, A. 'El vocabulario hispano-quechua (1608) de González Holguín, ¿heredero de Nebrija?', *Philologia Hispalensis*, 28.1/2 (2014), 7–23.
- Seidensticker, T. 'Lexicography: Classical Arabic', in *Encyclopedia of Arabic Language and Linguistics*, III.30–7.
- Sellheim, R. 'Mathal' in *Encyclopaedia of Islam*, 2nd edn, VI.815–25.
- Sengupta, M. *Becoming Assamese: Colonialism and New Subjectivities in Northeast India* (Abingdon, UK: Routledge, 2016).
- Seymour, D. J. (ed.). *From the Land of Ever Winter to the American Southwest: Athapaskan Migrations, Mobility, and Ethnogenesis* (Salt Lake City: University of Utah Press, 2012).
- Shaaban, K. 'Language policies and language planning' in *Encyclopedia of Arabic Language and Linguistics*, II.694–707.
- Shastree, K. K. 'Monolingual and bilingual dictionaries in Gujarat' in Misra (ed.), *Lexicography in India*, 29–40.

- Shaw, G. W. 'The Copenhagen copy of Henriques' *Flos Sanctorum*', *Fund og Forskning*, 32 (1993), 39–50.
- Shaw, G. W. *Printing in Calcutta to 1800: A Description and Checklist of Printing in Late 18th-Century Calcutta* (London: Bibliographical Society, 1981).
- Shaw, G. W. 'South Asia' in S. Eliot and J. Rose (eds.), *A Companion to the History of the Book* (Malden, MA, and Oxford: Blackwell, 2007), 126–37.
- Shaw, G. W. 'South Asia and the history of the book', *SALG Newsletter*, 43 (1996), 17–20.
- Shaw, J. *Contributions to a Study of the Printed Dictionary in France before 1539* (Toronto: EDICTA, 1997), homes.chass.utoronto.ca/~wulfric/edicta/shaw/.
- Sherānī, H. M. 'Dībācah-i avval' in H. M. Sherānī (ed.), *Hifẓ al-lisān (ma'rūf ba-nām-i khālīq bārī)* (Delhi: Anjuman-i Taraqqī-yi Urdū, 1944).
- Shinmura Takeshi. *Kōjien monogatari* (Tokyo: Geijutsu seikatsusha, 1970).
- Shinmura Yasushi. *Kōjien wa naze umareta ka: Shinmura Izuru no ikita kiseki* (Kyoto: Sekai shisōsha, 2017).
- Si, J. 'Reprinting Robert Morrison's dictionary: producers, literary audience, and the English language market in 19th-century Shanghai', *Frontiers of History in China*, 6.2 (2011), 229–42.
- Sidarus, A. Y. 'Coptic lexicography in the Middle Ages: The Coptic Arabic Scalae' in R. McL. Wilson (ed.), *The Future of Coptic Studies* (Leiden, Brill: 1978), 125–42.
- Sidarus, A. Y. 'La littérature Copte à la première époque Arabe (VIIe–XIe siècles)' in A. Boud'hors and C. Louis (eds.), *Études Coptes XIV, Seizième journée d'études = Cahiers de la bibliothèque copte*, 21 (Paris: Boccard, 2016), 193–221.
- Sidarus, A. Y. 'Medieval Coptic grammars in Arabic: the Coptic Muqaddimat', *Journal of Coptic Studies*, 3 (2001), 63–79.
- Sijpesteijn, P. M. 'Arabic papyri and Islamic Egypt' in Bagnall (ed.), *Oxford Handbook of Papyrology*, 452–72.
- Sim Gyeongho. 'Hanguk ui unseo wa unseo gwaryong bangsik', *Hanja, Hanmun yeongu*, 5 (2009), 137–86.
- Simmons, R. V. 'Mandarin, varieties of' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, II.672–82.
- Simmons, R. V. 'Rime tables and rime table studies', in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, III.609–19.
- Simmons, R. V. 'Transcription systems: Gwoyeu Romatzyh 國語羅馬字' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, IV.416–24.
- Simpson, J., R. Amery, and M.-A. Gale. 'I could have saved you linguists a lot of time and trouble: 180 years of research and documentation of South Australia's Indigenous languages, 1826–2006' in McGregor (ed.), *Encountering Aboriginal Languages*, 85–144.
- Simpson, J. A. '100, 000 entries published and counting' (2011), public.oed.com/the-oed-today/recent-updates-to-the-oed/previous-updates/december-2011/100000-entries-published-and-counting/.
- Simpson, J. A. 'English lexicography after Johnson to 1945', in Hausmann et al. (eds.), *Wörterbücher*, II.1953–66.
- Simpson, J. A., E. S. C. Weiner, and P. N. Durkin. 'The Oxford English Dictionary today', *Transactions of the Philological Society*, 102 (2004), 335–81.

- Sinclair, J. McH. (ed.). *Looking Up: An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary* (London: Collins ELT, 1987).
- Singh, R. A. 'Lexicography of New Indo-Aryan' in Hausmann et al. (eds.), *Wörterbücher*, III.2507–21.
- Singh, Y. *Modernization of Indian Tradition: A Systematic Study of Social Change* (Jaipur: Rawat, 1994).
- Sirjānī, 'A.-A. S. 'Dehkodā' in *Encyclopædia Iranica*.
- Sitoe, B. 'Lexicografia da língua Tsonga: uma proposta metodológica', unpublished MA, Universidade de Varsóvia (1991).
- Skinner, D. *The Story of Ain't: America, Its Language, and the Most Controversial Dictionary Ever Published* (New York: Harper, 2012).
- Sluiter, I. 'Ancient etymology: a tool for thinking' in Montanari et al. (eds.), *Brill's Companion to Ancient Greek Scholarship*, II.896–922.
- Smith, J. S. 'Japanese writing' in Daniels and Bright (eds.), *The World's Writing Systems*, 209–17.
- Smith, O. L. 'A note on Holkham Gr. 88 and Marc. Gr. 622', *Maia*, 27 (1975), 205.
- Smith-Stark, T. C. 'Lexicography in New Spain (1492–1611)' in Zwartjes et al. (eds.), *Missionary Linguistics IV*, 3–82.
- Smith-Stark, T. C. 'Phonological description in New Spain' in Zwartjes and Altman (eds.), *Missionary Linguistics II*, 3–64.
- Smitskamp, R. *Philologia Orientalis: A Description of Books Illustrating the Study and Printing of Oriental Languages in 16th- and 17th-Century Europe* (Leiden: Brill, 1992).
- Söderblom Saarela, M. 'Shape and sound: organizing dictionaries in late imperial China', *Dictionaries*, 35 (2014), 187–208.
- Song Eunha. 'Yeonghansajeon-ui yeogsawa sidaebyeol yeonghansajeon-ui teugjing', *Journal of Korealex*, 22 (2013), 127–59.
- Song, K.-J. *The Study of Foreign Languages in the Chosŏn Dynasty, 1392–1910* (Seoul: Jimoondang, 2001).
- Sonnino, M. 'I frammenti della commedia greca citati da Prisciano e la fonte del lessico sintattico del libro XVIII dell'Ars', in Martorelli (ed.), *Greco antico nell'Occidente carolingio*, 163–204.
- Spencer, J. 'S. W. Koelle and the problem of notation for African languages, 1847–1855', *Sierra Leone Language Review*, 5 (1966), 83–105.
- Speranzi, D. 'Il copista del lessico di Esichio (Marc. gr. 622)' in *Storia della scrittura e altre storie*, ed. D. Bianconi (Rome: Accademia Nazionale dei Lincei, 2014), 101–46.
- Splett, J. 'Abrogans deutsch' in *Die deutsche Literatur des Mittelalters: Verfasserlexikon*.
- Srinivas, M. N. *Social Change in Modern India* (Berkeley: University of California Press, 1966).
- Stachowski, M. 'Remarks on the phonetic value of the letters ⟨y⟩ and ⟨ü⟩ in Franciscus Meninski's Ottoman Turkish Thesaurus (1680)', *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 129 (2012), 189–97.
- Stachowski, S. 'François à Mesgnien Meninski und sein Thesaurus Linguarum Orientalium' in Meninski, *Thesaurus* (2000), xxiii–xxxiv.
- Stachowski, S. *Lexique turc dans le 'Vocabulaire' de P. F. Viguier (1790)* (Kraków: Księgarnia Akademicka, 2002).

- Staff of DOST. 'A re-editing of GIF', in C. Macafee and I. Macleod (eds.), *The Nuttis Schell: Essays on the Scots Language* (Aberdeen University Press, 1987), 25–58.
- Stammerjohann, H. (ed.). *Lexicon Grammaticorum: A Bio-Bibliographical Companion to the History of Linguistics*, 2nd edn (Tübingen: Niemeyer, 2009).
- Stankiewicz, E. *Grammars and Dictionaries of the Slavic Languages from the Middle Ages up to 1850: An Annotated Bibliography* (Berlin etc.: Mouton, 1984).
- Stark, U. *An Empire of Books: The Naval Kishore Press and the Diffusion of the Printed Word in Colonial India* (Ranikhet: Permanent Black, 2007).
- Starnes, D. T. *Renaissance Dictionaries, English–Latin and Latin–English* (Austin: University of Texas Press, 1954).
- Starnes, D. T., and G. E. Noyes. *The English Dictionary from Cawdrey to Johnson 1604–1755* (Chapel Hill: University of North Carolina Press, 1946).
- Starr, J. B. 'The legacy of Robert Morrison', *International Bulletin of Mission Research*, 22.2 (1998), 73–6.
- Steadman-Jones, R. *Colonialism and Grammatical Representation* (Oxford: Blackwell, 2007).
- Stein, G. *The English Dictionary Before Cawdrey* (Tübingen: Niemeyer, 1985).
- Steiner, R. C. 'Saadia vs. Rashi: on the shift from meaning-maximalism to meaning-minimalism in medieval biblical lexicology', *Jewish Quarterly Review*, 88.3/4 (1998), 213–58.
- Stimson, H. M. *The Jongyuan in yunn: A Guide to Old Mandarin Pronunciation* (New Haven: Far Eastern Publications, Yale University, 1966).
- Stockigt, C. 'Pama-Nyungan morphosyntax: lineages of early description', unpublished PhD thesis, University of Adelaide (2016).
- Strandberg, E. 'Lexicography of Middle Indo-Aryan' in Hausmann et al. (eds.), *Wörterbücher*, III.2497–2507.
- Strotmann, V. *Majd al-Dīn al-Fīrūzābādī: A Polymath on the Eve of the Early Modern Period* (Leiden: Brill, 2016).
- Sueiro Justel, J. *Historia de la lingüística española en Filipinas (1580–1898)* (Lugo: Axac, 2003).
- Sugimoto Tsutomu 杉本つとむ. *Jisho, jiten no kenkyū ichi* 辞書・事典の研究 I = Sugimoto Tsutomu chosaku senshū, 6 (Tokyo: Yasaka Shobō, 1999).
- Święcicka, E. 'Interpreter Yovhannes Ankiwrac'i also called Giovanni Molino', *Folia Orientalia*, 36 (2000), 329–42.
- Swiggers, P. 'La linguistique romane, de Friedrich Diez à l'aube du XXe siècle' in Klump et al. (eds.), *Manuel des langues romanes*, 43–63.
- Sybesma, R., et al. (eds.). *Encyclopedia of Chinese Language and Linguistics*, 5 vols. (Leiden and Boston: Brill, 2017).
- 'System of orthography for native names of places', *Proceedings of the Royal Geographical Society and Monthly Record of Geography*, new monthly ser., 7.8 (1885), 535–6.
- Tal, A. 'The Historical Dictionary of the Hebrew Language of the Academy of the Hebrew Language: aims and achievements' in S. Schorch and E.-J. Waschke (eds.), *Bibliche Exegese und hebräische Lexikographie: Das 'Hebräisch-deutsche Handwörterbuch' von Wilhelm Gesenius als Spiegel und Quelle alttestamentlicher und hebräischer Forschung*, 200 Jahre nach seiner ersten Auflage (Berlin: De Gruyter, 2013), 16–24.
- al-Ṭalībī, M. *al-Muḥaṣṣaṣ li-Ibn Sīda: Dirāsa, dalīl* (Tunis: al-Maṭba'a al-'Aṣriyya, 1965).
- Talmon, R. *Arabic Grammar in its Formative Age: Kitāb al-'Ayn and its Attribution to Ḥalīl b. Aḥmad* (Leiden: Brill, 1997).

- Tancke, G. *Die italienischen Wörterbücher von den Anfängen bis zum Erscheinen des 'Vocabolario degli Accademici della Crusca' (1612)* (Tübingen: Niemeyer, 1984).
- Tang, J. 'Xiàndài Hànyǔ cídiǎn 現代漢語詞典' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, IV.608–9.
- Tanret, M. *Per aspera ad astra: L'apprentissage du cunéiforme à Sippar-Amnanum pendant la période paléobabylonienne tardive* = *Mesopotamian History and Environment*, ser. III: Texts, 1/2 (University of Ghent, 2002).
- Tauer, F. 'Persian learned literature to the end of the 18th century' in K. Jahn (ed.), *History of Iranian Literature* (Dordrecht: Reidel, 1968), 419–82.
- Taylor, K. 'Literacy in early seventeenth-century northern Vietnam' in M. A. Aung-Thwin and K. R. Hall (eds.), *New Perspectives on the History and Historiography of Southeast Asia: Continuing Explorations* (London: Routledge, 2011), 183–98.
- Teleman, U. 'The role of language cultivators and grammarians for the Nordic linguistic development in the 16th, 17th and 18th centuries' in Bandle et al. (eds.), *Nordic Languages*, II.1379–96.
- Téllez Nieto, H. *Acercamiento filológico al vocabulario trilingüe en español-latín y náhuatl atribuido a fray Bernardino de Sahagún* (Mexico City: Universidad Nacional Autónoma de México, 2004).
- Téné, D., and J. Barr. 'Linguistic literature, Hebrew', in C. Roth et al. (eds.), *Encyclopaedia Judaica*, 16 vols. (Jerusalem: Keter; New York: Macmillan, 1971–2), XVI, cols. 1352–90 (corresponding to 2nd edn, updated by A. Maman, XIII.29–61).
- Thankappan Nair, P. 'Father Hanxleden – pioneer Indologist', in Pallath (ed.), *Arnos Padiri*, 46–60.
- Thieberger, N. 'Daisy Bates in the digital world' in P. K. Austin, H. Koch, and J. Simpson (eds.), *Language, Land, and Song* (London: EL, 2016), 102–14.
- Thieberger, N. 'The lexicography of indigenous languages in Australia and the Pacific' in Hanks and de Schryver (eds.), *International Handbook of Modern Lexis and Lexicography*, published online 2015.
- Thomas, J. *The Universal Dictionary of Biography and Mythology*, 2 vols. in 4 parts (New York: Cosimo, 1887).
- Thomason, A. K. *Luxury and Legitimation: Royal Collecting in Ancient Mesopotamia* (Aldershot: Ashgate, 2005).
- Thompson, D. J. 'The multilingual environment of Persian and Ptolemaic Egypt: Egyptian, Aramaic, and Greek' in Bagnall (ed.), *Oxford Handbook of Papyrology*, 395–417.
- Thompson, J. *Edward William Lane 1801–1876: The Life of the Pioneering Egyptologist and Orientalist* (London: Haus, 2010).
- Thomson, R. M. 'England and the twelfth-century Renaissance', *Past and Present*, 101 (1983), 3–21.
- Tiberius, C., A. Aalstein, and J. Hoogland. 'OMBI bilingual lexical resources: Arabic–Dutch/Dutch–Arabic' in A. Dykstra and T. Schoonheim (eds.), *Proceedings of the XIV Euralex International Congress* (Leeuwarden: Fryske Akademy, 2010), 188–9 (printed abstract) and 855–60 (chapter on accompanying CD), *euralex.org*.
- Tietze, A. 'Die fremden Elemente im Osmanisch-Türkischen' in Gy. Hazai (ed.), *Handbuch der türkischen Sprachwissenschaft*, part 1 (Wiesbaden: Harrassowitz, 1990), 104–18.

- Tietze, A. 'Die Lexikographie der Turksprachen. I: Osmanisch-Türkisch' in Hausmann et al. (eds.), *Wörterbücher*, III.2399–2407.
- Tobi, Y. 'Piyuṭe Rav Se'adya Ga'on: mahadura madda'it (šel ha-Yotsrot) u-mavo kəlali li-ytsirato', unpublished PhD thesis, Hebrew University of Jerusalem (1982).
- Tolkiehn, J. 'Lexikographie', in *Paulys Realencyclopädie*, XXIV.2432–82.
- Tomalin, M. 'Review of Zwartjes et al., *Missionary Linguistics IV*', *Historiographia Linguistica*, 37.1/2 (2010), 238–46.
- Toomer, G. J. *Eastern Wisdom and Learning: The Study of Arabic in Seventeenth-Century England* (Oxford: Clarendon, 1996).
- Torero, A. 'Entre Roma y Lima: el *Lexicon* quichua de fray Domingo de Santo Tomás (1560)', in K. Zimmermann (ed.), *La descripción de las lenguas amerindias en la época colonial* (Madrid: Iberoamericana; Frankfurt am Main: Vervuert, 1997), 271–90.
- Tosi, R. 'La lessicografia e la paremiografia in età alessandrina ed il loro sviluppo successivo' in F. Montanari (ed.), *La philologie grecque à l'époque hellénistique et romaine* = *Entretiens Hardt*, 40 (Vandœuvres-Geneva, Fondation Hardt, 1994), 143–209.
- Tosi, R. 'Polluce: struttura onomastica e tradizione lessicografica' in C. Bearzot, F. Landucci, and G. Zecchini (eds.), *L' Onomasticon di Giulio Polluce: tra lessicografia e antiquaria* (Milan: Vita e Pensiero, 2007), 3–16.
- Tosi, R. 'Typology of Greek lexicography', in Montanari et al. (eds.), *Brill's Companion to Ancient Greek Scholarship*, I.622–40.
- Trần Trọng Dương 'Vấn đề khai thác từ cổ qua hệ thống từ điển và các văn bản chữ Nôm', *Tạp Chí Hán Nôm* 1.98 (2010), 17–36.
- Trần Trọng Dương. *Khảo cứu từ điển song ngữ Hán Việt Nhật dụng thường đàm của Phạm Đình Hổ* (Hanoi: NXB Văn Học, 2016).
- Trần Trọng Dương 陳仲洋. 《中世紀越南漢字詞典的類型與特點》(Zhongshiji Yuenan hanzi cidian de leixing yu tedian), in He Huazhen 何華珍 and Nguyễn Tuấn Cường 阮俊強 (eds.), *東亞籍漢與越南漢喃古辭書研究* (*Dongya hanji yu Yuenan Han-nan gucishu yanjiu*) (Beijing: Zhong shehui kexue chubanshe, 2017), 15–41.
- Trautmann, T. R. 'Dr Johnson and the pandits: imagining the perfect dictionary in colonial Madras', *Indian Economic & Social History Review*, 38.4 (2001), 375–97.
- Trautmann, T. R. 'Hullabaloo about Telugu', *South Asia Research*, 19.1 (1999), 53–70.
- Trautmann, T. R. *Languages and Nations: The Dravidian Proof in Colonial Madras* (Berkeley: University of California Press, 2006).
- Trautmann, T. R. (ed.). *The Madras School of Orientalism: Producing Knowledge in Colonial South India* (New Delhi: Oxford University Press, 2011).
- Trench, R. W. C. *On Some Deficiencies in Our English Dictionaries* (London: Parker, 1857).
- Trotter, D. 'Gallo-Romance II: synchronic lexicography' in Gouws et al. (eds.), *Dictionaries: Supplementary Volume*, 663–72.
- Trudgill, P., and J. Hannah. *International English: A Guide to the Varieties of Standard English*, 5th edn (London, Hodder Education, 2008).
- Trudinger, D. 'Converting salvation: Protestant missionaries in Central Australia, 1930s–1940s', unpublished PhD thesis, Australian National University (2004).
- Truschke, A. 'Defining the other: an intellectual history of Sanskrit lexicons and grammars of Persian', *Journal of Indian Philosophy*, 40.6 (2012), 635–68.
- Tsantsanoglou, K. *To lexiko tou Photiou* (Thessalonica: Hetaireia Makedonikon Spoudon, 1967).

- Tsien Tsuen-Hsuin. *Paper and Printing* = J. Needham et al. (eds.), *Science and Civilisation in China*, vol. V.1, revised reprint (Cambridge University Press, 1987).
- Twine, N. *Language and the Modern State: The Reform of Written Japanese* (London: Routledge, 1991).
- Ucciardello, G. *Hypomnemata papiracei e lessicografia* (Università degli Studi di Messina, 2012).
- Ucciardello, G. 'Nuovi frammenti di oratori attici nel Vat. gr. 7', *Miscellanea Bibliothecae Apostolicae Vaticanae*, 14 (2007), 431–82.
- UNESCO. 'Education: literacy rate', data.uis.unesco.org.
- Uqashi, Z. 'Millon Šəmuʔel ha-Nagid ʿal pi širato', unpublished PhD dissertation, Hebrew University of Jerusalem (1998).
- Urbánková, E. *Soupis děl J. A. Komenského v Československých knihovnách, archivech, a museích* (Prague: Státní Pedagogické Nakladatelství, 1959).
- Vaiyapuri Pillai, S. 'History of Tamil lexicography' in *Tamil Lexicon*, I.xxv–l.
- Valente, S. 'Alphabetical dictionaries: from antiquity to the Byzantine period' in *Encyclopedia of Ancient Greek Language and Linguistics*, I.100–3.
- Valente, S. *The Antiatticist* (Berlin and Boston: De Gruyter, 2015).
- Valente, S. 'Dictionaries of onomastics: from antiquity to the Byzantine period' in *Encyclopedia of Ancient Greek Language and Linguistics*, I.483–5.
- Valente, S. 'Etymological dictionaries: from antiquity to the Byzantine period' in *Encyclopedia of Ancient Greek Language and Linguistics*, I.574–7.
- Valente, S. 'From Plato to the Byzantine Etymologica: the etymologies of [hērōes] in the Etymologicum Gudianum, in C. Charalambakis et al. (eds.), *Festschrift in Honour of John N. Kazazis* (Berlin and Boston: De Gruyter, forthcoming).
- Valente, S. 'Hesychius' in *Lexicon of Greek Grammarians of Antiquity*.
- Valente, S. 'Porphyrios und der "Kratylos" in einer Glosse des Etymologicum Gudianum (510, 15f. de Stefani)', *Zeitschrift für Papyrologie und Epigraphik*, 201 (2017), 64–7.
- Valente, S. 'Synonymica: from antiquity to the Byzantine period' in *Encyclopedia of Ancient Greek Language and Linguistics*, III.352–5.
- Van Dam, P. *Beschryvinge van de Oostindische Compagnie*, ed. F. W. Stapel with C. W. T. van Boetzelaer van Asperen en Dubbeldam, 4 vols. in 7 parts (The Hague: Nijhoff, 1927–1943).
- Van der Kuijp, L. W. J. 'On the vicissitudes of Subhūticandra's *Kāmadhenu Commentary* on the *Amarakoṣa* in Tibet', *Journal of the International Association of Tibetan Studies*, 5 (2009), 1–105.
- Van der Loon, P. 'The Manila incunabula and early Hokkien studies, Part 2', *Asia Major*, n. s., 13 (1967), 95–186.
- Van Hal, T. 'À la recherche d'une grammaire perdue: Johann Ernst Hanxleden's *Grammatica Grandonica* retrieved', *Historiographia Linguistica*, 37.3 (2010), 445–57.
- Van Hal, T. 'Protestant pioneers in Sanskrit studies (ca. 1700–1750): an overlooked chapter in South-Indian missionary linguistics', *Historiographia Linguistica*, 43.1/2 (2016), 99–144.
- Van Hal, T., A. Peetermans, and Z. Van Loon, 'Presentation of the RELiCTA database Repertory of Early Modern Linguistic and Catechetical Tools of America, Asia and Africa', *Beiträge zur Geschichte der Sprachwissenschaft* 28.2 (2018), 293–306.
- Van Rossem, S. *Portret van een woordenaar: Cornelis Kiliaan en het woordenboek in de Nederlanden* (Antwerp: Provincie Antwerpen, Departement Cultuur, 2007).

- van Spaandonck, M. *Practical and Systematical Swahili Bibliography: Linguistics 1850–1963* (Leiden: Brill, 1965).
- Veisbergs, A. 'Overview of the early development of the lexicography of the three Baltic nations (from 17th to 19th century)', *Baltic Journal of European Studies*, 1.1(9) (2011), 307–25.
- Veisbergs, A. 'Translators' tools: Latvian and Lithuanian bilingual lexicographical tradition compared', *Vertimo studijos*, 3 (2010), 131–49.
- Veldhuis, N. C. 'Domesticizing Babylonian scribal culture in Assyria: transformation by preservation' in W. S. van Egmond and W. H. van Soldt (eds.), *Theory and Practice of Knowledge Transfer: Studies in School Education in the Ancient Near East and Beyond* = Publications de l'Institut historique et archéologique néerlandais de Stamboul, 121 (Leiden: Nederlands Instituut voor het Nabije Oosten, 2012), 11–24.
- Veldhuis, N. C. *History of the Cuneiform Lexical Tradition* = Guides to the Mesopotamian Textual Record, 6 (Münster: Ugarit-Verlag, 2014).
- Veldhuis, N. C. 'How did they learn cuneiform? "Tribute/Word List C" as an elementary exercise' in P. Michalowski and N. C. Veldhuis (eds.), *Approaches to Sumerian Literature: Studies in Honour of Stip (H. L. J. Vanstiphout)* = Cuneiform Monographs, 35 (Leiden: Brill/STYX, 2006), 181–200.
- Veldhuis, N. C. 'Purity and access: A catalog of lexical texts dedicated to Nabû', *Journal of Cuneiform Studies*, 65 (2013), 169–80.
- Veldhuis, N. C. *Religion, Literature, and Scholarship: The Sumerian Composition 'Nanše and the Birds', with a Catalogue of Sumerian Bird Names* = Cuneiform Monographs, 22 (Leiden: Brill/STYX, 2004).
- Venås, K. 'General tendencies in Nordic language cultivation and language planning' in Bandle et al. (eds.), *Nordic Languages*, II.2013–24.
- Venkatachalapathy, A. R. "'Grammar, the frame of language": Tamil pandits at the College of Fort St George', in T. R. Trautmann (ed.), *The Madras School of Orientalism: Producing Knowledge in Colonial South India* (New Delhi: Oxford University Press, 2009), 113–25.
- Verdelho, T. 'Portugiesisch: Lexikographie/Lexicografia' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, VI.2.673–92.
- Verlagsteam Wissenmedia. 'Auf Wiedersehen' (16 July 2014), web.archive.org/web/20140716023211/http://www.wissenmedia.de/.
- Vernus, P. 'L'égypto-copte', in D. Cohen (ed.), *Les langues chamito-sémitiques*, part 3, of J. Perrot (ed.), *Les langues dans le monde ancien et moderne* (Paris: Editions du Centre National de la Recherche Scientifique, 1988), 163–206.
- Vernus, P. 'Entre Néo-égyptien et Démotique: la langue utilisée dans la traduction du Rituel de repousser l'agressif', *Revue d'Égyptologie*, 41 (1990), 153–208.
- Versteegh, K. *The Arabic Linguistic Tradition* (London: Routledge, 1997).
- Vielle, C. 'Devotional Christianity and pre-Indology in early eighteenth-century Kerala: Johann Ernst Hanxleden s.j., alias Arnos Padiri, scholar and poet', in K. Preisendanz and J. Buss (eds.), *Dimensions of the Christian Encounter with the Religions of India: Aims, Possibilities, Ramifications* (Vienna: De Nobili Research Library, 2018).
- Vielle, C. 'A new manuscript of J. E. Hanxleden's works on Malayalam language' in P. Kannampuzha (ed.), *Arnos Padiri Malayala Sahithyathile Mahamaneeshi* (Kochi: Liturgical Research Centre, 2018), 201–11.

- Vittmann, G. *Ägypten und die Fremden im ersten vorchristlichen Jahrtausend* (Mainz: Zabern, 2003).
- Vogel, C. *Die Anfänge des westlichen Studiums der altindischen Lexikographie* (Opladen: Westdeutscher, 1999).
- Vogel, C. *Indian Lexicography*, revised and enlarged edn, ed. J. Hanneder and M. Straube = *Indologica Marpurgensia*, 6 (Munich: Kirchheim, 2015).
- Vogel, C. *Zum Aufbau altindischer Sanskritwörterbücher der vorklassischen Zeit* = *Nordrhein-Westfälische Akademie der Wissenschaften, Vorträge G 341* (Opladen: Westdeutscher, 1995).
- Volten, A. 'An "alphabetical" dictionary and grammar in Demotic (Pap. Carlsberg XII verso)', *Archiv Orientální*, 20 (1952), 496–508.
- Vončina, J. 'Belostenčeva hibridni jezik', *Suvremena lingvistika*, 43–4 (1997), 325–39.
- von Soden, W. 'Leistung und Grenze sumerischer und babylonischer Wissenschaft', *Die Welt als Geschichte*, 2 (1936), 411–64 and 509–57.
- Vulfson, J. J. 'Slavica: USSR', *Quarterly Journal of Current Acquisitions*, 12.3 (1955), 161–79.
- Wafer, J., and H. M. Carey. 'Waiting for Biraban: Lancelot Threlkeld and the "Chibcha phenomenon" in Australian missionary linguistics', *Language & History*, 54.2 (2011), 112–39.
- Wagoner, P. B. "'Sultan among Hindu kings": dress, titles, and the Islamicization of Hindu culture at Vijayanagara', *Journal of Asian Studies*, 55.4 (1996), 851–88.
- Wagoner, P. B. 'Telugu literature of the Qutb Shahi period' in N. N. Haidar and M. Sardar (eds.), *Sultans of the South: Arts of India's Deccan Courts, 1323–1687* (New York: Metropolitan Museum of Art, 2011), 99–101.
- Wahrig-Burfeind, R. 'Das Wörterbuch als Datenbank: das Leben und Wirken Gerhard Wahrigs', *Der Sprachdienst* (2014), part 5–6, 196–208.
- Wakelin, M. F. 'The treatment of dialect in English dictionaries', in Burchfield (ed.), *Studies in Lexicography*, 156–77.
- Wallace-Hadrill, A. *Suetonius* (London: Duckworth, 1983).
- Walls, A. F. 'Crowther, Samuel Adjai (or Ajayi)', in G. H. Anderson (ed.), *Biographical Dictionary of Christian Missions* (New York: Macmillan, 1998), 160–1.
- Wang Li 王力. *漢語史稿 Hanyu shi gao* (Beijing: Kexue chubanshe, 1957).
- Wangdi, P. 'Language policy and planning in Bhutan' (2015), dzongkha.gov.bt/en/article/papers.
- Warder, A. K. *Indian Kāvya Literature*, 8 vols. in 9 parts (Delhi: Banarsidass, 1972–2004).
- Washbrook, D. "'To each a language of his own": language, culture, and society in colonial India' in P. J. Corfield (ed.), *Language, History and Class* (Oxford: Blackwell, 1991), 179–203.
- Watson, W. J. 'Ibrāhīm Mūteferriḳa and Turkish incunabula', *Journal of the American Oriental Society*, 88.3 (1968), 435–41.
- Webb, J. 'Cullers of words: writing the *Dictionary of Newfoundland English*', *Dictionaries*, 33 (2012), 58–82.
- Weijers, O. 'Lexicography in the Middle Ages', *Viator*, 20 (1989), 139–53 (reprinted in *Ashgate Critical Essays*, II.113–27).
- Westenholz, J. G. 'Thoughts on esoteric knowledge and secret lore' in J. Prosecký (ed.), *Intellectual Life of the Ancient Near East: Papers Presented at the 43rd Rencontre Assyriologique Internationale, Prague, July 1–5, 1996* (Prague: Academy of Sciences of the Czech Republic, Oriental Institute, 1998), 451–62.

- Whitney, W. D. *Language and the Study of Language* (London: Trübner, 1867).
- Wieder, N. *The Judean Scrolls and Karaism*, 2nd edn (Jerusalem: Ben-Zvi Institute for the Study of Jewish Communities in the East, Yad Izhak Ben-Zvi, and the Hebrew University of Jerusalem, 2005).
- Wiegand, H. E. 'Historische Lexikographie', in W. Besch, A. Betten, O. Reichmann, and S. Sonderegger (eds.), *Sprachgeschichte: ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung* (Berlin and New York: De Gruyter, 1998), 1.643–715.
- Wilkins, D., and D. Nash. 'The European "discovery" of a multilingual Australia: the linguistic and ethnographic successes of a failed expedition' in McGregor (ed.), *Encountering Aboriginal Languages*, 485–507.
- Wilkinson, E. P. *Chinese History: A New Manual*, 4th edn (Cambridge, MA, and London: Harvard University Asia Center, Harvard University Press, 2015).
- Willoughby, J. 'The provision of books in the English secular college' in C. Burgess and M. Heale (eds.), *The Late Medieval English College and Its Context* (York Medieval, 2008), 154–78.
- Wilson, N. G. 'On the transmission of the Greek lexica', *Greek, Roman, and Byzantine Studies*, 23 (1982), 369–75.
- Wilson, N. G. *Scholars of Byzantium*, 2nd edn (London: Duckworth, 1996).
- Wilts, O. 'Die nordfriesische Lexikographie' in Munske (ed.), *Handbook of Frisian Studies*, 353–66.
- Winkelmann, O. 'Rumänisch: Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, III.492–507.
- Winter, M. 'Kāngxī zìdiǎn 康熙字典' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, II.481–5.
- Winterbottom, A. *Hybrid Knowledge in the Early East India Company World* (Basingstoke: Palgrave Macmillan, 2016).
- Wippermann, D. 'Transcription systems: Zhùyīn fúhào 注音符號' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, IV.436–43.
- Witt, R. G. *The Two Latin Cultures and the Foundation of Renaissance Humanism in Medieval Italy* (Cambridge University Press, 2012).
- Wolfart, H. C., and J. F. Carroll. *Meet Cree: A Guide to the Cree Language*, 2nd edn (Edmonton: University of Alberta Press, 1981).
- Wolff, J. U. 'The *Vocabulario de lengua tagala* of Fr Pedro de San Buenaventura (1613)' in L. A. Reid, E. Ridruejo, and T. Stolz (eds.), *Philippine and Chamorro Linguistics Before the Advent of Structuralism* (Berlin: Akademie, 2011), 33–48.
- Woll, D. 'Portugiesische Lexikographie' in Hausmann et al. (eds.), *Wörterbücher*, II.1723–35.
- Wolski, W. 'Das Lemma und die verschiedenen Lemmatypen' in Hausmann et al. (eds.), *Wörterbücher*, I.360–71.
- Woods, C. E. 'Bilingualism, scribal learning, and the death of Sumerian' in S. L. Sanders (ed.), *Margins of Writing, Origins of Cultures* (Oriental Institute of the University of Chicago, 2006), 91–120.
- Wooldridge, T. R. *Les débuts de la lexicographie française: Estienne, Nicot, et le Thresor de la langue françoise (1606)* (University of Toronto Press, 1977).
- Wright, E. M. *The Life of Joseph Wright*, 2 vols. (London: Oxford University Press, 1932).

- Wu, J. 'Chinese-English dictionaries' in S. W. Chan and D. E. Pollard (eds.), *An Encyclopedia of Translation: Chinese-English, English-Chinese* (Hong Kong: Chinese University Press, 1995), 519–32.
- Wu Junxun 巫俊勳. 'Zihui' bianzuan lilun yanjiu 《字彙》編纂理論研究 (Taipei: Hua Mulan wenhua chubanshe, 2007).
- Xavier, Â. B., and I. G. Županov. *Catholic Orientalism: Portuguese Empire, Indian Knowledge (Sixteenth–Eighteenth Centuries)* (New Delhi: Oxford University Press, 2015).
- Xú Cháohuá 徐朝華. *Ēryǎ jīnshù 爾雅今注* (Tianjin: Nánkāi dàxué chūbānshè, 1994).
- Yamada Tadao. *Kindai Kokugo jisho no ayumi: Sono mohō to sōi to*, 2 vols. (Tokyo: Sanseidō, 1981).
- Yamada Tadao. *Sandai no jisho: Kokugo jisho hyakunen shōshi* (Tokyo: Sanseidō, 1967).
- Yamagiwa, J. K. *Bibliography of Japanese Encyclopedias and Dictionaries* (Ann Arbor: Panel on Far Eastern Language Institutes, 1968).
- Yamaguchi Akiho. 'Kogo jiten no hanashi' in Jiten kyōkai (ed.), *Nihon no jisho no ayumi*, 53–68.
- Yáñez Rosales, R. H. 'Presencia y ausencia de Antonio de Nebrija en vocabularios castellano-lengua indígena del siglo XVIII' in M. Mantilla Trolle and N. E. Jiménez Hernández (eds.), *Fondos del Tesoro: colección de lenguas indígenas Biblioteca Pública del Estado de Jalisco 'Juan José Arreola'* (Guadalajara: Universidad de Guadalajara and El Colegio de Michoacán, 2007), 45–63.
- Yang Gang 杨钢. 'Li Shi "Shuyu" lunlüe' 李实《蜀语》论略, *Tianfu xinlun*, 2 (1990), 85–7.
- Yang, H. 'The making of the first Chinese-English dictionary: Robert Morrison's *Dictionary of the Chinese Language in Three Parts* (1815–1823)', *Historiographia Linguistica*, 41.2/3, 299–322.
- Yeo, R. *Encyclopaedic Visions: Scientific Dictionaries and Enlightenment Culture* (Cambridge University Press, 2001).
- Yeşildal, S. 'Bir Türklük atlası, bir umman: Dîvân-ı Lügâti't Türk', *Kırmızı Beyaz*, 4 (2008), 9–13.
- 'Yiddish dictionaries', www.yivo.org/Yiddish-Dictionaries.
- Yiftach-Firanko, U. 'Law in Graeco-Roman Egypt: hellenization, fusion, Romanization' in Bagnall (ed.), *Oxford Handbook of Papyrology*, 541–60.
- Yıldız, H. 'Ahmet Vefik Paşanın *Lehçe-i Osmânî*'si ve Batı Kaynaklı Kelimeler', *Türk Dili. Dil ve Edebiyat Dergisi*, 634 (2004), 439–46.
- Yong, H., and J. Peng. *Bilingual Lexicography from a Communicative Perspective* (Amsterdam and Philadelphia: Benjamins, 2007).
- Yong, H., and J. Peng. *Chinese Lexicography: A History from 1046 BC to AD 1911* (Oxford University Press, 2008).
- Yoo Hyunkyung. 'Hangug-eosajeon pyeonchan-ui hyeonhwang-gwa ilonjeog jeongae', *Journal of Korealex*, 17 (2011), 186–208.
- You, R. 'Wú 吳 dialects' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, IV.602–7.
- Yuasa Shigeo. 'Edo jidai no jisho' in Nishizaki Tōru (ed.), *Nihon kojisho o manabu hito no tame ni* (Kyoto: Sekai shisōsha, 1995), 223–54.
- Zádrapa, L. 'Word and wordhood, premodern' in Sybesma et al. (eds.), *Encyclopedia of Chinese Language and Linguistics*, IV.549–54.

- Zastoupil, L., and M. Moir (eds.). *The Great Indian Education Debate: Documents Relating to the Orientalist-Anglicist Controversy, 1781–1843* (Richmond, UK: Curzon, 1999).
- Zaunmüller, W. *Bibliographisches Handbuch der Sprachwörterbücher: ein internationales Verzeichnis von 5600 Wörterbüchern der Jahre 1460–1958 für mehr als 500 Sprachen und Dialekte* (New York and London: Hafner; Stuttgart: Hiersemann, 1958).
- Zemzare, D. *Latviešu vārdnīcas* (Riga: Latvija PSR Zinātņu Akademijas Izdevniecība, 1961).
- Zenker, J. T. *Bibliotheca orientalis: manuel de bibliographie orientale*, 2 vols. (Leipzig: Engelmann, 1846–61).
- Zetzel, J. E. G. *Latin Textual Criticism in Antiquity*. New York: Arno, 1981.
- Zgusta, L. 'Copying in lexicography: Monier-Williams' Sanskrit dictionary and other cases (dvaikošyam)', *Lexicographica*, 4 (1988), 145–64.
- Zgusta, L. 'The influence of scripts and morphological language types on the structure of dictionaries' in Hausmann et al. (eds.), *Wörterbücher*, 1.296–305.
- Zgusta, L. *Lexicography Then and Now: Selected Essays*, ed. F. S. F. Dolezal and T. B. I. Creamer (Tübingen: Niemeyer, 2006).
- Zhang Zongxiang 張宗祥. Wang Anshi 'Zishuo' ji 王安石《字說》輯 (Fuzhou: Fujian renmin chubanshe, 2005).
- Zhōu Zǔmó 周祖謨. *Tang Wudai yunshu jicun* 唐五代韻書集存 (Beijing: Zhōnghuá shūjù, 1983).
- Zhōu Zǔmó 周祖謨. *Wén xué ji* 問學集, 2 vols. (Beijing: Zhōnghuá shūjù, 1966).
- Zide, N. 'Lexicography of other languages of the Indian subcontinent: the Munda languages' in Hausmann et al. (eds.), *Wörterbücher*, III.2533–47.
- Zinkevičius, Z. *Lietuvių kalbos istorija*, vol. IV: *Lietuvių kalba XVIII–IX a.* (Vilnius: Mokslas, 1990).
- Zolli, P. 'Italianisch: Lexikographie' in Holtus et al. (eds.), *Lexikon der romanistischen Linguistik*, IV.786–98.
- Zulay, M. *ha-ʾaskola ha-payṭanit šel Rav Saʿadya Gaʾon* (Jerusalem: Schocken Institute for Jewish Research of the Jewish Theological Seminary of America, 1964).
- Županov, I. G. 'Aristocratic analogies and Demotic descriptions in the seventeenth-century Madurai Mission', *Representations*, 41 (1993), 123–48.
- Županov, I. G. 'Compromise: India' in R. P.-C. Hsia (ed.), *A Companion to the Reformation World* (Oxford and Malden: Blackwell, 2004), 353–74.
- Županov, I. G. *Disputed Mission: Jesuit Experiments and Brahmanical Knowledge in Seventeenth-Century India* (New Delhi: Oxford University Press, 2001).
- Županov, I. G. 'Goan Brahmins in the Land of Promise: missionaries, spies, and gentiles in the 17th–18th century Sri Lanka', in J. Flores (ed.), *Portugal–Sri Lanka: 500 Years* (Wiesbaden: Harrassowitz and the Calouste Gulbenkian Foundation, 2006), 171–210.
- Županov, I. G. 'Language and culture of the Jesuit "early modernity" in India during the sixteenth century', *Itinerario*, 31.2 (2007), 87–110.
- Županov, I. G. 'Orientalist museum: Roman missionary collections and prints', in I. Banerjee-Dube and S. Dube (eds.), *From Ancient to Modern: Religion, Power, and Community in India* (New Delhi: Oxford University Press, 2009), 207–35.
- Županov, I. G. 'Prosélytisme et pluralisme religieux: deux expériences missionnaires en Inde aux XVIe et XVIIe siècles', *Archives des sciences sociales des religions*, 87 (1994), 35–56.
- Zvelebil, K. V. *Lexicon of Tamil Literature* (Leiden: Brill, 1995).

- Zwartjes, O. 'The historiography of missionary linguistics: present state and further research opportunities', *Historiographia Linguistica*, 39.2/3 (2012), 185–242.
- Zwartjes, O. *Portuguese Missionary Grammars in Asia, Africa, and Brazil, 1550–1800* (Amsterdam and Philadelphia: Benjamins, 2011).
- Zwartjes, O. 'Some remarks on missionary linguistic documentation and lexicography: Diaethnical markedness in Alexandre de Rhodes' *Dictionarium Annamiticum* (1651)' in C. Nunes de Almeida and M. Pacheco Pinto (eds.), *O Oriente em Tradução/The Orient in Translation* (Vila Nova de Famalicão: Húmus, 2017), 213–50.
- Zwartjes, O. 'El Vocabulario de letra china de Francisco Díaz (ca. 1643) y la lexicografía hispano-asiática', *Boletín Hispánico Helvético*, 23 (2014), 57–100.
- Zwartjes, O., and C. Altman (eds.). *Missionary Linguistics II/Lingüística misionera II: Orthography and Phonology* (Amsterdam and Philadelphia: Benjamins, 2005).
- Zwartjes, O., R. Arzápalo Marín, and T. C. Smith-Stark (eds.). *Missionary Linguistics IV/Lingüística misionera IV: Lexicography* (Amsterdam and Philadelphia: Benjamins, 2009).
- Zwartjes, O., and E. Hovdhaugen (eds.). *Missionary Linguistics/Lingüística misionera: Selected Papers from the First International Conference on Missionary Linguistics, Oslo, 13–16 March 2003* (Amsterdam and Philadelphia: Benjamins, 2004).

Index

- Aa* (Mesopotamian sign list), 31, 34
 Aasen, Ivar, 476, 738
Abba–Ababus, 270, 273
 ‘Abd-al-laṭīf ibnī Melek, 242, 738
 ‘Abd-al-Rashīd, 234, 739; *see also* *Farhang-i Rashīdī*
 Abdel-Nour, Jabbour, 425, 739
 Abenaki language, 599, 706
Abhidhānappadīpikā, 77, 78, 143–5
 Abhimānacihna, 141
 Abramović, Teodor, 730, 739
 abridged dictionaries,
 Arabic, 174, 423, 425, 429
 Chinese, 204, 214
 English, 308, 490–1, 498
 French, 534, 535
 Greek, 96, 99, 251, 257, 263, 297, 298
 Hebrew, 188
 Italian, 538
 Japanese, 619
 Korean, 220
 Latin, 90, 269, 271, 272, 275,
 284, 286
 Persian, 385
 Portuguese, 545
 Scots, 514
 Spanish, 541
 Tibetan, 147
Abrogans, 281
 Abū ‘Amr, 172
 Abū l-Barakāt ibn Kabar, 49, 739
 Abū l-Faraj Harūn, 201
 Abū Ḥatīm al-Siğṣṡānī, 163
 Abū Ja’far al-Bayhaqī, 231, 739
 Abū Miṣṣhal al-A ‘rābī, 164, 739
Abuṣṣka Liḡati, 240
 Abū l-Ṭayyib al-Luḡawī, 169, 739
 Abū ‘Ubayd, 163–5, 170, 739
 Abū Zayd al-Anṣārī, 160, 164, 739
 academies as producers of dictionaries, 304–5,
 311, 313, 418, 428, 433–4, 437, 451, 453–4, 461,
 466–7, 472–3, 474, 481, 486, 487–8, 531, 541,
 543, 545–6, 548–9, 551–2
 Accius, 90
 Achagua language, 556, 706
 Adam von Rottwil, 299
 Addison, Joseph, 486, 489, 517
 Addy, Sidney Oldall, 512–13, 739
 Adelung, Johann Christoph, 462–4, 466, 468,
 469, 470, 739
 Adīb Naṭanzī, 231, 739
 Adler, Ada, 254, 258
 Aelius Herodianus. *See* Herodian
 Aelius Stilo, 90–1
 Aeschylus, 257
 Afghānīnavīs, ‘Abdullāh, 387
 Afranius, 90, 91
 Afrikaans language, 480–1, 528, 679, 706
 Afroasiatic languages, 706
 Aggavaṃsa, 76, 144, 739
 Ahom language, 404, 706
 Aitken, Adam Jack, 514, 739
 Ajayapāla, 134, 139, 141, 739
Akarāti Nikanṭu, 149, 153, 155, 645
 Akkadian language, 11–35, 40
 Aktunç, Hülki, 375, 739
 Albanian language, 549, 707
 Albert, Roy, 604
 Alcalá, Pedro de, 302, 580, 621, 739
 Alcalay, Reuben, 439–40, 740
 Alcedo, Antonio de, 544, 740
 Aldhelm, 270, 280, 740
 Alévêque, Charles, 352
 Alexander of Masowia, 446
 Alfāsi, David ben Abraham, 188–9, 191, 193,
 198, 200
 Algaze, Guillermo, 16
 Algonquian languages, 527, 599–600, 707

- Algonquin language, 599, 707
 Ali Emîrî, 227
 'Alî Şîr Nevâî, 240
 Allentiac language, 556, 707
 Allon, Mark, 76
 Allsopp, Richard, 525–6, 740
 Alonso, Juan, 582, 740
 Alpers, Klaus, 248, 255, 257–8, 264
 alphabetization. *See* macrostructure, alphabetical or syllabic
 Alvarado, Francisco de, 581, 584–5, 592, 595, 740
 Alves, Albino, 663–4, 680, 740
 Amaral, Gaspar do, 361
 Amarasimha and *Amarakośa*, 68, 71–2, 131–2, 133–4, 135, 136–9, 140–1, 142, 143–4, 145–8, 390–1, 394, 396, 400, 643, 645, 650, 740
 Amaya, Diego, 563
 Amenemope, 39, 740
 Amharic language, 442
 Ammonius, 94, 263
 An = *Anum* (Mesopotamian wordlist), 29, 31, 34
 Anatolian languages, 95
 Anchieta, Joseph de, 558
 ancient South Arabian languages, 45, 707
 Andaqui language, 556, 707
 Andrews, Richard, 589
 Ange de Saint Joseph, 235, 299, 740
 Anglo-Norman language, 282–3, 286, 536, 707
 Anindilyakwa language, 704, 707
 Ankurac'i, Yovhannes. *See* Molino
 Annan Yiyu, 357
 Anshel, 736, 740
 Ansileubus. *See* *Liber glossarum*
 Anthony, Albert Seqaqkind, 603
 anthropology, 610, 671, 684, 685, 692, 696, 697, 743, 745, 748, 755, 764, 771, 776, 787, 798
 Antoine, Francesca, 604
 antonyms, in dictionary entries, 53, 181, 470, 479, 568, 593, 666
 Apachean languages, 606
 Apion, 92, 103, 740
 Apollonius Sophista, 92, 740
 Apple, 508
 Ara, Domingo de, 582, 586, 740
 Arabana language, 692, 707
 Arabic language, 36, 41, 48–50, 143, 159–81, 184–5, 187, 188–94, 195–6, 198–9, 200–1, 223, 226–8, 229–32, 233, 234–5, 238–40, 241, 284, 287, 299, 302, 369–74, 380–1, 385, 391–4, 398, 410–11, 414–30, 433, 570, 580, 621, 649, 669, 670, 707
 Egyptian Arabic, 179, 181, 239, 426, 427
 Judeo-Arabic, 185, 194
 Lebanese Arabic, 427
 Modern Standard Arabic, 414, 417, 420, 422, 426–7
 Moroccan Arabic, 424, 426–7
 Spanish Arabic, 420
 Aragona, Alonso de, 575, 740
 Aramaic languages, 41, 95, 183–4, 188–9, 190, 195–6, 199, 224, 232, 299, 433, 434, 707
 Arawakan languages, 559, 560, 577, 708
 archaic words
 Arabic, 160, 418
 Chinese, 54, 58, 322
 Dutch, 479
 Egyptian, 39, 40
 English, 406, 490, 504, 511–12, 516
 German, 461–2, 470
 Greek, 88, 94, 98, 260
 Guarani, 568
 Japanese, 216, 344, 346
 Korean, 355
 Latin, 89–90
 Latvian, 459
 Mesoamerican languages, 583
 Mesopotamian, 15, 19, 27, 29, 34
 Persian, 224
 Sanskrit, 69, 136
 South Asian languages, 413
 Spanish, 541, 542
 Tamil, 402
 Turkish, 374, 376
 Archilochos, 104
 Arenas, Pedro de, 580, 593, 741
 Argenti, Filippo, 243, 741
 Aristophanes of Byzantium, 87–9, 100, 741
 Aristophanes the dramatist, 86, 97, 98, 257, 262
 Aristotle, 95, 276
 Armenian language, 180, 244, 708
 Arrernte language(s), 685, 693, 708
 Arroyo de la Cuesta, Felipe, 598, 741
Arte y vocabulario de la lengua morocosi, 563, 568
Arte y vocabulario en la lengua general del Peru, 555, 561, 562, 563, 566
 Arts, Tressy, 417, 422, 741
 Ārzū, Sirāj al-Dīn 'Alī Khān, 378–9, 381, 741; *see also* *Sirāj al-Lughat*
 Asadi Ṭūsī, 224–5, 228–30, 741
 al-'Askarī, 168, 741
 Aslanov, M. G., 412
 al-Aşma'ī, 161, 741
 Assamese language, 402–4, 708

- Assumpçam, Manoel da, 634, 638, 641, 647, 652, 655, 741
- Athabaskan languages, 601, 606, 611, 708
- Athanasius of Qus, 50, 741
- Aubery, Joseph, 599, 741
- Auden, W. H., 502
- Augustinian missionaries as lexicographers, 582, 599, 614, 634, 741, 743, 744, 777, 786, 800
- Aulete, Francisco Júlio de Caldas, 545–6, 741
- Aurelius Opillus, 91
- Austen, Jane, 502
- Austin, Peter, 687
- Australian National Dictionary*, 527–8
- Avadhī language, 410
- Avestan language, 224, 708
- Avineri, Yitshaq, 184
- Avis, Walter Spencer, 522, 524, 741
- Avneyon, Eitan, 437, 443, 741
- Awabakal language, 685, 689–90, 708
- Aymara language, 555–6, 560–2, 563–4, 565–7, 569, 572, 574–5, 708
- ʿAynī, Ṣadr al-Dīn, 387
- Azar, Moshe, 435, 741
- al-Azharī, 162, 176, 200, 742
- Baalbaki, Mounir, 425, 742
- Baalbaki, Ramzi, 173, 425, 742
- Baalbaki, Ruhi, 417, 425, 429, 742
- Bacon, Francis, 489
- Badawi, El-Said, 426, 429, 742
- Badr-al-Dīn Ibrāhīm, 233, 392, 742
- Baevskij, Solomon I., 233
- Bahār, ʿĪk Chand, 379, 742; *see also Bahār-i ʿĀjam*
- Bahār-i ʿĀjam*, 379, 383–4, 385
- Bahat, Shoshana, 436
- Bailey, Nathan, 488, 490, 742
- Balbi, Giovanni, 268, 274–5, 279, 282, 286, 288–9, 290, 292–3, 742
- Balčikonis, Juozas, 457
- Baldauf, Ingeborg, 236, 237
- Baldinger, Kurt, 536, 742
- Baldwin, Caleb Cook, 330, 742
- Baltic languages, 445, 454–5, 456, 708
- Bambara language, 665, 708
- al-Bandanīgī, 173–4, 175, 179, 742
- Bandini, Simon, 563
- Banks, Joseph, 683, 742
- Bantu languages, 670, 672, 673, 681, 708
- Baraga, Frederic, 601, 742
- Barbier de Meynard, Charles Adrien Casimir, 372, 742
- Barbosa, Antonio, 361
- Bardhi, Frang, 707, 742
- Bardi language, 683, 697, 699, 701–2, 709
- Baretto, Joseph, 394
- Barkhudarov, A. S., 412
- Barlement, Noel van, 300–1, 743
- Barngarla language, 687, 691, 709
- Barthélemy, Jean-Jacques, 396
- Bartholomew, Doris, 600
- Bartlett, John Russell, 518, 743
- Barua, Jaduram Deka, 402, 766
- Barua, Rai Sahib Golap Chandra, 404
- Baruwā, Hemchandra, 403, 743
- Barzizza, Gasparino, 285, 743
- Basalencque, Diego, 569, 583, 584, 587, 590–1, 743
- Basilio, Tomás, 600, 743
- Basque language, 287, 709
- al-Baṭalyawṣī, 169, 743
- Bates, Daisy May, 684, 694–5, 743
- Bathe, William, 295, 743
- Battaglia, Salvatore, 539, 743
- Baudouin de Courtenay, Jan, 452
- Beal, Joan, 515
- Bedersi, Abraham ben Isaac, 197, 743
- Bēixī zīyī, 125–6
- Bélisle, Louis-Alexandre, 537, 743
- Bellini, Bernardo, 538, 539, 540, 743
- Belostenec, Ivan, 448, 743
- Belsare, M. B., 405
- Benedictines as lexicographers, 265, 273, 285, 693, 695, 700–1, 757, 773, 785, 790–1, 792
- Ben Elʿazar, Yaʿakov, 192
- Benešovský, Matouš, 447, 744
- Bengali language, 403, 634, 641, 647, 651–2, 655, 709
- Benson, T. G., 672
- Bentley, William Holman, 660–2, 744
- Ben-Yehuda, Eliezer, 199, 432–5, 437–9, 744
- Benzebra, Radia, 425
- Bergaño, Diego de, 618, 744
- Bernolák, Anton, 451, 744
- Bertonio, Ludovico, 555, 560–2, 563, 565–7, 568, 572, 574–5, 744
- Berynda, Pamva, 448, 744
- Bescherelle, Louis-Nicolas, 542
- Beschi, Costanzo Giuseppe, 149, 153, 154–5, 158, 640, 644, 649, 651, 655–6, 744
- Beskrovniĭ, Vasilij Matveevič, 412, 744
- Bethel, Rosalie, 604
- Bezmez, Serap, 375
- Bhāguri, 71
- Bhānuji Dīkṣita, 142

- Bhojpurī language, 410
 Bianchi, Thomas Xavier, 371, 744
 Bibbesworth, Walter de, 283, 744
 Bible
 in Afrikaans, 481
 in Arabic, 49, 184
 in Aramaic, 190, 196, 199, 299
 in Armenian, 244
 in Chinese, 328
 in Choctaw, 603
 in Coptic, 44, 46, 49
 in English, 489, 497
 in Gothic, 715
 in Greek, 47, 247, 250–1, 255, 264, 298
 in Hebrew, 182, 183, 184–8, 189–92, 194–6,
 197, 199, 200, 201, 277, 284, 299, 667
 in Kikongo, 662
 in Korean, 351
 in languages of New England, 602
 in Latin, 269–70, 273, 277, 279, 284
 in Lithuanian, 455
 in Muskogee, 603
 in Russian, 452
 in Welsh, 307
 in Western Apache, 612
 in Worrorra, 696
 in Xhosa, 678
 Bicol languages, 618, 709
 Bignon, Jean-Paul, 653
 Bilodid, Ivan Kostjantynovyč, 454, 744
 Biraban, 689, 744
 Bird, William, 695
Birds (Mesopotamian wordlist), 17
 Birwé, Robert, 72
 Bischofs, Joseph, 695, 745
 Blas Pretovio. *See* Pablo Restivo
 Blau, Joshua, 191
 Bleek, Wilhelm Heinrich Immanuel, 672, 674,
 680–1, 745
 Blemmydes, Nikephoros, 255, 745
 Bloch, Oscar, 535, 745
 Blochmann, Henry, 383–4, 385
 Bluteau, Rafael, 545, 745
 Bobowski, Wojciech, 245
 Boccaccio, 304
 Bodo language, 404, 709
 Bohorič, Adam, 448, 450
 Böhrling, Otto von, 399, 745
 Bonelli, Luigi, 375, 745
 Bonet, Jean Pierre Joseph, 362, 745
 Bonvini, Emilio, 665
 Bopp, Franz, 492
 Bosnian language, 241, 454, 709
 Boucher, Jonathan, 516, 745
 Bourke, John Gregory, 610, 745
 Bourzès, Louis Noël de, 655
 Bowrey, Thomas, 642, 745
 Bradley, Henry, 494, 745
 Brahmanism. *See* Hinduism and Brahmanism
 Braj language, 398, 410, 709
 Brambila, David, 601
 Branner, David Prager, 330
 Brasche, Gustav, 458
 Bray, Dorothy, 612–13
 Brender, Franz, 457, 745
 Breton, Peter, 411
 Breton, Raymond, 559, 560, 569, 577, 745
 Breton language, 278–9, 709
 Briggs, Keith, 515
 Bright, William Oliver, 522, 745
 Brinton, Daniel Garrison, 603
 Brinton, Laurel, 523
 Brito, Guillelmus, 268, 273, 282, 287, 746
 Brockhaus dictionaries, 386, 471
 Brodowski, Jacob, 456, 746
 Brollo, Basilio, 327, 746
 Bronson, Miles, 402–3, 746
 Brooks, Maureen, 528
 Brown, C. H., 375
 Brown, Charles Philip, 392–3, 400, 746
 Brugiotti da Vetralla, Giacinto, 660
 Bruno, Ignaccio, 154
 Bruyas, Jacques, 599, 746
 Bryant, Alfred Thomas, 677, 746
 Buarque de Holanda Ferreira, Aurélio,
 547, 746
 Buchi, Éva, 532, 552
 Buckwalter, Tim, 422–3, 429
 Buddhaghosa, 76–7, 79
 Buddhism, 63, 65, 72–6, 77, 112, 125, 127–8,
 130–1, 134–5, 137, 139, 140, 143–4, 145–7, 148,
 207, 208, 218–19, 358, 619, 627, 632, 770, 793
 Budé, Guillaume, 297, 746
 Buechel, Eugene, 601, 746
 Būga, Kazimieras, 457, 746
 Bugarin, José, 618, 746
 Bühler, Georg, 142
 Bulgarian language, 445–6, 450, 709
 Bunuban languages, 699, 709
 Burarra-Gun-Narpta language, 704, 709
 Burchfield, Robert William, 500–2, 505, 746
Burhān-i Qāṭi‘, 233, 377–8, 381–4, 386
 Burke, Scott, 524
 Burmese language, 641
 Burnell, Arthur Coke, 527, 747
 Busbecq, Ogier Ghiselin de, 715, 747

- al-Bustānī, Buṭrus, 417–18, 429, 747
 Buxtorf, Johannes, the elder, 299, 747
 Buxtorf, Johannes, the younger, 299, 747
 Bydgoszczy, Bartłomiej, 447, 747
 Byington, Cyrus, 603, 747
 Byron, George Gordon, 491
- Cài Biàn, 124, 747
 Calepino, Ambrogio, 7, 292–4, 297, 300–1, 306, 309, 447, 565, 585, 619, 623, 624–5, 626, 628, 747
 Callary, Edward, 522
 Callimachus of Cyrene, 87, 747
 Calvo Pérez, Julio, 556, 563
 Camerarius, Joachim, 265
 Campbell, A. D., 401
 Campbell, William, 332, 747
 Candrea, Ioan Aurel, 549, 747
 Candy, George, 406, 747
 Candy, Thomas, 406, 747
Cāngjié piān, 52–3, 63
 Canicatti, Bernardo Maria da, 660, 662, 748
 Cantemir, Dimitrie, 547, 748
 Capell, Arthur, 685, 697, 748
 Capuchin missionaries as lexicographers, 639, 651, 660, 748, 788, 790, 797, 798, 802
 Cardoso, Jerónimo, 302, 624–5, 748
 Cardoso, Mattheus, 660
 Carey, William, 406, 748
 Carib language, 556, 559, 560, 710
 Cariban languages, 559, 710
 Carmelite missionaries as lexicographers, 639, 645, 649, 740, 783, 795, 802
 Carochi, Horacio, 587
 Carradori, Arcangelo, 244, 748
 Carrier language, 601, 604, 710
 Caspar, Marie-Antoine-Louis, 362
 Cassidy, Frederic Gomes, 509, 520–2, 525–6, 748
 Castell, Edmund, 235, 383–4, 385, 748
 Castillo, Martín del, 563
 Castro, Andrés de, 581, 748
 Cat, Simon, 641, 647, 748
 Catalan language, 550–1, 710
 Cauzim, Mirza Mahomed, 405
 Cawdrey, Robert, 308, 748
 Cebuano language, 629, 710
 Celtic languages, 277–9, 668, 710
 Cēṇāvaraiyar, 81
 censorship, 452, 453, 459, 647, 651–2
 Chabata, Emmanuel, 674, 675
 Chagatay language, 238, 240, 372, 710
 Chaliha, Debesvar, 403
 Champollion, Jean-François, 41
 Changuion, Antoine Nicholas Ernest, 480, 749
 Chao, Eduardo, 542, 749
 Chao, Yuen Ren, 325, 749
 Chatelain, Héli, 663
 Chaucer, Geoffrey, 308, 497
 Chekayri, Abdellah, 426
 Chén Chún, 125, 749
 Chén Dì, 116, 749
 Chén Péngnián, 62, 114, 749
 Chén Shiyuán, 126, 749
 Chen Xiang, 413
 Chevillard, Jean-Luc, 648
 Cheyenne language, 603, 710
 Chibcha language, 556, 710
 children, language of, 525, 627
 Chindhundu, Herbert, 675, 680–1
Chi nam ngọc âm giải nghĩa, 357–8, 360
 Chinese language, 51–66, 109–29, 202–22, 302, 317–39, 340–65, 412–13, 501, 614–17, 619–20, 622, 626–8, 632–3, 710–11
 Chiquitano language, 569, 711
 Chirino, Pedro, 326, 615, 616, 749
 Choctaw language, 603, 711
 Choeroboscus, George, 264, 749
 Chökê language, 388
 Cholinus, Petrus, 309, 749
 Chos skyong bzang po, 147, 749
 Choueka, Yaacov, 436, 442, 749
 Church Missionary Society (later Church Mission Society), 666–7, 679, 690, 761, 762, 769, 791
 Church Slavic language, 445–6, 448, 450, 451, 547, 711
 Cicero, 96, 294–5
 Cigale, Matej, 450, 749
 Cihac, Alexandru, 549, 749
Cíhǎi, 322–3, 324, 325
 Cinyanja language, 673, 711
 Cioranescu, Alejandro, 549, 750
 Ciudad Real, Antonio de, 6, 582, 586, 750
Cíyuán, 322–3, 324, 325
 Claretus de Solentia, 288, 446, 750
 Cloatius, 91
 Clodius, Johann Christian, 367, 750
 Coate, Howard, 688, 697, 702–3, 750
 Coblin, W. South, 328, 624
 Codex Comanicus, 236–7, 289
 Coeurdoux, Gaston-Laurent, 396–7, 401
 Cohen de Lara, David, 196
 Colebrooke, Henry Thomas, 394, 750
 Colenso, John William, 676, 750
 Collado, Diego, 622, 626, 750

- Collard, Elisabeth Scott, 604
 Collard, Howard, 604
 Collins Cobuild dictionaries, 503–5
 colloquial words, 56, 58, 98, 126–7, 215, 334, 370, 435, 440, 444, 499, 500, 504, 679
 Comenius, Johannes Amos, 295–6, 298, 309, 312, 314, 750
 Cometas, 257
 commentaries and lexicography
 Arabic, 159, 167, 174, 196
 Chinese, 56, 58, 60, 64, 113, 116, 118–20, 122, 124–5, 129, 348
 Greek, 86, 254, 257, 264, 297
 Hebrew, 184–6, 189, 195, 196, 197, 199, 200
 Latin, 268, 292
 Mesopotamian, 29, 34
 Pali, 74, 76–7, 78, 143
 Persian, 377, 380
 Prakrit, 74
 Sanskrit, 69–70, 71, 133–4, 136–7, 140–1, 142, 143, 145, 146–8, 395
 Tamil, 81–2, 156
 Tibetan, 147
 compounds
 Chinese, 66, 125, 213, 324, 347–8
 Dutch, 479
 English, 528
 German, 470
 Greek, 87, 96, 100
 Guarani, 570
 Japanese, 347–8, 619, 630
 Mapudungun, 573
 Mesopotamian, 25, 34
 Persian, 225
 Sanskrit, 132
 Tagalog, 618, 628
 computational methods in lexicography, 6, 426, 436, 471, 479–80, 605, 687; *see also* corpus lexicography
Concise Oxford Dictionary, 399, 402, 498, 504, 508
Concise Scots Dictionary, 512, 514, 529
 Confucius and Confucianism, 53, 116, 125–6, 206, 216, 217–18, 222, 318, 321, 323, 358, 360
 Congregatio de Propaganda Fide, 622, 642, 651, 653, 660
 Congregationalist missionaries as lexicographers, 329, 331, 603, 689, 692, 796
 Constantin, Robert, 297, 750
 Constantine Porphyrogenetos, 254, 263
 Cook, James, 529, 683, 750
 Cooper, Thomas, 306, 750
 Coptic language, 36, 46–50, 299, 711
 Córdova, Juan de, 561, 581, 584, 585, 591–2, 595, 750
 Cormac mac Cuileinnáin, 277, 750
 Corneille, Thomas, 305, 751
 Cornish language, 278, 711
 Coromines i Vigneaux, Joan, 544, 551, 751
 corpus lexicography
 Arabic, 417, 423, 429–30
 Danish, 474
 English, 487, 498, 502–5, 506
 French, 536
 German, 471
 Hebrew, 438, 444
 Korean, 355
 Norwegian, 476
 Portuguese, 547
 Romance languages, 532
 Shona, 676
 Spanish, 541, 543, 544–5
 Swedish, 475
 Corriente Córdoba, Federico, 424, 751
 Cortés y Zedeño, Jerónimo Tomás de Aquino, 583, 751
 Coste, Eugène-Jean-Georges, 350
 Cotgrave, Randle, 304, 307, 751
 Cotton, Josiah, 602, 751
 Courtois, Victor-Joseph, 673–4, 751
 Covarrubias Horozco, Sebastián de, 302, 305, 314, 541, 751
 Cowan, J Milton, 422, 751
 Cox, Barrie, 515
 Craigie, William Alexander, 494, 496, 513–14, 519–20, 522, 751
 Crastoni, Giovanni, 296, 751
 Cree language, 601, 711
 creoles, 525, 551, 642, 663, 682, 704
 Crisostomo, Jay, 25
 Croatian language, 241, 288, 447–9, 450, 454, 712
 Crooke, William, 527, 751
 cross-references in dictionaries, 64, 309, 347, 436, 467, 513, 620
 Crowther, Samuel Ajayi, 666–9, 680–1, 751
 Cua, Paulus, 363, 752
 Cubero, Fulgencio, 695, 700
 Cuervo, Rufino José, 544, 752
 Cūla Moggallāna, 143–4, 752; *see also* *Abhidhānappāḍipikā*
 Cumanagoto language, 559, 561, 562, 567, 570–1, 712
 Cuman language, 236–7, 712
 Cunha, Antônio Geraldo da, 547
 Cunningham, Ian, 248, 252

- Cuoq, Jean-André, 599, 752
 Curr, Edward Micklethwaite, 684, 752
 Curtis, Edward Sheriff, 609, 610
Cūṭamaṇi Nikāṇṭu, 149, 152–3,
 155–6, 644
 Cyril of Alexandria, 96, 249–51, 252, 259,
 276, 752
 Czech language, 287–8, 300, 309, 417, 446–8,
 449, 451, 452, 712
- Dabīr Siyāqī, Muḥammad, 386
 da Costa, Balthasar, 652
 Dagut, Menachem, 440
 Dài Tóng, 119–20, 752
Daigenkai, 344–5
Daijirin, 345, 349
Daijisen, 345, 349
Daijiten (1917), 347–8
Daijiten (1934–6), 344
 Dakota language, 603–4, 712
 Dal', Vladimir Ivanovič, 452, 752
 d'Alberti di Villanuova, Francesco, 538,
 752
 Dalby, Andrew, 1
 d'Alembert, Jean le Rond, 531, 534
 Dampier, William, 683
 Daniil of Moscopole, 450, 752
 Danish language, 284, 291, 311–12, 472, 474,
 475–7, 482, 602, 712
 Dankoff, Robert, 226
 Dante, 274, 304
 Đào Duy Anh, 364, 752
 Daoism, 63, 125
 Dard, Jean, 665, 667, 752
 Dareau, Margaret, 514
 Darmesteter, Arsène, 534, 752
 Darviš Muḥammad, 240
 Dās, Śyām Sundar, 410, 753
 Dasypodius, Petrus, 309–10, 753
 Dave, Narmada Śankara, 405, 753
 Daveluy, Marie-Nicolas-Antoine, 351, 753
 David ben Yesha^c, 196
 Davis, William Jafferd, 678, 753
 Dawes, William, 683, 753
 al-Dawlah, Muḥammad Ḥasan Ṣanī', 384
 definitions
 based on quotation evidence, 488, 491
 extensive, 63, 125, 185, 344, 346, 562, 577, 689,
 691, 693, 699, 702–3
 high quality of, 305, 345, 420, 469, 475, 479,
 488, 518, 543, 678
 not always obligatory, 60, 104, 116, 181, 198,
 212, 466, 467
 sentence-style, 503–4
 unsatisfactory, 177, 181, 466, 485, 496,
 507, 512
- Defoe, Daniel, 486
 DeFrancis, John, 331, 336–7, 753
 Delaware languages, 603, 712
 del Canto, Francisco, 561
 D'Elia, Pasquale, 616
 Della Bella, Ardelio, 448, 753
 Delsaerd, Pierre, 649
 De Mauro, Tullio, 540, 753
 Demosthenes, 97
 Densușianu, Ovid, 549, 753
 de Reuse, Willem, 611
 descriptivism, 334, 410, 432, 435–6, 463–4, 479,
 485, 487, 496, 499, 501–2, 503, 507, 533, 537,
 539, 542, 543, 544, 548, 550, 632, 679
 Deshpande, B., 399
 Deshpande, P. G., 399
 design of dictionaries. *See also* headwords:
 script used for
 affixes in margin, 588
 affixes in separate columns, 590
 headwords in right-hand column, 559
 miscellaneous observations in separate
 column, 672
 multiple typefaces, 649
 new line for each entry, 85, 101
 page size, 215, 271, 280, 281, 294–5, 297–8,
 300–1, 302, 304–5, 308, 309, 311, 313, 334, 352,
 385, 399, 424, 436, 438–40, 490, 498,
 548–9, 575
 size of lettering, 372
 special symbols, 375, 459, 461, 496, 501, 520,
 568, 588, 592, 628–9, 673
 text in continuous block, 85, 94, 101, 225
 text in multiple columns, 14, 15, 215, 300,
 395, 670, 672, 693
 text in rows and columns, 615, 616
 text in three columns, 15, 237, 244, 271, 372,
 649, 702, 703
 text in two columns, 14, 28, 29, 85, 86, 94,
 96, 237, 298, 308, 326, 369, 524, 558, 617,
 620, 668, 670, 678
 two-page spreads, 300, 649
 use of coloured inks, 44, 85, 225, 243
- Deuri language, 404, 713
Deutsches Wörterbuch, 465–7, 469, 475,
 478–9, 493
 Devarāja, 141
 de Vries, Matthias, 479, 753
 Dhadphale, M. G., 75
 Dhanapāla, 141–2, 753

- Dharanidāsa, 140, 753
 dialects and regionalisms
 Arabic, 160, 164, 168, 175, 414–15, 426–7
 Aymara, 566
 Chinese, 54–5, 57, 59, 61, 110, 115, 126–7,
 318–20, 322, 328–30
 Croatian, 448–9
 Danish, 312
 Dutch, 310, 477, 479
 English, 484–5, 496, 501, 508, 509–29, 602
 Francoprovençal, 550
 French, 287, 537, 552
 Frisian, 481–2
 German, 461, 463, 470
 Greek, 87, 89, 92, 95, 96, 99, 251–2
 Hebrew, 185
 Hindi, 410
 Ilocano, 629
 Italian, 284–5, 539
 Japanese, 215, 216, 343, 627
 Kikongo, 660–1
 Korean, 355
 Latin, 237
 Latvian, 458–9
 Lithuanian, 456, 457
 Lower Sorbian, 454
 Luxemburgish, 483
 Mapudungun, 567
 Marathi, 406
 Mesoamerican languages, 583, 592
 Nahuatl, 580
 Norwegian, 475–6
 Occitan, 550
 Oromo, 670
 Persian, 224, 379, 384, 387
 Polish, 452
 Portuguese, 545, 552
 Prakrit, 74, 78–9
 Quechua, 556, 564, 566
 Romance languages, 530–1, 532
 Russian, 452
 Shona, 674–5
 Slavic languages, 447
 Slovak, 451
 South Asian languages, 388
 Spanish, 541, 542–4, 552
 Sumerian, 15
 Swahili, 670
 Swedish, 472
 Tagalog, 628
 Tamil, 153, 157
 Turkic languages, 226–7, 228–30, 238–40
 Turkish, 242, 370, 374
 Urdu, 410
 Visayan, 629
 Western Apache, 613
 Western Desert Language, 702
 Wolof, 666
 Xhosa, 679
 dialogues, 243, 300, 367, 368
 Díaz, Francisco, 616, 617, 620, 622, 626, 631, 753
Diccionario da lingua portuguesa, 546, 551
Diccionario de autoridades (formerly
 Diccionario de le lengua castellana), 541,
 542, 543
 Dickens, Charles, 497, 502
 Dickey, Eleanor, 86
 dictionaries, judgements on
 ‘Almost every word is mistranscribed’, 700
 ‘boring’, 466
 ‘entertainment in perpetuity’, 37
 ‘entirely a compilation of entries in earlier
 dictionaries’, 320
 ‘Great Tzotzil Disaster, The’, 7
 ‘history of thought and civilization’, 495
 ‘huge, avant-garde, and influential’, 273
 ‘Listenwissenschaft’, 11
 ‘living scriptures, the’, 502
 ‘more like the author’s notebook than a
 formal alphabetized dictionary’, 562
 ‘most comprehensive and
 sophisticated’, 297
 ‘most popular [in the world]’, 334
 ‘most terrible lexicon ever’, 250
 ‘no-one will read it for pleasure’, 248
 ‘one of the finest ever written in any
 language’, 420
 ‘rather uneven’, 474
 ‘surely not a practical instrument’, 274
 ‘*Tom Thumb* performance, a’, 398
 ‘world’s biggest dictionary’, 478
 dictionaries and encyclopedias, 2, 11, 39, 222,
 271, 274, 314, 464, 533
 dictionaries and other lexicographical genres,
 3, 12–13, 47, 57, 59, 64, 85, 232, 271, 293, 357,
 589, 594, 630, 686, 688, 694, 699
Dictionarium latino lusitanicum, ac iaponicum,
 618, 621, 622, 624, 627
Diționarul limbii române, 548–9
 ‘dictionary’, words for
 Arabic, 178, 371
 Chinese, 65, 110, 111, 320
 Japanese, 212
 Persian, 223
 Russian, 223
 Sanskrit, 69, 71

- Spanish, 565, 585, 624
 Tamil, 148, 149
 Turkish, 223, 371
 Turkmen, 223
Dictionary of American English, 513, 519–21, 522, 524–5
Dictionary of Americanisms, 520, 522, 524–5, 529
Dictionary of American Regional English, 509, 513, 520–2, 524–5
Dictionary of Canadianisms on Historical Principles, 523–4
Dictionary of Jamaican English, 509, 525–6
Dictionary of South African English on Historical Principles, 528
Dictionary of the Older Scottish Tongue, 511, 513–14, 516, 519
Dictionnaire coréen-français, 350–2
Dictionnaire de l'Académie française, 305, 451, 533–4, 536, 541
Dictionnaire du moyen français, 536
 Diderot, Denis, 531, 534
 Diez, Friedrich Christian, 532, 754
 Dihkhudā, 'Alī Akbar, 377, 382, 386–7, 754
 Dihlavī, Muḥammad Ḥabībullah Isfahānī Qannawjī, 380
 Dihlavī, Sayyid Aḥmad, 6, 396, 408–9, 410, 754
 Dihlawī, Qāzī Khan, 225, 233
 al-Dīnawarī, 173, 754
 Dīng Shēngshù, 334, 754
 Diogenianus, 102–3, 251, 252, 254, 754
 Dionysius of Halicarnassus, 96
 Dioskoros of Aphroditopolis, 47, 754
 Diouf, Jean-Léopold, 666
 Dirī (Mesopotamian sign list), 25, 34
 Dixon, Robert Malcolm Ward, 528, 682, 700, 754
 Diyari language, 692–3, 713
 Dobson, Edward Scott, 515, 754
Doctrina Christiana y catecismo, 555, 565, 566, 569
 Doke, Clement Martyn, 658, 662, 671–2, 677–8, 681, 754
 Dolan, Terence P., 515
 Dollinger, Stefan, 523
 Dominguez, Ramón Joaquín, 542, 754
 Dominicans as lexicographers
 of Arawakan languages, 745
 of Chinese, 616, 626, 753, 800
 of Ibanag, 746
 of Japanese, 750
 of Kaqchikel, 800
 of Kurdish, 759
 of Latin, 742
 of Mesoamerican languages, 580, 599
 of Mixe, 786
 of Mixtec, 740
 of Quechua, 555, 791
 of South Asian languages, 637
 of Tzeltal, 740
 of Zapotec, 750
Dongguk jeong-un, 219
 Doniach, Nakdimon Shabbethay, 422, 754
 Doroszewski, Witold, 452, 754
 Douglas, Carstairs, 330, 331–2, 754
 Douglas, Gavin, 510
 Douglas, Wilfrid Henry, 697, 701–2, 755
 Dozy, Reinhart Pieter Anne, 420, 755
 Dravidian languages, 143, 401, 527, 639, 713
 Droṇa, 141
 Dror, Michael, 441
 Drummond, Robert, 405
 Dryden, John, 486, 489, 491, 497
 du Cange, Charles du Fresne, 295, 298, 755
 Duckert, Audrey R., 525
 Duden, Konrad Alexander Friedrich, 469, 755
 Duden (publishing company), 469, 471–2
 Duncan, Andrew, 510, 755
 Dunglison, Robley, 517, 518, 755
 Durbin, Marshall, 607
 Đurđević, Bartol, 447, 755
 Durga, 70
 Dürr, Michael, 586, 593
 Dutch language, 214, 216–17, 235, 281, 291, 300–1, 302, 303, 306, 307, 309–10, 342, 422–4, 477–82, 642, 648, 676, 713
 Dyck, Andrew, 264
 Dyirbal language, 700, 713
 Dzongkha language, 388
Ea (Mesopotamian sign list), 21, 25, 31, 34
 Early English Text Society, 494, 512
 Echuwabo language, 673, 713
 Eckart, Anselm, 558, 755
 Edkins, Joseph, 330, 755
 Egede, Poul Hansen, 602, 755
 Egyptian languages, 36–46, 50, 713
 El'azar ben Ḥalfon, 197
 electronic dictionaries, 4, 474, 484, 505, 513, 514, 523, 532
 Arabic, 423–4, 429–30
 Australian languages, 687
 Catalan, 551
 Chinese, 317, 338
 Dutch, 478, 480
 English, 505–8, 523
 French, 536–7

- electronic dictionaries (cont.)
 German, 468, 471, 472
 Hebrew, 436, 438, 442–3, 444
 Icelandic, 477
 Italian, 538, 540
 Japanese, 348–9
 Korean, 355, 356
 North American languages, 606
 North Germanic languages, 477
 Norwegian, 476
 Old Occitan, 550
 Persian, 386
 Portuguese, 547
 Spanish, 541, 544
 Turkish, 374, 376
 Urdu, 410
 Vietnamese, 365
- Elger, Georg, 457–8, 755
 Elias, Elias Antoon, 420–1, 422, 425, 755
 Eliot, John, 602
 Elkin, Adolphus Peter, 685, 688, 697, 702, 755
 Elliott, William Allan, 675, 755
 Ellis, Francis Whyte, 401
 Elyot, Thomas, 306, 756
 Emakhuwa language, 673, 713
 Emesal word list, 15
 Emo, Nicolas, 694
 encyclopedias and encyclopedism. *See also*
 dictionaries and encyclopedias
 Arabic, 178, 418
 Arawakan, 577
 Catalan, 550
 Chinese, 54, 60, 61, 125, 126, 208–9, 212, 221
 Dutch, 480
 East Asian languages, 630
 Egyptian, 38–9, 42, 44, 50
 English, 491, 504
 French, 306, 531, 533–5, 548
 German, 462, 464, 470
 Greek, 88, 253
 Gujarati, 404
 Hebrew, 196, 437
 Italian, 538–9
 Japanese, 208–9, 212, 215, 340–1, 345
 Korean, 221, 222
 languages of the Philippines, 631
 Latin, 272, 275, 296, 309
 Latvian, 458
 Mapudungun, 565
 Marathi, 407
 Mesoamerican, 594
 Mesopotamian, 24
 Persian, 223, 379
 Portuguese, 545
 Romanian, 548
 Sanskrit, 145
 South Asian languages, 642
 Spanish, 302, 542
 Swedish, 475
 Tamil, 150, 155, 647
 Tangut, 203
 Turkic languages, 226, 229, 230
 Turkish, 223, 371
 Endzelins, Jānis, 459, 756
English Dialect Dictionary, 510, 512–13, 516, 520, 524
 English Dialect Society, 512–13
 English language, 216, 291, 306–10, 370, 388, 442, 484–508, 509–29, 713
 Middle English, 282–3, 290, 308
 Old English, 6, 278–80, 286, 308, 510
 English language, bilingual and multilingual dictionaries,
 Afrikaans, 480
 Arabic, 419–24, 425–6, 428–9
 Assamese, 402–4
 Australian languages, 683–705
 Cheyenne, 603
 Chinese, 327, 329–32, 335–8
 Dakota, 603
 Danish, 477
 Dutch, 307, 477
 French, 304, 306–7
 French and Latin, 296
 Frisian, 482
 German, 468
 Gujarati, 398–9, 404–5
 Hausa, 668
 Hebrew, 439–40, 442–3
 Hindi, 399, 411
 Hindustani, 397–8, 400, 405, 408
 Hopi, 605
 Icelandic, 477
 Indic languages, 391
 Inuktitut, 601
 Italian, 307
 Kikongo, 661
 Korean, 351–2
 Latin, 306
 Malay, 642
 Marathi, 400, 405–7
 Oromo, 670
 Persian, 377, 380, 383, 384–5, 386
 Persian and an Indian language, 381
 polyglot, 300–1, 361, 433, 603, 670, 672–3
 Punjabi, 399

- Russian, 307, 448
 Sanskrit, 395
 Shona, 674–5
 Sinhalese, 646
 Spanish, 303, 307
 Swahili, 670, 671
 Tamil, 157, 402, 652
 Turkish, 372, 375–6
 Umbundu, 664
 Urdu, 385, 411
 Welsh, 307
 Western Apache, 610–13
 Xhosa, 678–9
 Yiddish, 483
 Yoruba, 666–8
 Zulu, 676–7
- Eojeong Gyujiang jeonun*, 221–2
Epimerismi, 264
 'Épinal-Erfurt Glossary', 280
 Eratosthenes, 101
Erimhuš (Mesopotamian wordlist), 28–9, 31, 34
Ēryā, 53–7, 59, 60–1, 63–4, 124–6, 203, 216, 218
 Es'ad Efendi, Mehmet, 369, 756
 Eskimo-Aleut languages, 597, 713
 Estienne, Henri, 297–8, 305, 756
 Estienne, Robert, 293–4, 297, 303, 304–5, 306, 309, 310, 756
 Estonian language, 473, 713
 ethnography, 408, 419, 452, 577, 593, 596, 601–2, 608, 631, 693, 764, 767, 784, 786, 796
Etymologicum genuinum, 252, 255, 257–9
Etymologicum Gudianum, 258–9, 265
Etymologicum magnum, 257, 259–60, 263, 264–5
Etymologicum parvum, 256–7
Etymologicum Symeonis, 255, 259, 264
 etymology, 291, 399, 469, 605
- Assamese, 403
 Australian languages, 699
 Catalan, 551
 Chinese, 58, 60
 Dutch, 479
 English, 308, 485, 490, 491, 494, 498, 500, 505, 517, 521, 525
 French, 534–6
 Frisian, 482
 Gaulish, 287
 German, 311, 461, 462–3, 465–6, 468
 Greek, 47, 89, 95, 100, 255–60
 Gujarati, 405
 Hausa, 668
 Hebrew, 185, 188, 194, 198–9, 433
 Hindustani, 409
 Indo-European, 235, 465
- Irish, 278
 Italian, 538, 539–40
 Japanese, 216, 343–4
 Kimbundu, 663
 Latin, 90, 295
 Latvian, 459
 Mesopotamian, 28
 Norwegian, 476
 Pali, 77
 Persian, 235, 384
 Portuguese, 545, 547
 Romance, 532
 Romanian, 548, 549
 Sanskrit, 68, 70, 132–3, 145
 Slavic languages, 447
 Spanish, 302, 544
 Tamil, 657
 Turkic languages, 226
 Turkish, 374–5, 376–7
 universal, 312, 668
 Urdu, 385
 Wolof, 666
 Xhosa, 678
 Zulu, 677, 678
- Eudeve language, 600, 714
 Eulogius, 103, 251
 Euripides, 251, 257
 Eustathius, 263
 Even-Shoshan, Avraham, 434–6, 756
 examples invented by lexicographers, 305, 336, 350, 436, 462, 485, 534, 557, 564–5, 566, 578, 593, 594, 604, 616, 618, 624, 627, 630, 631, 666, 671, 677, 678, 689, 691–2, 695, 699, 701, 703, 704
Expositio notarum, 268, 270
- Faber, Basilius, 295, 756
 Fabra i Poch, Pompeu, 550, 756
 Fabricius, Johann Philipp, 652
 Fabricius, Otto, 602, 756
 Fabvre, Bonaventure, 599, 756
 Fakhr-al-Dīn Mubārakshāh Qawwās Ghaznawī, 232, 756
 Fakhr-al-Dīn Nakhjawānī, 232, 756
 Fallon, Samuel William, 408–9, 411, 756
 Fāng Yǐzhì, 125, 756
Fāngyán, 53, 57–8, 59, 63, 65, 126, 218, 329
 al-Fārābī, Ishāq bin Ibrāhīm, 169, 173, 228, 757
Farhang-i Jahāngīrī, 225, 233, 383
Farhang-i Rashīdī, 234, 378, 383
Farhang-i Shu'ūrī, 384
 Farmer, John Stephen, 518, 757
 Faroese language, 472, 476, 714

- Faruqi, Harith Suleiman, 428
 Favorino, Guarino, 265, 296, 757
 Fay, William Edwards, 664, 680, 757
 Febrès, Andrés de, 556, 561, 563, 564–5, 573, 757
 Fee, Margery, 523
 Fenne, Tönnies, 448, 757
 Fenton, James, 515, 757
 Féraud, Jean-François, 534, 757
 Ferdousi, 232
 Fergusson, John, 653
 Fernández de Palencia, Alfonso, 285, 757
 Ferraguto, Pietro, 243, 757
 Férron, Stanislas, 350
 Figueiredo, Cândido de, 545, 757
 Figueredo, Juan de, 556, 757
 Figueroa, Jerónimo de, 600, 757
 financing of dictionaries, 235, 298, 390, 401,
 409, 419, 425, 433, 466–7, 479, 481, 494–5,
 498, 500, 521, 652
 Finnish language, 312, 472–3, 714
 Finno-Ugric languages, 714
 Firdawsī, 377
 al-Firūzābādī, 178, 181, 223, 371, 380, 418,
 420, 757
 Fischer, August Wilhelm Hermann Gustav,
 420, 758
Fish (Mesopotamian wordlist), 17
 Fitzherbert, John Aloysius, 696
 Flemming, Barbara, 239
 Flierl, Johann, 692, 758
 Florio, John, 307, 758
 Fogelius, Martin, 312
 folklore, 91, 227, 374, 408, 451, 455, 511, 747, 751,
 753, 760, 790, 795, 802
 Fonseca, José da, 674
 Fonseca, Maria do Céu, 652
 Fonte, Juan, 600, 758
 Fortune, George Glynn, 674
 Francard, Michel, 537, 758
 Franciosini, Lorenzo, 303, 537, 758
 Franciscans as lexicographers,
 of Bicol, 772
 of Chinese, 327, 746
 of Cumanagoto, 789, 796
 of Cuman and Persian, 237
 of Czech, 801
 of Irish, 313, 781
 of Kaqchikel, 740
 of languages of the Philippines, 614
 of Latin, 746
 of Matlatzinca, 748
 of Maya, 750, 760, 794
 of Mesoamerican languages, 599
 of Mutsun, 598, 741
 of Nahuatl, 579–80, 778, 781
 of Navajo, 601, 761
 of Nubian, 748
 of Otomi, 799
 of Salinan, 598, 794
 of Slovene, 801
 of South Asian languages, 637, 639, 654
 of Tagalog, 618, 790
 of Tarascan, 759, 771
 of Tupi, 558
 of Turkish, 748, 784
 Francoprovençal language varieties, 532, 535,
 550, 714
 Franklin, Benjamin, 491, 517
Französisches etymologisches Wörterbuch, 535,
 536, 549
 Frawley, William, 597
 Freire-Marreco, Barbara, 609, 610
 French language, 217, 235, 242, 243, 245, 246,
 279, 284, 286–7, 290–1, 293, 296, 299–301,
 302–8, 310, 327, 336–7, 350–1, 352, 361–4,
 367, 371, 384, 386, 420, 424–5, 428–9, 433,
 442, 477, 483, 533–8, 539, 542, 546, 548,
 551–2, 559, 597, 616, 622, 624, 640, 642,
 665–6, 669, 672, 673, 680, 694, 714; *see also*
 Anglo-Norman
 Old French, 185, 278, 286
 Freytag, Georg Wilhelm, 418, 758
 Frisch, Johann Leonhard, 461–2, 758
 Frisius, Johannes, 309, 758
 Frisian language, 481–2, 714
 Friulian language, 539, 551, 714
 Fruchtmann, Maya, 437
 Fula language, 668
 Fürecker, Christoph, 457
 Furetière, Antoine, 306, 314, 533–4,
 541, 758
 Fychan, Simwnt, 279, 758

Gagen shūran, 343, 344, 346
 al-Gāhiz, 181
 Gaisford, Thomas, 260
 Gaj, Ljudevit, 449–50
 Gale, James Scarth, 351, 758
 Gale, Thomas, 253
 Galician language, 551, 714
 Gamilaraay language, 687, 690, 714
 Gangler, Jean-François, 722, 758
Gānlù zishū, 111, 118
 al-Ġarbādaqānī, 169, 758
 García-Medall, Joaquín, 623
 Gardiner, Alan H., 38

- Garo language, 404, 714
 Garzoni, Maurizio, 649, 655, 759
 Gascon language, 535, 550, 714
 Gastelú, Vázquez, 583
 Gatschet, Albert Samuel, 608–10, 759
 Gaul, Niels, 263
 Gaulish language, 287, 715
 al-Ġawāliqī, 166, 759
 al-Ġawharī, 173, 177, 178, 418, 420, 759
 Ge'ez language, 299, 715
 Gelenius, Sigismund, 447
 Gellius, Aulus, 294
 Ġemāl-al-Dīn ibnī Mühennā, 238, 759
 Génibrel, Jean-François-Marie, 362, 759
Genkai, 343–4, 346
 Germanic languages, 277, 279, 280, 283, 285, 306, 310, 311, 482, 715
 German language, 196, 237, 241, 243, 246, 281–2, 285, 290–1, 295, 296, 299–301, 302, 303, 309–11, 336, 367, 384, 421–2, 424, 433, 438, 446–7, 448, 450, 451, 454–6, 457–9, 461–72, 478, 482, 483, 547–8, 603, 669, 672, 692, 693, 715
 Old High German, 279, 280–1, 286, 465
 Gerov, Najden, 450, 759
 Gesenius, Wilhelm, 191, 199
 Gessner, Conrad, 310, 311
 Geytenbeek, Brian, 703
 Geytenbeek, Helen, 703
 Ghālib, Mīrzā Asadullāh Khān, 382, 759
 Gharieb, Gharieb Mohamed, 424, 759
 Ghatage, A. M., 131, 132
 Ghāzī al-Dīn Ḥaydar Shāh, 381, 759; *see also* *Haft Qulzum*
 Ghiyās al-Dīn Muḥammad Jalāl al-Dīn bin Shārif al-Dīn Rāmpūrī, 382, 386; *see also* *Ghiyās al-Lughāt*
Ghiyās al-Lughāt, 378, 382, 383, 386
 Giattini, Giovanni Battista, 245
 Gidabal language, 703, 715
 Gil, Thomas, 700
 Gilbert, G. K., 608
 Gilberti, Maturino, 581, 584, 585–6, 588, 595, 759
 Gilchrist, John Borthwick, 397–8, 400, 759
 Giles, Herbert Allen, 331, 336, 352, 759
 Gilg, Adán, 600, 759
 Gilliéron, Jules, 532
 Giorda, Joseph, 601, 760
 Gitonga language, 673, 715
 Glasgow, Kathleen, 704
 Glass, Amee, 702
 'Glosas Emilianenses', 285, 287
 glossaries to particular texts
 Chinese, 126, 208, 216, 222
 Coptic, 49
 English, 307–8, 516
 Greek, 84–5, 89, 92, 102–3
 Gujarati, 404
 Hebrew, 184–5, 195, 200, 201
 Latin, 84, 284
 Persian, 232
 Romanian, 547
 Scots, 510
Glossarium Gallico-Latinum, 287, 290
 'Glossarium Philoxeni', 92, 276
Glossarium Salomonis, 271–2
 glosses, collections of,
 Breton, Cornish, and Welsh, 278
 Chinese, 51, 53, 56, 57, 58, 60, 64
 Greek, 89, 103, 251
 Irish, 278
 Latin, 269–71
 Old English, 280
 Old French, 286
 Old High German, 281
 glosses, interlinear and marginal,
 Breton, Cornish, and Welsh, 278
 Church Slavic, 446
 Coptic, 47
 Czech, 287
 Danish, Icelandic, and Swedish, 283
 Egyptian, 44, 47
 German, 237
 Irish, 277, 281
 Italian, 284
 Latin, 237, 269, 277
 Matlatzinca, 581
 Old and Middle English, 6, 279, 281, 282–3, 286
 Old French and Anglo-Norman, 282, 286–7
 Old High German and Old Saxon, 280, 286
 Otomi, 581
 Sorbian, 449
 Spanish, 285
 Goddard, Cliff, 686, 688
 Godefroy, Frédéric-Eugène, 535, 536, 760
 Goethe, Johann Wolfgang von, 470
 Goi, 343
 Golding, William, 493
 Golius, Jacobus, 235, 299, 385, 418, 760
 Golla, Victor, 597
 Gonçalves, Jacome, 641, 656, 760
 González de Nájera, Gaspar, 582, 760
 González Holguín, Diego, 555, 561, 563, 571–2, 760

- Goode, Phillip, 611
 Goodwin, Grenville, 606, 609
 Google, 349, 508
 Gopāla, 141
 Goro d'Arezzo, 285, 760
 Gothic language, 312, 715
 Gottsched, Johann Christoph, 463
 Gouws, Rufus, 2
 Gove, Philip Babcock, 499, 760
 Grags pa rgyal mtshan, 147, 760
 grammars
 Achagua, 556
 Arabic, 159–61, 165–6, 168, 179, 201
 Arernte, 693
 Australian languages, 682, 684, 700
 Awabakal, 685, 689–90
 Aymara, 555, 560–1, 565–6, 574
 Bengali, 634
 Carib, 559
 Chibcha, 556
 Chinese, 327, 616, 626
 Chiquitano, 569
 Church Slavic, 448
 Coptic, 49, 50
 Cumanagoto, 559, 562, 567
 Diyari, 692
 early modern European vernaculars, 291
 English, 486, 512
 French, 245, 302, 306–7
 Greek, 248, 256, 257, 262
 Guarani, 557, 563, 575–6
 Gujarati, 404
 Gunin/Kwini, 701
 Hausa, 668
 Hebrew, 186–7, 190, 191–2, 196–7, 563
 Hindustani, 651
 Irish, 278
 Italian, 245
 Japanese, 618
 Kurna, 690
 Kikongo, 660–1
 Kiriri, 557
 Kriol (Australian), 705
 Latin, 98, 273, 274, 287, 292, 448, 584
 Latvian, 458
 Lithuanian, 455–6
 Lule, 561
 Lule and Tonocoté, 567
 Mapudungun, 567, 573
 Marathi, 404, 406
 Matlatzinca, 584, 591
 Mazahua, 594
 Mixtec, 584, 592, 594
 Moxo, 563, 568, 573
 Nahuatl, 580, 583–4, 588
 Nyulnyul, 694
 Oto-Manguean languages, 587
 Pali, 76, 144
 Persian, 235
 Pitjantjatjara, 701
 Polish, 245
 Quechua, 555, 561, 563, 572
 Romance, 532
 Romanian, 548
 Ronga, 673
 Sanskrit, 71, 73, 130, 136, 145, 645
 Slavic languages, 447
 Slovene, 448
 South Asian languages, 634, 651, 653
 Swahili, 671
 Tagalog, 625, 629
 Tamil, 151–2, 153, 155, 158, 650–1
 Tarahumara, 600
 Telugu, 401
 Tupi, 558, 567
 Turkish, 243, 366–7, 368, 372
 Umbundu, 664
 Ungarinyin, 702
 Urdu, 408
 Western Apache, 609–11
 Wiradjuri, 685
 Worrorra, 696
 Yaqui, 600
 Yoruba, 667, 680
 Zapotec, 584, 591
 Zulu, 677
 grammatical information in dictionary
 entries, 94, 97, 175, 177, 192, 196,
 198, 201, 203, 207, 238, 246, 256,
 260, 263–4, 355, 462, 467, 471, 476, 490,
 503–4, 548, 557–8, 568, 570, 592–3, 611, 616,
 625, 627–8, 666, 675, 691, 693, 694, 699,
 701, 703
 Gravier, Jacques, 599, 760
 Gray, Edward, 598
 Grazovski, Yehuda. *See* Gur
 Greek language, 41–3, 44, 45–8, 84–105, 166,
 178, 180, 183–4, 185, 196, 223, 233, 247–66,
 269, 276–7, 278, 284, 287, 290, 296–8,
 300–1, 308, 310, 368, 369, 396–7, 434, 468,
 496, 547, 549, 602, 715
 Greenlandic language, 602, 715
 Gregory of Nazianzus, 262
 Grey, George, 684, 691, 760
 Grimm, Jacob Ludwig Carl, 465–8, 469, 471,
 474, 479, 481, 492–3, 532, 760

- Grimm, Wilhelm Carl, 465–8, 479, 493, 532, 760
 Grimm brothers. *See* individual names;
Deutsches Wörterbuch
 Grønbech, Kaare, 236
 Grossman, Reuben, 439
 Grunau, Simon, 454
 Gù Yěwáng, 62–5, 760
 Guadalaxara, Thomás de, 600
Guǎngyǔn, 112, 113–14, 117
 Guarani language, 557, 558, 561, 563–4, 568, 570, 574–6, 622, 715
 Gude, Marquard, 258
 Guðmundur Andræsson, 717, 761
 Guevara, Miguel de, 583, 593, 761
 Gugu-Yalanji language, 703, 716
 Guignes, Chrétien-Louis-Joseph de, 327, 761
 Gujarati language, 398–9, 404–5, 716
Gǔjīn yùnhuì jǔyào, 204, 219–20
 Gundert, Hermann, 655
 Gunin/Kwini language, 695, 700–1, 716
 Gunn, Edward, 322
 Günther, James, 690, 761
 Günther, Johann Christian, 461
 Gunwinjguan languages, 716
Guóyīn chángyòng zìhuì, 322
Guóyīn zìdiǎn, 322, 333
Guóyǔ cídiǎn, 325, 333
 Gur, Yehuda, 434–5, 761
 Guugu Yimidhirr language, 683, 716
 Guzana, Z., 681
- Haack, Friedrich Wilhelm, 455, 761
 Haas, Mary Rosamond, 598, 761
 Habelić, Juraj, 449, 761
 Hackett, Dorothy, 702
 Hadley, Sydney, 695
 Hadrian of Canterbury, 279, 289
 Haensch, Günther, 544, 761
 Haex, David, 642
Haft Qulzum, 381, 384
 Hai Gaon, 195, 200, 761
 Haile, Berard, 602, 761
 Hakka language variety, 319, 330, 716
 Halāyudha, 132, 134, 137–9, 761
 Halbertsma, Justus Hiddes, 481, 761
 Hale, Horatio Emmons, 685, 762
 al-Ḥalīl bin Aḥmad, 159, 163, 164, 171–2, 174–6, 179–80, 200, 762
 Ḥālīṣ, Yūsuf, 242, 762
 Hall, Joan Houston, 522, 762
 Hall, Joseph S., 522
 Halliwell-Phillipps, James Orchard, 512, 762
- Halma, François, 217, 762
 Hamann, Byron Ellsworth, 561, 576, 584–5, 586, 623, 631
 Hán Dàozhāo, 114, 762
 Ḥan Dehlawī, 762
 Ḥanaguid, Samuel, 190, 191–2, 762
 Handt, Johann Christian Simon, 690, 762
 Hansen, Kenneth, 702
 Hansen, Lesley, 702
 Hanxleden, Johann Ernst, 641, 645, 646–8, 656, 762
Hànyǔ dà cídiǎn, 335, 338
Hànyǔ dà zìdiǎn, 335
 Hanzeli, Victor, 597, 599
 Ḥaq, 'Abdul, 409–10, 762
 al-Harawī, Abū 'Ubayd, 164, 762
 Haraya language, 618, 629, 716
 al-Ḥarbī, 164, 762
 al-Ḥarīrī, 166, 763
 Harṣakīrti, 134, 138–9, 763
 Hartmann, Andrew M., 674–5, 763
 Ḥasan, N., 396
 Hasitzka, Monika, 47
 Hassanein, Ahmad Taher, 417
 Haßler, Gerda, 590
 Hatzfeld, Adolphe, 534, 763
 Hausa language, 668–9, 716
 Hausmann, Franz Josef, 1, 389, 471, 597
 Hauzenberga, Edite, 459, 763
 Haywood, John, 418, 419, 420, 422
 Hayyim, Sulayman, 386, 763
 Ḥayyūj, Yehuda, 189–90, 192–3, 196, 763
 headwords
 bound morphemes as, 323, 574, 629–30, 664, 691, 699
 citation form, 436–7, 569–71, 590, 592, 613, 676
 complex, 126, 303, 557, 589–90, 646, 676
 not always obligatory, 3, 156
 redundant, 613
 repeated for each meaning of a word, 662
 script used for, 60, 118, 122, 276, 352, 360, 361, 363, 367, 374, 458, 615, 648, 669
 selection of, 18, 91, 94, 118, 177, 178, 181, 305, 328, 334, 336–7, 343, 398, 434, 446, 453, 466–7, 486, 496, 512, 518, 520, 531, 533, 538, 541, 571–3, 576, 675, 698
 Heath, Jeffrey G., 686, 688, 763
 Hebrew language, 4, 179, 182–201, 277, 278, 284, 286, 287–8, 298–9, 301, 308, 386, 431–44, 482–3, 563, 570, 604, 649, 657, 697, 716
 Heckewelder, John Gottlieb Ernestus, 603, 763

- Heflin, Woodford A., 519
 Heinmiller, Jennifer H., 522
 Heliodorus, 103
 Hemacandra, 131–2, 134, 137, 139, 141–2, 145, 763
 Henisch, Georg, 311, 763
 Hennig, Christian, 449
 Henriques, Henrique, 153, 650
 Herbert, Thomas, 234, 763
 Herennius Philo, 94, 263, 763
 Hermanson, Eric, 676
Hermeneumata, 269, 270, 272, 276, 278, 281, 289
 Hermolaus, 263
 Herodian, 103, 252
 Hershberger, Hank, 703
 Hershberger, Ruth, 703
 Hesiod, 257
 Hesychius of Alexandria, 86, 94, 102–5, 250–1, 264–5, 763
 Hesychius of Miletus, 254
 Hewson, John, 599
 Hidāyat, Rizā Qulī Khān, 383
 Higgins, N. S., 608, 610
 Hiligaynon language, 618, 629, 716
 Hill, Faith, 611–12
 Hill, Kenneth, 597
 Hillier, Henry, 693
 Hilton, K. Simon, 604
 Hince, Bernadette, 529
 Hindi language, 385, 395, 398, 399, 404, 407, 410–12, 716–17
 Hindawī, 392
 Hindi-Urdu, 384
 Hindustani, 393, 397–8, 400, 405, 407, 408–9, 639, 641, 648, 651
 Hinds, Martin, 426, 429, 764
 Hinduism and Brahmanism, 72, 130, 131, 134, 138, 139, 379, 391, 393, 639, 643, 644
 Hiraethog, Gruffudd, 279, 764
Historical Dictionary of the Hebrew Language, 433, 437, 444
 historical orientation,
 Arabic, 420
 Catalan, 551
 Chinese, 321, 322, 335
 Czech, 451
 Danish, 474
 Dutch, 478
 English, 488, 489, 491–5, 505, 508, 510, 513–14, 519–21, 523, 525–6, 528–9
 French, 534–5, 536
 Frisian, 482
 German, 462, 465, 466, 469
 Hebrew, 432–4, 435–6, 437–8, 444
 Icelandic, 477
 Italian, 539–40
 Japanese, 346–7
 Polish, 451, 452–3
 Portuguese, 545, 547
 Romance languages, 532, 552
 Romanian, 548, 549
 Scots, 510, 513–14
 Spanish, 545
 Swedish, 475
 Tamil, 157
 Turkish, 374, 377
 Urdu, 410
 Hitti, Yusuf, 428
 Hittite language, 15, 28, 717
Hobson-Jobson, 527
 Hodge, David McKellop, 603, 764
 Hodge, John, 352
 Hoff, August, 693
 Hogarth, Richard, 308, 764
 Hoijer, Harry, 609, 610–11, 764
 Holdermann, Jean-Baptiste Daniel, 367, 764
 Holm, John, 525, 764
 Homer, 84–5, 86, 89, 92, 98, 102–3, 250–1, 256, 257, 264
 homonym lists
 Arabic, 161, 164, 167, 170
 Hebrew, 197, 201
 Sanskrit, 69–70, 131, 132, 136, 138, 139–40, 148
 homonymy, 57, 58, 59, 70, 80, 94, 132–3, 134, 166, 194, 198, 385, 390, 394, 627
Hongmu jeong-un yeokhun, 219–20
Hóngwú zhèngyùn, 112, 114, 219
 Hoogland, Jan, 417, 423, 424, 764
 Hooker, Richard, 491
 Hopi language, 604–5
 Hosayn Zawzanī, 231, 764
 Houaiss, Antônio, 547, 764
 Houtman, Frederik de, 642, 649, 656, 764
 Houtsma, Martijn Th., 236, 238
 ‘Houtsma’s Glossary’, 238–9
 Hrinčenko, Borys Dmytrovyč, 453, 764
 Huarpean languages, 717
 Hudson, Joyce, 703, 705
 Hue, Gustave, 364, 764
 Hugga, 71
 Hughes, Earl J., 703
 Hugutio, 268, 274–5, 285, 290, 292, 764
 Hulbert, James Root, 519, 764
 Humphreys, David, 516, 765
 Hundertpfund, Rochus, 558, 765
 Hungarian language, 300–1, 309, 312, 447, 449, 451, 547–8, 549, 717

- Hunger, Herbert, 255
 Hunter, Joseph, 511–12, 513, 516, 765
 Hunter, William, 400, 765
 Huron language, 599, 717
 Hurrian language, 15, 28, 717
 Hūsayn Injū, 225, 233, 765; *see also* *Farhang-i Jahāngīrī*
Hwa-Dong jeon-geum tongseok un-go, 221
Hyōta jiruishō, 211, 213
- Ibanag language, 618, 717
 Ibn al-Anbārī, Abū Bakr, 167, 765
 ibn Balʿam, Yehuda, 196–7, 765
 ibn Barūn, Isaac, 199
 Ibn Durayd, 162, 165, 172, 174–5, 200, 765
 Ibn Durustawayhi, 167
 Ibn Ezra, Moshe, 197, 765
 Ibn Fāris, 167, 174, 176, 765
 Ibn Ḥālawayhi, 164, 765
 Ibn Janāh, Yona, 189–94, 197, 198–9, 765
 Ibn Manzūr, 178, 180–1, 418, 765
 ibn Parḥon, Shelomo, 192
 Ibn al-Qaṭṭāʾ, 170, 765
 Ibn Quraysh, Yehuda, 188, 191, 198, 199, 765
 Ibn al-Ṣaġarī, 167, 766
 Ibn Sīda, 161, 162, 170, 174, 177, 181, 766
 Ibn al-Sikkīt, 167–8, 173, 179, 766
 ibn Tibbon, Yehuda, 191–2, 766
 Ibn Wallād, 169
 Icelandic language, 283, 312, 472, 476–7, 717
 idioms. *See* multi-word expressions
 Idriss, Souheil, 425, 766
 Igbo language, 668, 717
 illustrated dictionaries, 215, 296, 404, 429, 491, 523, 604, 647, 697, 701–2
 Ilocano language, 618, 625, 629, 717
 Inbal, Shimson, 440
 indexes
 dictionary-like, to non-lexicographical works, 558, 565, 594
 to lexicographical works. *See* paratexts
 Indo-Aryan languages, 67, 72–3, 380–1, 384, 391–2, 639, 718
 Indo-European languages, 2, 396, 401, 456, 465, 492, 531, 532, 673, 718
 Inshā, M. I. K., 407
 Inuktitut languages, 589, 601, 718
 Inupiaq language, 604, 718
 Iranian languages, 95, 718
 Irish language, 277–8, 281, 313, 515, 718
Iroha jiruishō, 210
 Iroquoian languages, 599–600, 718
 al-Iṣfahānī, Ḥamza, 165, 766
- Ishikawa Masamochi, 343, 766; *see also* *Gagen shūran*
 Isidore, 271, 274, 278, 766
 Islam, 48, 159–61, 165, 179, 228, 232, 239, 241, 267, 292, 366, 392, 393, 409, 415, 595, 640; *see also* Qurʾān
 Italian language, 185, 193, 235, 241, 243–5, 246, 284–5, 291, 299–301, 303–5, 307, 310, 367–70, 375, 447, 448, 537–40, 547, 551, 616, 622, 649, 660, 694, 718
 Italic languages, 718
 Iulius Vestinus, 102
 Iyura language, 683, 718
 Izi (Mesopotamian wordlist), 25
- Jabirrabirr language, 699, 718
 Jablonskis, Jonas, 456
 Jacobson, Steven, 604
 Jainism, 72–3, 74, 78, 130, 134–5, 139, 156, 763
 Jakobsen, Jakob, 714, 725, 766
 Jakubaš, Filip, 454, 766
 James, Gregory, 635, 640, 647, 648
 Jamieson, John, 493, 510–11, 512–13, 766
 Japanese language, 63, 207–17, 222, 328, 335, 340–50, 353, 618–21, 622, 624, 626–7, 629, 630, 631–2, 719
 Jarrakan languages, 699, 719
 Jauncey, Dorothy, 528
 Jawi language, 695, 699, 719
 Jenssøn, Christen, 726, 766
Jeon-un-ogpyeon, 222, 351
 Jerome, 277, 766
 Jesuits as lexicographers
 of Abenaki, 741, 786
 of Achagua, 780, 788
 of Allentiac, 799
 of Aymara, 555, 744, 757, 798
 of Basque, 771
 of Carib, 783
 of Chinese, 326, 328, 616, 626, 749, 787, 789, 798
 of Coptic, 50, 769
 of Croatian, 753, 777
 of East Asian languages, 614
 of Eudeve, 759
 of French, 533, 757, 778, 785
 of Guarani, 557, 740, 787, 789
 of Huron, 785
 of Japanese, 214, 620–1, 630, 793
 of Kalispel, 760
 of Kiriri, 775
 of Koyukon, 766
 of Lakota, 746

- Jesuits as lexicographers (cont.)
 of languages of northern Mexico, 600
 of Latin, 295, 743, 785
 of Latvian, 457, 755
 of Lithuanian, 793
 of Lule, 774
 of Malayalam, 762
 of Mapudungun, 757, 799
 of Mesoamerican languages, 599
 of Miami-Illinois, 760, 785
 of Millcayac, 799
 of Mohawk, 746
 of Montagnais, 756, 793
 of Moxo, 775
 of North American languages, 598, 599, 601
 of Northern Tepehuan, 758, 788
 of Nyungwe, 751
 of O'odham, 768, 792
 of Opata, 773
 of Quechua, 555, 757, 760, 800
 of Sanskrit, 397, 645, 762, 785, 788
 of Shona, 674, 763
 of South Asian languages, 637, 638
 of Spanish, 542, 797
 of Tagalog, 781, 791
 of Tamil, 149, 396, 644, 655, 744, 786
 of Tarahumara, 757, 794
 of Telugu, 641
 of Tepehuan, 757
 of Tupi, 558, 755, 765, 776
 of Turkish, 243, 367, 757, 764
 of Vietnamese, 787
 of Yaqui, 743, 800
- Jetté, Jules, 601, 766
Jǐjiù piān, 52–3
 Jñānaśrīmitra, 135
 John Samannudi, 49–50, 767
 Johnson, Francis, 385, 767
 Johnson, Frederick, 672, 681, 767
 Johnson, Samuel, 398, 406, 484–5, 486–93, 495, 498, 503, 517–18, 767
 Jones, Eliza, 601, 767
 Jones, William, 397, 401, 465, 655, 767
 Joyce, James, 502
 Ju Si-kyeong, 353, 767
 Jukun language, 699, 719
 Jungmann, Josef, 451, 767
 Junius, Franciscus, 715, 767
 Junod, Henri-Alexandre, 673, 767
 Jurchen language, 203, 205–6, 719
 Juška, Anatanas, 456–7, 767
 Juwaliny language, 704, 719
- Kaberry, Phillis M., 685
Kagakushū, 212, 215
 Kakati, B. K., 404
 Kalispel language, 601, 719
 Kallierges, Zacharias, 260, 265
 Kamba language, 670, 719
Kāngxī zhidān, 110, 112, 117, 120–2, 207, 218, 320–1, 322–3, 328, 335, 338, 347
 Kañkēyar, 149, 767; *see also* *Uric col nikañtu*
 Kannada language, 393, 401, 635, 719
 Kant, Immanuel, 655
 Kaqchikel language, 581–3, 587, 592, 719
 Karadžić, Vuk, 449–50, 767
 Karajarri language, 704, 719
 Karbi language, 404, 719
 Karłowicz, Jan Aleksander Ludwik, 452, 767
 Kaspi, Josef, 194, 198, 768
 Katakī, D., 404
 Katre, Sumitra, 67, 71, 135
 Kātya, 71
 Kauria language, 690–1, 720
 Kayātarar, 149, 768
 Kelham, Robert, 707, 768
 Kenbō Hidetoshi, 345, 768; *see also* *Meikai kokugo jiten*
 Kennedy, Vans, 400, 768
 Kenrick, William, 490, 768
Keunsajeon, 353–5
 Khal'fin, Sagit, 733, 768
 Khan, M. M. T., 411
 al-Khatib, Ahmad, 428
 Kholar, V. L., 404
 Khoury, Marielle, 424
 Khro phu Lo tsā ba Byams pa'i dpal, 146
 Kḥurāsānī, Yūsuf bin Muḥammad Yūsuf, 391, 768
 K'iche' language, 581, 583, 587, 592–3, 595–6, 720
 Kieffer, Jean Daniel, 371, 768
 Kikongo language, 659–62, 720
 Kilham, Christine, 704
 Kiliaan, Cornelis, 310, 768
 Kilpatrick, Kelly, 515
 Kimbundu language, 660–1, 662–4, 680, 720
 Kimchi, David, 193–4, 768
 Kimwani language, 673, 720
 Kindaichi Kyōsuke, 768
 Kinkade, M. Dale, 597
 Kino, Eusebio Francisco, 600, 768
 Kipchak languages, 236, 237–9, 733
 Kipokomo language, 670, 720
 Kircher, Athanasius, 50, 299, 327, 769
 Kiriri language, 557, 720

- Kirticandra, 147, 769
 Kirwin, William James, 524, 769
 al-Kisā'ī, 166, 769
 Klaatsch, Hermann, 685
 Klappenbach, Ruth, 470, 768
 Klausner, Joseph Gedaliah, 438, 769
 Klein, Daniel, 455
 Kleinschmidt, Samuel Petrus, 602, 769
 Klöter, Henning, 615
 Knaani, Yaakov, 435–6, 769
 Knox, Robert, 646
 Koelle, Sigismund Wilhelm, 672, 769
Kōjien, 345, 348
 Kǒng Ānguó, 56, 64, 769
 Kong Julian, 412
 Konkani language, 635, 641, 646, 655, 720
 Kopitar, Jernej, 450
 Korean language, 63, 216, 217–22, 328, 349–56, 720
 Koyukon language, 601, 720
 Krah, Günther, 424, 769
 Kramer, Johannes, 539
 Kramer, Matthias, 478, 769
 Kramvant, Jaganathshastri, 406, 770
 Krapf, Johann Ludwig, 670–1, 680, 770
 Krasnodembsky, V., 412
 Krauss, Michael, 597
 Kriol language (Australian), 705, 720
 Kro, B. S., 404
 Kropf, Johann Heinrich Albert, 678–9, 680, 770
 Kropfitsch, Lorenz, 417, 424, 770
 Krünitz, Johann Georg, 464
 Kryński, Adam Antoni, 452
 Kṣīrasvāmin, 136, 770
 Kūkai, 65, 208, 770
 Kukatja language, 702, 704, 720
 Kun dga' rgyal mtshan, 147, 770
 Kurā' al-Naml, 167, 180, 770
 Kurdish languages, 649, 655, 720
 Kurmin, Jan, 458, 770
 Kurschat, Alexander Theodor, 457, 770
 Kurschat, Friedrich, 456–7, 770
 Kusumaprajñā, Samanī, 79
 Kuun, Géza, 236
 Kuyani language, 692, 720
 Ký, Pétrus Trương Vinh, 363, 770

 labelling, 104, 336, 426, 461, 463, 470–1, 484–5, 487, 490, 496, 499, 501, 503–4, 505, 507, 513, 521, 546, 566–8, 591, 627–9, 632, 648
 Lacombe, Albert, 601, 770

 Ladin language variety, 539, 721
 Lagadeuc, Jehan, 279, 771
 Lagunas, Juan Baptista de, 581, 585, 588–9, 771
 Lakota language, 601, 721
 Lambrecht, Joos, 310
 Landsberger, Benno, 35
 Lane, Edward William, 179, 418–20, 424–5, 429, 771
 Lane-Poole, Stanley Edward, 419, 771
 Lange, Jacob, 458, 771
 Langenscheidt dictionaries, 417, 424, 433, 468
 language death,
 Australian languages, 684, 705
 Hebrew, 182
 Polabian, 449
 South Asian languages, 635
 Sumerian, 15, 20, 26
 la Parra, Francisco de, 587
 Lara, Luis Fernando, 544, 771
 Larousse, Pierre, 534, 771
 Larousse dictionaries, 364, 386, 534, 535, 548
 Larramendi, Manuel, 709, 771
 Las Casas, Cristóbal de, 303, 771
 Laski, Marghanita, 500
 Latif, Muhammad, 411
 Latin language, 6, 39, 45–9, 89–94, 95–6, 98–9, 102, 104–5, 110, 183–4, 196, 235, 237, 241, 243, 246, 267–89, 290–302, 303–7, 308–12, 327, 351, 361, 367–8, 384, 396–7, 406, 418, 434, 446–50, 451, 455, 457–8, 467, 472, 477, 478, 486, 510, 517, 531–2, 533, 545, 547–8, 549, 550, 557, 580, 582, 602, 618, 620, 622, 624, 626, 633, 637, 640–2, 645, 647–8, 649, 659–60, 662, 721
 Latvian language, 457–9, 721
 Lauden, Edna, 442
 Laughlin, Robert Moody, 6–7, 771
 Laurell, Yngve, 685
 Laves, Gerhardt Kurt, 685, 771
 Lawrence, D. H., 502
 al-Layṭ bin al-Muẓaffar, 175, 771
 Lê Ngọc-Trụ, 365
 learners' dictionaries, 93, 337–8, 354, 363, 429, 440, 441–2, 444, 473, 484, 503–4, 532, 548, 560, 594, 605, 622, 675
 Lebedev, K. A., 412
 LeBoullenger, Antoine-Robert, 599
 Lee, Siu-Yau, 334
 Lee Yoonjae, 354
 Leeding, Velma, 704

- legal vocabulary, 136, 268, 278, 308, 313, 411, 428
 Legrand de la Liraye, Théophile Marie,
 362, 771
 Le Guennec, Grégoire, 665, 771
 Leibniz, Gottfried Wilhelm, 461, 462
 Lemoine, Georges, 601
 Lěng Yùlóng, 335, 771
 Leonhardi, Moritz von, 693
 Leo the Mathematician, 257
 Le Page, Robert Brock, 525–6, 772
 Lepsius, Karl Richard, 671, 672–3, 678, 680
 Le Roux, Jean-Marie, 669
 LeSourd, Philip, 604
Lessico etimologico italiano, 540, 549
 Le Talleur, Guillaume, 287, 772
 Levenston, Edward, 439
 Le Ver, Firmin, 286–7, 772
 Levita, Eliahu, 196, 772
 Lévy, Cerf, 350
 Levy, Emil, 550, 772
 Lewis, Geoffrey, 373
 lexicography, concept of, 2–3, 288
Lexicon Vindobonense, 262, 265
 Lí Jīnxī, 325, 334, 772
 Lǐ Lùxīng, 337, 772
 Lǐ Róng, 334, 772
 Lǐ Shí, 127, 772
 Lǐ Sī, 52, 772
 Lǐ Tǎo, 119, 772
 Lǐ Yángbīng, 118–19, 772
 Liang Shih-chiu, 336–7, 772
Liber glossarum, 271–3
Libù yùnlüè, 111, 114, 218
 Lin Yutang, 336–7, 772
 Linde, Samuel Bogumił, 451, 772
 Lisboa, Marcos de, 618, 772
 literacy, 19, 27, 35, 36–7, 41, 212, 213, 268, 277,
 333, 340, 341, 348, 358, 486, 603, 639, 693,
 697, 701
 Lithuanian language, 455–7, 721
 Littré, Émile Maximilien Paul, 468, 493, 534–7,
 542, 773
 Liú Xī, 58–9, 773
Liùshū gù, 119, 124
 Lloyd, William, 309, 773
 loanwords,
 in Arabic, 160, 161–2, 164, 165–6, 178, 223,
 426, 427
 in Aramaic, 196
 in Australian languages, 699
 in Chinese, 126
 in Czech, 451
 in Danish and Swedish, 472
 in Dravidian languages, 401
 in Dutch, 479
 in English, 308, 501, 527, 528
 in German, 463, 465, 466, 468, 470
 in Hebrew, 183, 196, 199, 434
 in Hindi, 410–11
 in Italian, 539
 in Japanese, 341, 345
 in Kimbundu, 663
 in Latvian, 459
 in Luxemburgish, 483
 in Mesoamerican languages, 592, 595
 in Persian, 223, 233, 380, 381
 in Prakrit, 141, 143
 in Romanian, 547–8
 in South American languages, 576
 in South Asian languages, 393, 407, 413, 636
 in Spanish, 541–2
 in Sumerian, 33
 in Swahili, 670
 in Tagalog, 628
 in Tamil, 154
 in Telugu, 393
 in Turkic languages, 223, 239
 in Turkish, 223, 240, 367–74
 in Vietnamese, 362
 in Zulu, 676
 Lobscheid, Wilhelm, 330, 773
 Loderecker, Peter, 448, 773
 Loderus, Andries Lambert, 642
 Loew, O., 608
 Lokono language, 556, 721
 Lombardo, Natale, 600, 773
 Lommatzsch, Erhard, 536, 773
 Lomonosov, Mikhail, 452
Lóngkān shǒujiàn, 128
 Longman dictionaries, 338, 498
 Lopadiotes, Andreas, 262, 773; *see also* *Lexicon*
 Vindobonense
 López, Gaspar, 563
 los Reyes, Alonso de, 584, 592
 los Santos, Domingo de, 617, 621, 626, 628
 lost dictionaries, 192, 327, 615
 African languages, 659, 681
 Arabic, 170
 Burmese, 641
 Chinese, 51–2, 62, 65, 204, 326, 615, 626
 Greek, 98, 251, 263
 Guarani, 575
 Japanese, 618
 Kikongo, 660
 languages of Philippines, 617, 621
 Latin, 286

- Nepali, 641
 Northern Tepehuan, 600
 Persian, 229
 Sanskrit, 69, 79, 142
 South Asian languages, 638–9, 654
 Tamil, 640, 652
 Tibetan, 641
 Vietnamese, 361, 620
 Loughridge, Robert McGill, 603, 773
 Love, James Robert Beattie, 697, 704, 773
 Lovell, Charles J., 522, 773
Lu A (Mesopotamian wordlist), 16–17, 18
 Lữ Chén, 62
 Lữ Diàn, 124, 773
 Lữ Ērkuí, 322, 773; *see also* *Cíyuán*
 Lữ Fáyán, 112, 774
 Lữ Gǔsūn, 336, 774
 Lữ Shūxiāng, 334, 774
 Ludolf, Hiob, 299, 774
 Lule language, 556, 561, 567, 721
 Lule-Vilela languages, 721
 Luó Zhúfēng, 774
 Luritja language, 693, 702, 721
 Luther, Martin, 462
 Lutheran missionaries as lexicographers, 602,
 611, 637, 639, 640, 678, 690, 692–3, 756, 758,
 770, 777, 787, 792, 795, 796, 805
Luxemburger Wörterbuch, 483
 Luxemburgish language, 483, 722
 Lyon, Robert M., 684

 Maaler, Josua, 309, 774
 Maas, Paul, 250
 Mabrou, Abdelouahed, 426
 Macafee, Caroline, 515, 774
 Macaulay, Thomas Babington, 487
 Macedonian language, 445, 450, 454, 722
 MacGowan, John, 330, 774
 Machado, José Pedro, 546, 774
 Machoni de Cerdeña, Antonio, 561, 567, 774
 MacIver, Donald, 330, 774
 Maclay, Robert Samuel, 330, 774
Macquarie Dictionary, 508, 509
 macrostructure, alphabetical and syllabic
 German, 311
 macrostructure, alphabetical or syllabic,
 292, 313
 Arabic, 161–2, 163, 164–5, 166, 167, 168–9,
 171–4, 175, 176, 177–8, 180, 231, 381,
 416–17, 419
 Arawakan, 570
 Australian languages, 698
 Awabakal, 689
 Aymara, 566, 574–5
 Barngarla, 691
 Breton, 279
 Carib, 559
 Chinese, 204, 318, 326, 330, 332–4, 336–7, 338,
 615, 616
 Chiquitano, 569
 Diyari, 692
 Dutch, 217
 English, 308–9, 423, 494, 670, 685, 690, 692,
 695, 698, 700
 European languages, 317
 French, 287, 303, 305, 533, 694
 German, 282, 457, 468, 692
 Greek, 84, 85–6, 92, 93, 95, 102–3, 249–50,
 251–2, 254, 255, 256–7, 262, 263–5, 269,
 296, 298
 Guarani, 576
 Gunin/Kwini, 695
 Hebrew, 179, 187, 189, 194, 195–7, 200, 235,
 434, 439
 Irish, 313
 Italian, 304, 368
 Japanese, 210, 211–12, 213, 214–16, 342, 619
 Kaurna, 691
 Kimbundu, 663
 Korean, 222, 351
 Latin, 84, 90, 93, 95, 237, 269–70, 271, 273,
 275, 282, 289, 290, 292–3, 295, 367
 Malayalam, 649
 Manchu, 206
 Marathi, 406
 Matlatzinca, 587
 Mesoamerican languages, 586, 590
 Moxo, 569
 North American languages, 606
 Nunggubuyu, 703
 Persian, 224–5, 230, 231, 233, 381, 387
 Prakrit, 142
 Quechua, 561
 Ramindjeri, 691
 Romanian, 549
 Sanskrit, 71, 138, 139–40, 148
 South American languages, 569
 South Asian languages, 392, 643, 648
 Spanish, 559, 561, 571
 Tamil, 152–3, 155, 644, 650
 Tarascan, 588–9
 Telugu, 401
 Tupi, 570
 Turkic languages, 228, 230
 Ungarinyin, 703
 Vietnamese, 361, 362–4

- macrostructure, alphabetical (cont.)
 Western Apache, 611–12
 Western Desert Language, 702
 Zapotec, 591
 Zulu, 677
- macrostructure, other,
 acrographic, 12–13, 21, 25, 45
 alliterative, 152, 155
 associative, 25, 70
 by form of character or sign, 12, 53, 59–60,
 61, 62–3, 66, 110, 118–21, 206–8, 211, 213,
 214, 218, 220, 317, 320, 322, 325, 334, 337,
 342, 347, 615
 by grammatical category, 71, 140, 180, 228,
 231, 243, 255, 282, 306, 367, 689–90
 by length of definition, 139–40, 142
 by length of headword, 76, 140, 142, 148,
 168, 174–5, 176, 215, 228, 230, 324, 332, 334–5
 by number of meanings, 153
 by number of objects designated, 151, 155
 by order of occurrence in a given text, 49,
 84–5, 164, 200, 264, 269
 by rhyme, 65–6, 110, 112–18, 119, 170, 172–4,
 177, 178, 179–80, 197, 204, 207, 208, 211, 213,
 214, 218–22, 321, 329, 380
 by shared root consonant, 13, 28
 derivational, 100, 246, 273, 274, 293, 297–8,
 305, 309, 311, 313, 416–17, 419–20, 422, 423,
 434, 468, 491, 533, 589, 630
 designed for continuous reading, 62, 125
 entries from same source grouped
 together, 262, 269–70
 etymological, 28
 geographical, 43, 61
 homophones treated together, 118
 inflected forms disturb alphabetical
 sequence, 570
 ‘Instant Index System’, 337
 loanwords treated separately, 370
 morphological, 71, 170, 173, 176, 230, 231
 names of deities treated separately, 69–70
 permutational, 164, 171–2, 174–5, 176, 177,
 180, 195, 200–1
 phonological, 169, 174, 228, 230, 317, 325
 roots listed separately, 588
 synonyms and homonyms treated
 separately, 69, 72, 131, 136, 142
 thematic groups interrupt alphabetical
 ordering, 694
 unsystematic, 13, 29, 87, 100, 119, 139, 142,
 148, 161, 163–4, 169, 173, 187, 204, 225, 615
- macrostructure, thematic,
 Arabic, 161–3, 165, 166, 167–8, 170, 231
 Aramaic, 232
 Australian languages, 684, 690
 Carib, 559
 Chinese, 52, 54, 56–8, 61, 63, 64–5, 119–20,
 123, 125, 205, 207, 209, 210–12, 213, 216, 220,
 222, 357, 360, 413, 615
 Coptic, 49
 Czech, 288, 446
 Danish, 311
 Egyptian, 37–9, 43
 European languages, 296, 313
 French, 303, 559
 German, 281, 282, 296, 299, 454, 457
 Greek, 84, 87–8, 93, 100, 102
 Hebrew, 197
 Hindawi, 391
 Italian, 285, 299, 539
 Japanese, 207, 208, 209, 210–12, 213, 215–16,
 217, 222, 342, 619
 Korean, 220, 222
 Latin, 269, 281, 282, 285, 288, 295–6, 303,
 446, 645
 Latvian, 457
 Manchu, 206, 222
 Marathi, 406
 Mazahua, 594
 Mesopotamian, 12–13, 17, 21, 22, 24, 26,
 31–2, 34
 Mongolian, 205, 222
 Nyangumarta, 704
 Old Prussian, 454
 Pali, 143–4
 Persian, 224, 231–3
 Pitjantjatjara, 701
 Prakrit, 142
 Sanskrit, 70, 136–9, 147, 643, 645
 South Asian languages, 391, 394, 413
 Swedish, 311
 Tamil, 150, 155
 Tangut, 203
 Telugu, 400
 Tibetan, 146–8
 Turkish, 367
 Urdu, 413
 Vietnamese, 357, 360–1, 363
 Western Desert Language, 702
- Maćczyński, Jan, 447, 774
 Madan, Arthur Cornwallis, 671–2, 774
 Maehler, Herwig, 96
 Magnús Ólafsson, 717, 774
Mahāvīyutpatti, 146
 Maheśvara, 134, 147, 775
 Mahican language, 602–3, 722

- Maḥmūd al-Kāšġarī, 173, 225–30, 238, 366, 775
 Maimonides, Moses, 195
 Majid, ‘Abdul, 396
 al-Mālaqī, 168, 775
 Malaret, Augusto Nicolás, 544, 775
 Malayalam language, 393, 640, 641, 646–9, 655, 722
 Malay language, 477, 636–7, 638, 640, 642, 646, 648–9, 722
 Malayo-Polynesian languages, 639, 722
 Malcolm, Daniel McKinnon, 677
 Malige-Klappenbach, Helene, 470, 775
 al-Malik al-Afdal, 180, 775
Malku = *šarru* (Mesopotamian wordlist), 29, 31, 34
 Mallinātha, 136, 141, 775
 Mamiani, Luis Vincencio, 557, 775
 Mancelius, Georg, 457–8, 775
 Manchu language, 205–7, 214, 216, 222, 722
 Mandarin language (pǔtōnghuà), 318, 337, 722;
 see also Chinese
 Mandarin topolects, 318, 723
 Mangala language, 704, 723
 Mañkha, 132, 136, 141, 145, 775
 Mannheim, Bruce, 572, 577
 Manutius, Aldus, 251
 Máo Qíling, 127, 775
 Mapudungun language, 556, 561, 562, 564–5, 567, 572–3, 723
 Marak, H. M., 404
 Marathi language, 393, 400, 404–7, 723
 Marbán, Pedro, 563, 568, 569, 573, 775
 Marchesinus, Joannes, 447
 Marlett, Stephen, 604
 Marsden, William, 1, 648
 Marsh, James, 702
 Martial, 292
 Martin, Benjamin, 488, 775
 Martin, Willy, 478
 Martini, Martino, 626
 Martu Wangka language, 702, 723
 Marwaḥ language, 410
 Maslai, D. U., 404
 Massachusetts language, 602, 603, 723
 Masthay, Carl, 603
 Mathews, Mitford M., 519–20, 522, 775
 Mathews, Robert Hamilton, 684, 776
 Mathews, Robert Henry, 331, 336, 776
 Matlatzinka language, 581, 582–4, 586–7, 590–1, 593, 723
 Matsui Kanji, 344, 346, 347, 776
 Matsui Shigekazu, 346, 776; *see also* *Nihon kokugo daijiten*
 Matsumura Akira, 345, 776; *see also* *Daijirin; Daijisen*
 Matta, Joaquim Dias Cordeiro da, 663, 776
 Mavoungou, Paul Achille, 668
 Mawer, Allen, 515, 776
 Maya language, 6, 582, 586, 723
 Mayan languages, 580, 582, 586, 593, 723
 al-Maydānī, Abu’l-Faẓl, 165, 231, 776
 Mayo language, 604, 723
 Mazahua language, 594, 723
 al-Māzinī, 163
 Mbarrumbathama language, 687, 723
 McCarus, Ernest, 416
 McConnel, Ursula H., 685
 McGregor, William B., 688
 McKelson, Kevin, 704, 776
 McMillan, James B., 520
 Medhurst, Walter Henry, 329–30, 350, 776
 medical vocabulary,
 Arabic, 171, 178, 181, 391, 429
 English, 308
 German, 462
 Greek, 98, 102, 103, 105, 256
 Hindawī, 391
 Japanese, 216
 Persian, 380, 391
 Sanskrit, 131, 133, 136–7
 Urdu, 411
 Medinikara, 140, 776
 Megiser, Hieronymus, 241, 243, 448, 449, 776
 Mehta, B. B., 405
 Mehta, B. N., 405
 Méi Yǐngzuò, 120–1, 776
Meikai kokugo jiten, 345, 347
 Meisterburg, Anton, 558, 776
 Melanchthon, Philipp, 448
 memorization, 15–16, 22, 52, 62, 79–80, 131, 145, 148, 150, 204, 278, 288, 359–60, 380, 390–1, 394–5, 594
 Menaḥem ben Saruq, 188–9, 198, 199–200
 Menander, 97
 Mencken, H. L., 518
Měnggǔ zìyùn, 204–5
 Meninski, François à Mesgnien, 241, 244–6, 366, 367–8, 372, 383–4, 385, 777
 Mennonite missionaries as lexicographers, 603, 784
 Métrida, Alonso de, 618, 622, 623, 628, 777
 Merriam-Webster dictionaries, 491, 498–500, 505, 508, 509, 518
Metal (Mesopotamian wordlist), 17
 metaphor, 68, 77, 89, 94, 100, 177, 233, 380, 592, 624, 627, 628

- Meyer, Heinrich August Eduard, 691, 777
 Meyer-Lübke, Wilhelm, 532, 777
 Mgharfaoui, Khalil, 426
 Miami-Illinois language, 599, 723
 microstructure,
 Arabic, 160, 167, 177–8, 181, 418, 422, 429
 Australian languages, 699
 Chinese, 54, 55–6, 57, 60, 63, 122, 126–8, 209, 213, 320, 616
 Egyptian, 45
 English, 485, 491
 French, 305, 537
 German, 463
 Greek, 86, 92, 104, 249, 256, 264
 Hebrew, 190, 198, 441
 Italian, 538
 Japanese, 209, 213, 342, 343, 344
 K'iche, 593
 Korean, 352
 Latin, 90, 94, 275, 289, 294
 Mesoamerican languages, 593
 Mesopotamian, 12, 21–2, 25, 29, 31
 Pali, 76, 144
 Sanskrit, 70, 131
 Scots, 514
 Spanish, 543
 Swahili, 671
 Tagalog, 623
 Tamil, 150
 Turkish, 246, 374, 376
 Wolof, 666
 Zulu, 677
Middle English Dictionary, 506, 514
 Mielcke, Christian Gottlieb, 455, 777
 Mikalja, Jakov, 448, 777
 Miklošič, Franc, 450, 777
 Millcayac language, 556, 723
 Milton, John, 489, 491, 497
 M'n language varieties, 319–20, 326–7, 329–30, 331–2, 615–16, 624, 626, 724
 Minsheu, John, 307, 623, 777
 Minucius Pacatus, 96
 Mirzā Khān, 381, 777; *see also* *Tuḥfat al-Hind*
 Mischlich, Adam, 669, 777
 Mishnah, 183–5, 187, 190, 195, 200, 201
 Mishor, Mordechai, 436
 Mishra, B. V., 399
 missionary lexicography, 5, 149, 153, 214, 217, 235, 243, 317, 319, 325–8, 329–32, 349–51, 360–2, 364, 368, 396–7, 400, 402–4, 555–705
 Mistral, Frédéric, 550, 777
 Mixe language, 583, 594, 724
 Mixtec languages, 580, 581, 584–5, 592, 594, 595, 724
 Moeris, 261–2, 777
 Moḥammad Nakhjawānī, 225, 778
 Mohawk language, 599–600, 724
 Molesworth, James Thomas, 405, 406–7, 778
 Molina, Alonso de, 580–2, 583–4, 585–6, 587–9, 592, 595, 623, 778
 Moliner, Maria, 543, 778
 Molino, Giovanni, 243–4, 778
 Molnár, Albert Szenci, 717
 Monet, Philibert, 305, 778
 Mongolian language, 180, 205, 206, 222, 238–40, 724
 Mono language, 604
 Montagnais language, 599, 601, 724
 Montenegrin language, 454
 Montgomery, Michael B., 515, 522, 778
 Monzón, Cristina, 585
 Moon Seyoung, 354, 778
 Moore, Bruce, 528
 Moore, George Fletcher, 683, 778
 Morais Silva, Antônio de, 545–6, 551, 778
 Moravian missionaries as lexicographers, 603, 763, 769, 791, 804
 Morgan, William, 602
 Morice, Adrien-Gabriel, 601, 778
 Morohashi Tetsuji, 335, 347–8, 778
 Morrison, Robert, 327–9, 330–1, 778
 Morrison, William Thomas, 330, 778
 Morrobolam language, 687, 724
 Moschopoulos, Manuel, 262
 Moser, Edward W., 604, 778
 Moser, Mary Margaret Beck, 604, 778
 Moth, Matthias, 312–13, 779
 Moxo language, 561, 563, 568, 569, 573, 724
 Mtuze, Peter Tshobisa, 679, 680, 779
 Mu 'arriḡ al-Sadūsī, 165, 779
 al-Mufaḍḍal al-Ḍabbī, 165, 779
 Muḥamed Hevāi Uskufī, 241, 779
 Muḥammad Ḥusayn Tabrīzī, 233, 779; *see also* *Burhān-i Qāṭi'*
 Mühlenbach, Karl, 459
 Mu'in, Muḥammad, 386, 779
 Muka, Arnošt, 454, 779
 Mukhliṣ, Ānand Rām, 379, 779
 Múlji, Karsandás, 398, 779
 Mullawillaburka, 779
 Muller, Jean-Claude, 558
 multi-word expressions. *See also* proverbs
 Arabic, 168
 Australian languages, 699
 Aymara, 566, 574–5

- Carib, 559
 Chinese, 221–2, 321, 324, 326, 328–9, 359, 615
 Cumanagoto, 562, 571
 East Slavic, 448
 Egyptian, 40, 47
 English, 488, 493, 495, 516, 524
 German, 311
 Greek, 93, 97–8, 100, 103
 Guarani, 564, 576
 Gugu-Yalanji, 703
 Gujarati, 399
 Hausa, 668
 Hebrew, 195, 197
 Igbo, 668
 Japanese, 215
 K'iche', 593
 Kaurna, 691
 Konkani, 646
 Korean, 352
 Latin, 94, 98, 281, 293, 295, 307, 308
 Latvian, 457
 Lithuanian, 456
 Malay, 646
 Malayalam, 646
 Manchu, 206
 Mapudungun, 565
 Matlatzinca, 593
 Mesopotamian, 21, 24, 25
 Mongolian, 205
 Mutsun, 598
 Narragansett, 602
 Nyulnyul, 694
 Old High German, 281
 Persian, 224, 225, 379, 382
 Portuguese, 557
 Quechua, 556, 572
 Ramindjeri, 691
 Russian, 452
 Shona, 676
 South Asian languages, 400, 646
 Spanish, 571, 573, 623, 631
 Swahili, 671
 Tagalog, 623, 624, 631
 Tamil, 402
 Tangut, 203
 Tupi, 557
 Turkic languages, 226
 Turkish, 242–3, 246, 367, 368, 373, 374, 375–6
 Urdu, 400, 407–8
 Western Apache, 613
 Worora, 696
 Yoruba, 667
 Zulu, 677
 Munro, Pamela, 597
 al-Murādī, 168, 779
Murgud (Mesopotamian wordlist), 15
 Murko, Anton, 450, 779
 Mummellius, Joannes, 295, 447, 779
 Murray, James Augustus Henry, 488, 492, 494–6, 499, 512, 518, 527, 780
 Muru, Cristina, 647
 Muruwari language, 704, 724
 Muskogean languages, 724
 Muskogee language, 603, 724
 Muṣṭafā, Ibrāhīm, 418, 428
 Musurus, Marcus, 251, 260
 Müteferrika, Ibrāhīm, 366
 Mutis, Celestino, 557
 Mutsun language, 598, 724
 Mužar, Andrija, 449
 Mymer, Franciszek, 447, 780
Nabritu (Mesopotamian wordlist), 15, 28, 29, 31, 34
 Na-Dene languages, 606, 724
 Nágera Yanguas, Diego de, 583, 594, 780
 Nahuatl language, 580–2, 583–8, 589, 592, 595, 623, 725
 Naish, Constance, 604
 Nakhat, 407–8, 780
 names. *See* personal names; place-names
 Naoumides, Mark, 85
 Narragansett language, 602, 725
 Nascimento, José Pereira do, 664, 780
 Nathan ben Yehiel, 195, 780
 Nathan, David, 687
 nationalism, 321, 349, 352, 364, 390, 393, 402–3, 410, 449, 451, 464, 465, 480, 483, 490
 Nauck, August, 262
 Navajo language, 601, 606, 611, 725
 Nayanappamutaliyār, 156
 Nda language, 673, 674, 680, 725
 Ndebele language, 675, 725
 Nebrija, Elio Antonio de, 286, 301–2, 306, 541, 545, 562, 563, 565, 576–7, 580, 584–7, 591–2, 593, 595, 622–5, 626, 630–1, 646, 780
 Neira, Alonso de, 556, 780
 Nekes, Hermann, 698, 704, 780
 neologisms
 Arabic, 160, 181, 415, 421, 425, 427–9
 Chinese, 317, 322, 324, 337, 364
 Czech, 451
 English, 424, 486, 498, 500, 506, 508, 527
 German, 470
 Hebrew, 187, 432, 439
 Italian, 538–9

- neologisms (cont.)
- Japanese, 340, 345
 - Latvian, 459
 - Lithuanian, 456
 - Mesopotamian, 27
 - Nahuatl, 595
 - North Germanic languages, 473
 - Pali, 144
 - Persian, 377
 - Romanian, 547–8
 - Sanskrit, 135
 - South American languages, 576
 - South Asian languages, 402
 - Spanish, 541
 - Turkish, 373, 375
 - Ukrainian, 453
 - Umbundu, 665
 - Vietnamese, 629
 - Zapotec, 595
- Nepali language, 641
- Nesselmann, Georg Heinrich Ferdinand, 456, 780
- Newman, Paul, 669
- Newman, Roxana Ma, 669
- newspapers as sources, 399, 412, 467, 468, 470, 497, 499, 500, 506–7, 543
- Ngag dbang 'jig rten dbang phyug grags pa'i rdo rje, 148, 780
- Ngamini language, 692, 725
- Ngarrindjeri language, 691–2, 725
- Ngô Thi Nhậm, 359, 780
- Ngumbarl language, 699, 725
- Nguni languages, 481
- Nguyễn Văn Minh, 365
- Nicot, Jean, 303, 305, 780
- Niedermann, Max, 457, 781
- Niedźwiedzki, Władysław Marcin, 452
- Niger-Congo languages, 668, 725
- Nighaṇṭu*, 68, 69–71, 72, 390
- Nihon kokugo daijiten*, 346, 349
- Nilambikai Ammaiyar, T., 402, 781
- Nilo-Saharan languages, 725
- Nimanburru language, 699, 725
- Niṣāb al-ṣibyān*, 391
- Niṣanyan, Sevan, 376–7, 781
- Nizolius, Marius, 294, 781
- Nlemvo, 661, 680, 781
- Nobili, Roberto di, 154, 637, 643
- Noceda, Juan José de, 614, 617, 623, 625, 629–30, 781
- Nogueira, Rodrigo de Sá, 673, 680, 781
- Nokele, A. B. B., 681
- Nonius Marcellus, 95, 781
- Noongar language, 683, 694, 725
- Norman, K. R., 76, 144
- Norn language, 476, 725
- Norwegian language, 312, 472–3, 475–6, 725
- Nowell, Laurence, 510, 781
- Nubian languages, 244, 726
- Nunggubuyu language, 686, 688, 703, 726
- Nxumalo, Otty Ezrom Howard Mandlakayise, 677, 680–1, 781
- Nyangumarta language, 704, 726
- Nyembezi, Cyril Sibusiso, 677, 680–1, 781
- Nyika language, 670, 726
- Nyikina language, 699, 726
- Nyulnyul language, 685, 694–5, 697, 699, 726
- Nyulnyulan languages, 695, 699, 726
- Nyungwe language, 673–4, 726
- Oates, Lynette, 703, 704
- Oates, William J., 703
- Oblates as lexicographers, 598, 601, 770, 778, 783, 791
- obscene words. *See* taboo and obscene words
- Occitan language, 535, 550, 726
- Ó Cléirigh, Míchél, 278, 313, 781
- 'O'Davoren's glossary', 278, 313
- official languages, 42, 45, 161, 318, 370, 384, 387, 388, 399, 402, 407, 409, 411, 414, 432, 435, 452, 454, 480–1, 636
- Officials* (Mesopotamian wordlist), 17
- Ogilvie, John, 399
- O'Grady, Geoffrey, 686
- Oguz languages, 239, 733
- Ojibwa language, 599, 601, 604, 726
- Oldendorp, C. G. A., 717
- Old Prussian language, 454, 727
- Oliver, Juan, 629
- Ölmez, Mehmet, 226
- Olmos, Andrés de, 580, 583, 781
- Ó Muirithe, Diarmaid, 515, 781
- Onions, Charles Talbut, 494, 519, 782
- Onkochishinsho*, 212–13
- online dictionaries. *See* electronic dictionaries
- Onofre, 558
- Onondaga language, 599, 603, 727
- O'odham language, 600, 604, 727
- Opata language, 600, 727
- Oprea, Ioan, 548
- oral transmission of dictionaries, 156, 391
- Ordbog over det Danske Sprog*, 474
- Ordbok över Svenska språget*, 474–5
- Orion of Thebes, 256, 257, 782
- Orlović, Jerolim, 449
- Orman, Uzzi, 431

- Oromo language, 670, 727
 orthography
 African languages, 663, 679–80
 American languages, 604, 605–6
 Arabic, 427
 Arremte, 693
 Assamese, 403
 Australian languages, 683, 684, 698, 701
 Awabakal, 689
 Chinese, 52, 60–1
 Dutch, 479
 English, 494, 517
 French, 533
 German, 462–3, 465–6, 468–9, 471
 Greek, 103, 251, 257, 262
 Gunin / Kwini, 695, 700–1
 Inuktitut, 601
 Inupiaq, 604
 Japanese, 344
 Kikongo, 661
 Kimbundu, 663
 Korean, 351, 352–4, 355
 Luxembourgish, 483
 Mesopotamian, 26
 Norwegian, 476
 Nyulnyul, 694, 695
 Portuguese, 546
 Romance languages, 531
 Romanian, 549
 Sanskrit, 131, 145
 Serbian, 450
 Shona, 675
 Slovene, 450
 South Asian languages, 636
 Spanish, 541
 Swedish, 472
 Umbundu, 665
 Ungarinyin, 703
 Vietnamese, 361, 363, 365
 Western Apache, 607, 610–11
 Xhosa, 678
 Zulu, 676
 Oscan language, 92
 Osing, Jürgen, 43
 Ōta Zensai, 343, 782
 Otomaco language, 556, 727
 Oto-Manguean languages, 580, 587, 727
 Otomi languages, 581, 623, 727
 Ōtsuki Fumihiko, 343, 782; *see also* Genkai
 Oudin, César, 302–3, 541, 782
 Oxford Arabic Dictionary, 417, 423, 429
 Oxford dictionaries, 399, 498, 501, 505
 Oxford English Dictionary, 410, 475, 478, 484, 488, 492–503, 505–8, 510, 512–13, 518–19, 521, 524, 527, 528, 536, 678
 Ožegov, Sergej Ivanovič, 453, 782
 Pādalīptācārya, 141
 Pahl, Herbert Walter, 679, 782
 Pali language, 65, 73–8, 130, 143–4, 727
 Pallottines as lexicographers, 695, 698, 702, 704, 745, 776, 780
 Palmer, E., 608, 610
 Palsgrave, John, 304, 306–7, 782
 Pama-Nyungan languages, 2, 699, 703–4, 727
 Pampango language, 618, 727
 Pamphilus of Alexandria, 102, 251, 782
 Pamulkan, Mabel, 704
 Pančev, Todor, 450
 Pāṇini, 71, 145, 782
 Papias, 268, 272–4, 275, 281, 285–6, 288–9, 782
 paratexts,
 alphabetical index, 235, 246, 274, 296, 297, 311, 367, 374, 395, 428, 521, 569, 574, 617, 620
 appendix, 361, 439, 589, 617, 619, 667, 668, 671
 bibliography, 466, 493, 524
 colophon, 14, 31, 244, 407
 encomiastic verses, 656
 epigraph, 302
 frequency graphs, 523
 frontispiece, 270
 grammar, 49, 192, 233, 243, 245, 311, 361, 368, 490, 555, 612, 638, 667, 670, 691
 guidance for language learners, 575
 guidance for users, 136, 233, 564, 572, 676
 history of lexicography, 157
 history of the language, 490
 history of the speakers of a language, 666
 information about macrostructure, 561, 585
 information about sources, 60, 98, 136, 148, 162, 178, 233, 251, 258, 352, 405, 563, 625, 660, 676
 introductory benediction, 133–5, 142–3
 list of contributors, 521, 524
 maps, 227, 230, 521
 non-alphabetical index, 62, 100, 218, 220, 221–2, 338, 347–8
 postface, 59, 61, 575

- paratexts (cont.)
 preface by lexicographer, 39, 54, 62, 97, 99,
 101, 102–4, 123, 125, 135, 143, 154, 155, 162,
 175–6, 177, 178, 199, 226, 232–3, 251, 272,
 275, 282, 295, 298, 332, 351, 352, 354, 362,
 367, 371, 372, 375, 385, 393, 397, 403, 409,
 421, 447, 458, 466, 489–90, 491, 493, 501,
 518, 560, 565, 572–3, 575, 585, 588, 591, 617,
 625, 635, 640, 649, 661, 676–7
 preface not by lexicographer, 175, 221, 260,
 309, 412, 667, 676
 rhyme tables, 116–17, 128–9, 212
 supplement, 125, 233, 297, 302, 374, 677
 table of contents, 99
 testimonials, 409
 title page, 297, 308, 400, 591, 599, 622, 649,
 667, 672
 Parkinson, Dilworth, 422–3, 429
 particles, 142, 163, 168, 170, 180, 197, 565–6,
 570–1, 572–6, 578, 584, 590–1, 611, 628, 676,
 691–2
 Pashto language, 387, 412, 727
 Pasor, Georg, 298, 782
 Passamaquoddy language, 604, 728
 Passow, Franz Ludwig Carl Friedrich, 468,
 492, 782
 Paul the Deacon, 90, 269, 783
 Paul, Hermann Otto Theodor, 469, 783
 Paulinus a Sancto Bartholomaeo, 645, 646,
 650, 654, 783
 Pausanias, 96
 Peacock, Edward, 512, 783
 Peciar, Štefan, 452, 783
 Peile, Anthony, 702, 704
 Pellas, Sauveur André, 726, 783
 Pelleprat, Pierre-Ignace, 559, 560, 783
 Pellicer, Dora, 594
 Percy, Algernon, 419
 Perera, S. G., 635
 Perotti, Niccolò, 292–4, 783
 Perrin, James, 676, 680, 783
 Perry, William, 490, 783
 Persian language, 143, 166, 180, 185, 196, 223–5,
 226, 228–37, 238–40, 299, 366, 367, 369–73,
 374, 377–87, 391–4, 395, 397–8, 399–400,
 410–11, 434, 636, 651, 728
 personal names,
 Awabakal, 689
 Chinese, 52, 204, 322
 Egyptian, 38, 43
 Greek, 100, 257
 Hebrew, 277, 437
 Japanese, 344
 Latin, 91
 Lithuanian, 456
 Mesopotamian, 23–4, 29
 Pali, 143
 Prakrit, 142
 Tamil, 150, 152, 156
 Tibetan, 146
 Western Apache, 610
 Peters, Wilhelm Karl Hartwig, 672, 674, 783
 Petitnicolas, Michel Alexandre, 351, 783
 Petitot, Émile-Fortuné-Stanislas-Joseph,
 601, 783
 Petrarch, 274, 304
 Petrocchi, Policarpo, 538, 784
 Petter, Rodolphe Charles, 603, 784
 Pfeiffer, Rudolf, 87
 Pfuhl, Christian Traugott, 454, 784
 'Phags-pa Lama, 203
 Phạm Đình Hổ, 360, 784
 Philitas the Coan, 87, 784
 Phillips, Edward, 308, 784
 Philological Society, 494, 513
 philology and lexicography
 African languages, 672, 677
 Arabic, 159–60
 Bantu and Indo-European, 673
 Chinese, 51, 109, 126, 127–8
 Coptic, 49
 Danish, 474
 Dravidian, 401
 Finnish, 312
 German, 465
 Germanic languages, 482
 Hebrew, 186–7, 189, 190, 198–9
 Indo-European, 396–7, 401, 456, 465, 492,
 531, 656
 Japanese, 343
 North Germanic languages, 476
 Persian, 233, 381
 Romance, 531, 532
 Romance languages, 535
 Slovene, 450
 South Asian languages, 388, 413, 657
 philosophical terms, 125, 136, 146, 171, 254, 255,
 256, 282, 284
 Philostratus of Athens, 99
 Philoxenos, 89, 92
 Photius, 252–3, 784
 Phrankopoulos, George, 262–3
 phrases. *See* multi-word expressions
 Phrynichus, 96–7, 261–2, 784
 Pianzola, Bernardino, 368–9, 784
 Pichardo y Tapia, Esteban, 544, 784

- Pickering, John, 602
 pidgins, 682, 704
 Piekarski, Edward, 230, 784
 Pigafetta, Antonio, 710, 722
 Pigneau de Behaine, Joseph Georges Pierre, 361, 784
 Pima Bajo language, 600, 728
 Pinet, Pierre-François, 599, 785
Pinkalam, 149, 150–3, 156
 Pinnock, Osbern, 273, 283, 290, 785
 Pintupi/Luritja language, 702, 728
 Pitjantjatjara language, 697, 701, 728
 Pittau, M., 730
 place-names,
 Arabic, 171, 178
 Australian languages, 693
 Awabakal, 689
 Chinese, 53, 61, 322
 Egyptian, 38, 43
 English, 515, 522, 526
 Greek, 257, 263
 Hebrew, 195, 196, 437
 Japanese, 344, 619
 Lithuanian, 456
 Mazahua, 594
 Mesopotamian, 12, 24
 Mixtec, 594
 Pali, 144
 Western Apache, 612
 Placidus, 270, 272, 785
 plagiarism, 298, 327, 400, 408, 471, 534, 638
 Plantin, Christophe, 310
Plants (Mesopotamian wordlist), 16–17
 Plato, 97, 256
 Platts, John Thompson, 385, 785
 Plautus, 273, 294
 Pleteršnik, Maks, 450, 785
 Plocher, Johannes, 609, 611
 Podolsky, Baruch, 441, 785
 poetic language,
 Arabic, 160, 162, 187
 Chinese, 54, 114, 115, 208, 211, 324
 German, 463
 Greek, 87, 93, 97, 257
 Hebrew, 187, 197
 Japanese, 210, 212, 343, 346, 627, 632
 Persian, 224, 228–9, 232, 233, 234, 377–9
 Prakrit, 142
 South Asian languages, 390–1, 413
 Tamil, 81, 151–2, 154, 639, 643
 Turkic languages, 240
 Urdu, 408
 Welsh, 278
 Poirier, Claude, 537
 Polabian language, 449, 728
 Polish language, 245, 246, 288, 295, 300–1, 446–7, 448, 451, 452–3, 455, 457–8, 477, 728
 Politis, Linos, 253
 Pollux, Julius, 87–8, 99–101, 261, 785
 Polomé, Edgar, 671, 672
 polyglot dictionaries and wordlists, 180, 206, 241, 243, 246, 293, 299–301, 302, 307, 447–8, 449, 477, 619, 624, 657, 672–3, 690, 692, 698–700
 Polynesia, languages of, 689
 polysemy, 57, 58, 60, 63, 88, 92, 94, 103–4, 132, 150, 152–3, 194, 198, 251, 253, 391, 488
 Pomey, François-Antoine, 303, 645, 785
 Pompeius Festus, Sextus, 90, 269, 785
 Pond, Samuel, 604, 785
 Pons, Jean-François, 645, 785
 Pontani, Filippomaria, 257, 262, 264
 Pootchemunka, Jennifer, 704
 Pope, Alexander, 486, 489
 Poqomchi' language, 582, 728
 Portuguese language, 214, 302, 326, 361, 396, 545–7, 551–2, 557–8, 616, 618, 620, 622, 624–5, 627, 634, 640–2, 647–9, 652, 655, 660, 662–5, 673–4, 680, 728
 Potier, Pierre-Philippe, 599, 785
 Pott, August, 729, 785
 Prakrit languages, 65, 73–4, 78–9, 130, 141–3, 729
 Pratt, Terry Kenneth, 524, 786
 Presbyterian missionaries as lexicographers, 332, 603, 696, 742, 754, 773, 773, 785, 788, 799, 803
 prescriptivism, 97, 118, 197, 229, 261–2, 320, 334, 352, 378, 410, 432, 436, 463–4, 467, 485, 487, 496, 501, 503, 506, 531, 544, 551, 632, 679
 price of dictionaries, 297, 332, 334, 344, 393, 396, 443, 469–70, 472, 495, 498
 printing, 4, 112, 156–7, 193, 205, 213–14, 244, 245, 251, 260, 268, 272, 276, 282, 283, 291–2, 301, 313, 327–8, 332, 340, 366, 390, 393–5, 407, 483, 486, 560–1, 620, 650–1
 difficulties and setbacks, 245, 300, 328, 332, 394, 395, 419, 440, 456, 560, 562, 587, 598, 647, 650, 651–2
 lithographic, 378, 394, 395
 xylographic, 111
 Priscian of Caesarea, 98, 590
 Proença, Antão de, 153–4, 158, 620, 640, 647, 648–9, 651, 652
Promptorium parvulorum, 283, 290, 306, 308

pronunciation

- Arabic, 169, 426
- Assamese, 403
- Awabakal, 689
- Chinese, 58, 59, 60, 62, 63, 65, 112–18, 124, 127–9, 204, 207, 208–9, 214–15, 219–20, 221, 320–2, 325, 328, 330, 334–5, 357, 362, 364, 619
- Egyptian, 44
- English, 485, 490, 494, 495, 516–17, 528
- French, 352
- Greek, 254
- Japanese, 207, 209, 210, 214, 341, 344, 345, 619
- Kimbundu, 664
- Korean, 219–20, 221, 351, 352
- languages of the Philippines, 628
- Lithuanian, 456
- Mongolian, 205
- Sanskrit, 67, 403
- Sumerian, 15, 19
- Tagalog, 624
- Tamil, 649
- Turkic languages, 238
- Turkish, 242, 246, 367, 369, 373
- Vietnamese, 362–3, 364
- Western Apache, 612
- Wolof, 666
- Xhosa, 679
- Yoruba, 667
- Zulu, 676–7

proverbs

- Arabic, 161, 162, 163, 164–5, 178, 426
- Greek, 98, 103, 251, 254
- Guarani, 577
- Hindustani, 409
- Kimbundu, 663
- Konkani, 646
- languages of the Philippines, 631
- Matlatzinca, 593
- Mesopotamian, 23, 25
- Shona, 676
- Tagalog, 624
- Turkic languages, 226
- Urdu, 407–8
- Vietnamese, 632
- Yoruba, 667
- Zulu, 677

purism

- Arabic, 160, 176, 177, 187
- Czech, 451
- Dravidian, 401
- Dutch, 310, 480
- English, 486–7, 489
- German, 462, 465, 466

- Greek, 88, 93, 96–7, 104, 250, 252, 255, 260–3
- Hebrew, 434, 444
- Hindi, 410
- Italian, 538–9
- Japanese, 348
- Latin, 293, 295
- Latvian, 459
- Lithuanian, 456
- Persian, 233, 382
- Russian, 451
- Sanskrit, 73
- South Asian languages, 393, 413
- Spanish, 540–2, 544, 552
- Tamil, 402
- Telugu, 393
- Turkic languages, 229, 239
- Turkish, 371, 373
- Zulu, 676

- Puruṣottamadeva, 134, 137, 139–41, 145, 786
- Putisillo, Mikhail Pavlovich, 350, 786

- Qafisheh, Hamdi, 426

- al-Qālī, 175–6, 786

- Qazvīnī, Muḥammad Yahyā bin Muḥammad Shafīq, 380

- Qièyūn, 65–6, 112–15, 117

- Quay, Margaret, 612

- Quechua language, 555–6, 560–1, 563–4, 565–6, 569, 572, 577, 729

- questionnaires, 521, 525, 684, 695

- Quintana, Augustín de, 583, 594, 786

- Quintão, José Luís, 673

- Quirk, Randolph, 502

quotations in dictionaries

- Arabic, 165, 175, 178, 418, 420, 429

- Assamese, 404

- Aymara, 561

- Chinese, 56, 57, 60, 63–5, 122, 126, 320, 328–9, 335, 348

- English, 485, 487–98, 500, 502, 503, 507, 510, 513, 519, 521, 523–4, 526, 527–9

- European languages, 488, 493

- French, 305, 534–5

- German, 461, 463, 466–8, 470

- Greek, 86, 93, 98, 104, 248, 251, 254, 256, 258, 260–4

- Guarani, 563

- Hebrew, 436

- Italian, 304, 538, 540

- Japanese, 209, 216, 342–3, 344, 346, 632

- Latin, 90, 275, 293–4, 304

- Manchu, 206

- Mesopotamian, 15

- Pali, 144
 Persian, 224, 229, 234, 377, 379, 381, 383
 Polish, 451, 453
 Portuguese, 545–7
 Sanskrit, 136
 Scots, 514
 Spanish, 541, 542–3
 Tagalog, 630
 Turkic languages, 229
 Turkish, 374–5
 Ukrainian, 453
 Ungarinyin, 703
 Urdu, 408–9
 Zulu, 677
 Qur'ān, 159–60, 164–5, 179, 232, 414–16
 Quṭrub, 169, 786

 Rada Cruzat, Martín de, 615, 786
 Radak. *See* Kimchi
 Radcliffe-Brown, Alfred R., 685
 Radloff, Friedrich Wilhelm, 230, 236, 786
 Rāhulaka, 141
Rakuyōshū, 214, 619–20, 626
 Rale, Sébastien, 599, 786
 Rāmamūrti, G. V., 394
 Ramanujaswami, P. K., 143
 Ramaswamy, Sumathi, 402
 Ramindjeri language, 691, 729
 Ramsey, S. R., 319
 Ramson, William Stanley, 528, 786
 Rangīn, 408
 Raphelengius, Franciscus, 299, 786
 rare words
 Arabic, 160, 163, 422
 English, 496, 497, 504
 German, 461
 Greek, 84, 87, 93–4, 103, 104
 Hebrew, 187
 Japanese, 345, 346
 Latin, 91, 96, 273
 Latvian, 459
 Mesopotamian, 15, 27, 34
 Persian, 224, 377
 Sanskrit, 69, 135
 Turkish, 371
 Vietnamese, 360
 Rashi, 185, 189, 198, 286, 786
 Rask, Rasmus, 465, 492
 Rastell, John, 307–8, 787
 Rasulid Hexaglot, 180
 Ravius, Christian, 234–5, 787
 Ray, John, 511, 513, 787
 Rāyamukuṭa, 134

 Raynouard, François Just Marie, 550, 787
 Read, Allen Walker, 517, 519, 520, 787
 Reagan, Albert B., 609–10
 recitation and lexicography, 67, 80, 115,
 128, 149
 Reda, Yussof Mohamed, 425
 Redhouse, James William, 372, 375–6, 787
 Reece, Laurie, 703
 regionalisms. *See* dialects and regionalisms
 register
 English, 487, 501, 504, 505
 Greek, 93, 97, 260
 Japanese, 215
 Korean, 350
 Portuguese, 546
 Spanish, 542
 Turkish, 368
 Reichmann, Oskar, 1
 Reig, Daniel, 424
 Reitzenstein, Richard, 257, 258–9
 Relandus, Hadrianus, 645
 Restivo, Pablo, 557, 563, 568, 574, 576, 787
 Reuchlin, Johannes, 282, 299, 447, 787
 Reuther, Johann Georg, 692, 787
 Rézeau, Pierre, 537, 787
 Rheede, Hendrik van, 642
 Rhodes, Alexandre de, 361, 620, 622, 626, 628,
 629, 632, 787
 Rhodes, Richard, 604
 rhyming dictionaries. *See* macrostructure,
 other: by rhyme
 Ribeiro, Diogo, 646
 Ricardo, Antonio, 561
 Ricci, Matteo, 302, 326, 616, 620, 622, 625,
 632–3, 787
 Richards, Eirlys, 703
 Richardson, Charles, 490–2, 518, 787
 Richardson, John, 384–5, 788
 Richelet, Pierre, 305–6, 533, 534, 541, 788
 riddles, 226, 237, 631, 663
 Ridel, Félix-Clair, 350, 788
 Ridley, Mark, 307, 448, 788
 Ridley, William, 690, 788
 Ridruejo, Emilio, 573
 Rifat, Kilisli, 227, 788
 Riggs, Stephen Return, 603, 788
 Rimanggudinhma language, 687, 729
 Rinaldini, Benito, 600, 788
 Rincón, Antonio del, 587
 Rishi, Weer Rajendra, 412
 Ritchie, Joan, 528
 Rivero, Juan, 556, 788
 Rivola, Francesco, 708, 788

- Robert, Paul, 535, 788
 Robert dictionaries, 535, 549
 Robinson, Julia, 528
 Robinson, Mairi, 514
 Roboredo, Manuel de, 660, 680, 788
 Roebuck, Thomas, 384
 Roget, Peter Mark, 3, 399
 Romance languages, 277, 284, 285, 319, 369, 530–2, 535, 547, 550, 552, 729
 Romani language, 729
 Romanian language, 547–9, 551–2, 729
 Romansh language, 551, 729
 Ronga language, 672–3, 680, 729
 Roquete, José Ignacio, 674
 Rosenbach, Zacharias, 298
 Roth, Heinrich, 645, 788
 Roth, Rudolf von, 69, 399, 789
 Rothwell, William, 536
 Routledge dictionaries, 498
 Ruby, C., 610
 Ruddiman, Thomas, 510, 789
 Rufus of Ephesus, 98, 789
 Ruggieri, Michele, 302, 326, 616, 620, 622, 625, 633, 789
 Ruhig, Philipp, 455, 789
 Ruiz Blanco, Matias, 559, 561, 567, 570–1, 789
 Ruiz de Montoya, Antonio, 557, 558, 561, 563–4, 568, 570, 574, 575–6, 577, 622, 789
 Rumphius, Georg Everhard, 638, 789
 Russian language, 216, 223, 307, 350, 387, 412, 433, 438–9, 440–2, 448, 451–3, 458–9, 477, 547, 729
 Ruthenian language, 448, 730
 Saadia Gaon, 179, 184, 185–8, 191, 195, 197, 198–9, 200, 789
 Sachse, Frauke, 586, 593
 Sa' dī, 377
 Ṣafīpūrī, 'Abd al-Raḥīm, 381
 al-Ṣaḡānī, 177, 181, 789
 Sagard, Gabriel, 599, 789
 Ṣahbā'ī, 408
 al-Ṣāḥib bin 'Abbād, 176, 789
 Said, Edward, 419
 Saikia, S., 404
 Șăineanu, Lăzar, 548, 790
 Sakaeda Takei, 347, 790
 Salakṣa, 392, 790
 Sale, Flaminio da, 729, 790
 Salesbury, William, 307, 790
 Salinan language, 598, 730
 Salle de l'Etang, Simon Philibert de la, 559, 790
 Salvá, Vicente, 542–3, 790
 Salvado, Rosendo, 693, 695, 790
 Salys, Antanas, 457, 790
 Samuels, David, 613
Samun seonghwi, 221
Samun tonggo, 220–1
 San Antonio, Francisco de, 618, 621, 625, 790
 San Buenaventura, Pedro de, 617–18, 620–2, 623–4, 625, 628–31, 790
 Sánchez Méndez, Juan Pedro, 562
 Sandefur, John, 705
 Sandefur, Joyce, 705
 Sanders, Daniel, 467–9, 471, 474, 790
 Sanders, Mark, 677
 Sanders, William Henry, 664, 680, 790
 San Lúcar, Pedro, 614, 617, 623, 625, 629–30, 791
 Sanskrit language, 65, 67–74, 79–81, 82, 112, 127–8, 130–48, 151, 154, 205, 216, 390–2, 394, 396–7, 399–403, 410, 411, 413, 531, 639, 643, 645, 647, 656, 730
 Santo Tomás, Domingo de, 555, 560–1, 563–4, 566, 569, 577, 791
 Santra, 90
 Sanz de Galdeano, Seraphim, 701, 791
 Sapir, Edward, 602, 611
 Sara, S. I., 171
 al-Saraquṣṭī, 170, 791
 Sardinian language, 539, 551, 730
 Śarmā, Vajñāth, 406, 791
 Sarup, Lakshman, 69, 70
 Sarvānanda, 134
Saseong tonggo, 219–20
 Śāśvata, 134, 139, 141, 791
 Saxton, Dean, 604
 Saxton, Lucille, 604
 al-Ṣaybānī, Abū 'Amr, 161, 163, 172, 175, 791
 Scapula, Johannes, 298, 791
 Schele de Vere, Maximilian Rudolph, 518, 791
 Schiller, Johann Christoph Friedrich von, 467, 470
 Schleicher, August, 456
 Schmick, Johann Jacob, 602, 791
 Schneider, Edgar, 526, 528
 Schneider, Lucien, 601, 791
 Schoenhals, Louise, 600
 Schoknecht, Carl H. M., 692
 Schön, James Frederick, 668–9, 791
 Schoonees, Pieter Cornelis, 791
 Schottelius, Justus Georgius, 311, 791
 Schregle, Götz, 424, 792
 Schrevelius, Cornelius, 298, 792
 Schroderus, Ericus Johannis, 714, 792
 Schroue, Thomas, 448
 Schueren, Gerard van der, 310, 792

- Schultze, Benjamin, 657, 792
 Schumann, Christian Ludwig, 731, 792
 Schürmann, Clamor Wilhelm, 690–1, 792
 Schweickard, Wolfgang, 532, 552
 scientific terms, 317, 322, 355, 410, 412, 425, 428, 446, 470, 501, 539, 545, 548
 Scots language, 493, 510, 513–14, 661, 668, 730
 Scott, James, 351
 Scott, Walter, 497
Scottish National Dictionary, 514
 Se' adiah ibn Danan, 194
 Seaman, P. David, 605
 Seco Raymundo, Manuel, 543, 792
 Sedelmayr, Jacobo, 600, 792
 Seelakkhandha Mahathera, 145
 Segal, Moshe, 433, 439, 440
 Semitic languages, 14, 20, 45, 198–9, 235, 433, 636, 730
 Şemseddin Sami, 370–1, 790
 Sena language, 673, 730
 Senisio, Angelo, 276, 285, 792
 Senn, Alfred, 457, 792
 Serbian language, 450, 454, 730
 Serbo-Croatian language, 449, 454
 Seri language, 604, 730
 Servius, 268, 272, 792
Setsuyōshū, 212–13, 342
 sGom sde Nam mkha' rgyal mtshan, 146
 Shād, Muḥammad Pādshāh, 382, 386, 793
 Shahīdī, Muḥammad Ja'far, 386
 Shakespear, John, 405
 Shakespeare, William, 307, 489, 497, 502
 Shapiro, Feliks, 440, 793
 Shaul, David, 604
 Shaw, Graham, 394
 Shelomo ben Meborakh, 193
 Shelomo ben Samuel, 196, 793
 Sheridan, Thomas, 490, 793
 Sherwood, W. L., 608, 610
 Shì Yóu, 52, 793
 Shilling, Alison Watt, 525
Shiming, 53, 58–9, 65, 216
 Shin Gicheol, 354
 Shin Yongcheol, 354
 Shinmura Izuru, 345, 793; *see also* *Kōjien*
 Shinmura Takeshi, 345, 793
 Shintoism, 627, 632
Shǐzhòu piān, 51
 Shlomo Ben-Elia, 197
 Shōjū, 208, 793
 Shona languages, 674–6, 730
 Shona language (standard Shona), 675–6, 681, 730
Shorter Oxford English Dictionary, 498, 504
 Shū Xinchéng, 322, 793; *see also* *Cǐhǎi*
Shūbun inryaku, 211–12, 213, 214, 221
Shuōwén jiàngjiān, 120
Shuōwén jiǐzǐ, 51, 59–62, 63–4, 65, 118–24, 129, 615
 Sībawayhi, 159, 165–6, 169, 175, 793
 Sikhakhana, Jonathan Mandlenkosi A., 677
 Šilāṅka, 141
 Silva, Duarte de, 599, 618, 793
 Silva, Penny, 528
 Silvy, Antoine, 793
 Sim Uilin, 354, 793
 Sinclair, John McHardy, 503, 793
 Sinhalese language, 388, 641, 642, 646, 730
 Sennius Capito, 90
 Sino-Tibetan languages, 639, 731
 Siona language, 556, 731
 Siouan languages, 731
Sirāj al-Lughat, 378, 383
 Sirvydas, Konstantinas, 455, 457–8, 793
 Sisam, Kenneth, 494
Sishēng piānhǎi, 114, 120
 Sitārāmasvāmī, I. V., 395
 Sitjar, Buenaventura, 598, 794
 Si tu paṇ chen, 146, 147, 794
 Sivan, Daniel, 437
 Sivan, Reuben, 439
 Skeat, W. W., 512, 513
 Skinner, Stephen, 308, 794
 slang, 127, 375, 475, 500, 519
 Slavic languages, 241, 287, 445–8, 451–2, 547, 549, 731
 Sleeman, William, 381
 slips, 271, 409, 453, 457, 467, 470, 512, 611
 Slovak language, 449, 451–2, 731
 Slovene language, 448, 450, 454, 731
 Sluiter, Ineke, 256
 Smart, C., 608, 610
 Smithsonian Institution, 6, 603, 610
 Smith-Stark, Thomas, 565, 581, 635
 Solana, Alonso de, 582, 794
 solecisms, 94, 159, 163, 166
 Solomon ben Abraham, 198, 794
 Somner, William, 308, 794
 songs, 226, 409, 459, 577
 Sorbian language, 449, 454, 731
 Sotho-Tswana languages, 481
 Spanish language, 185, 194, 243, 272, 285–6, 291, 295, 300–3, 307, 336, 424, 429, 442, 477, 538, 540–5, 551–2, 555–96, 600–1, 612, 615–18, 620–6, 627–8, 629–31, 659–60, 696, 731

- spoken language in dictionaries
 Arabic, 190
 Assamese, 403
 Chinese, 115, 318, 328, 617, 627
 Danish, 312
 English, 499, 503, 513, 525
 Greek, 88, 98
 Icelandic, 312
 Italian, 538
 Japanese, 209, 343, 632
 languages of the Philippines, 631
 Latin, 104
 Latvian, 458, 459
 Lithuanian, 456
 Norwegian, 475
 Portuguese, 547
 Slavic languages, 448
 South Asian languages, 400
 Spanish, 544
 Tamil, 155, 158, 640
 Telugu, 641
 Turkic languages, 228, 229
 Turkish, 368–9, 371, 373
 Vietnamese, 361
 Western Apache, 612
 Yoruba, 667
- Sranan Tongo language, 477, 731
 Śrīdharasena, 135, 137, 139–40, 147–8, 794
 Stahl, Heinrich, 713, 794
 standards and standardization, 460, 474, 480, 510
 African languages, 670
 Afrikaans, 481
 American languages, 605
 Arabic, 427–8
 Catalan, 550
 Chinese, 52, 59, 109, 111–12, 114–15, 118, 129, 204, 211, 219, 317–18, 320–2, 325, 333–4
 Danish, 472, 475
 Dutch, 478, 479–80
 Egyptian, 39
 English, 484–508, 509, 516, 518, 526
 French, 533, 537
 Frisian, 482
 German, 461, 463–4, 465–6, 468–9, 471, 478
 Hebrew, 432, 435–6, 439, 444
 Italian, 305, 538–9
 Japanese, 340
 Korean, 219, 222, 352–4, 355–6
 Latvian, 459
 Luxembourgish, 483
 Norwegian, 475
 Persian, 387
 Polish, 453
 Portuguese, 545–6
 Romance languages, 531–2, 551–2
 Russian, 451, 453
 Serbo-Croatian, 449
 Slavic languages, 452
 Slovak, 451
 South Asian languages, 636
 Spanish, 541, 544
 Swedish, 472
 Tajik, 387
 Tibetan, 146
 Turkish, 373, 375
 Vietnamese, 363
 Yiddish, 482
- Steere, Edward, 671, 794
 Steffel, Matthäus, 600, 794
 Steinbach, Christoph Ernst, 461, 463, 794
 Steiner, Richard, 198
 Steingass, Francis Joseph, 234, 385, 794
 Steinitz, Wolfgang, 470, 794
 Stender, Gotthard Friedrich, 458, 795
 Stenton, Frank Merry, 515, 795
 Stephanus a SS Petro et Paulo, 649, 795
 Stephanus of Byzantium, 255, 263, 795
 Steuerwald, Karl, 376, 795
 Stewart, George Rippey, 522, 795
 Stieler, Kaspar, 311, 313, 795
 Stiernhielm, Georg, 312, 795
 Story, George Morley, 524, 795
 Story, Gillian, 604
 Strandberg, Elisabeth, 73
 Strehlow, Carl Friedrich Theodor, 693, 795
 Strehlow, Theodor George Henry, 685, 795
 Štúr, L'udovit, 451
 Sturz, Friedrich Wilhelm, 258
 Subhūticandra, 142, 145, 147–8, 796
 Sūda, 87, 102, 252, 253–5, 264
 Sueiro Justel, J., 621
 Suetonius Tranquillus, Gaius, 98, 796
 Sughdī, 223, 796
 Sumerian language, 11–35, 732
 Summer Institute of Linguistics, 603–4, 611, 686, 687, 698, 700–1, 702–3, 705
 Sūn Qiáng, 63
 Surūrī, Muḥammad Qāsim, 382, 796
 Svabo, Jens Christian, 714, 796
 Swahili language, 670–2, 680–1, 732
 Swedish language, 283, 311–12, 472–5, 732
 Swētlík, Jurij Hawštyn, 449, 796
 Swift, Jonathan, 486, 517
 Symeon, 259, 264, 796
 Synagoge, 248, 252–4, 257

- synonym lists, 2, 81
 Arabic, 161, 162, 164, 167
 Chinese, 51, 53–5, 57, 218
 Danish and Swedish, 311
 Egyptian, 43
 Greek, 99, 102, 263–4
 Hebrew, 197
 Mesopotamian, 15, 28–9, 33
 Pali, 75, 76–7, 143–4
 Prakrit, 141–2
 Sanskrit, 69, 71, 72, 131, 136–8, 139, 140, 645
 Tamil, 150, 151
 Telugu, 394
 synonyms, in dictionary entries, 246, 479,
 666, 675
 synonymy, 68–9, 70, 77–80, 99, 131–2, 166, 168,
 390, 394
 Syriac language, 166, 640, 732

 Taberd, Jean-Louis, 361, 362, 796
 taboo and obscene words, 98, 242, 409, 463,
 495–6, 500–1, 592, 631, 647–8
 Tabrizi, Muhammad Karīm, 383
 Tagalog language, 617–18, 620–1, 623–5,
 628–31, 732
 Tajik language, 387
 Talmud, 183–5, 189, 191, 195, 196, 200, 201
 Tamil language, 3, 6, 80–3, 148–58, 388, 393,
 396, 401–2, 620, 635, 639–41, 643–5, 647–9,
 650–2, 655, 657, 732
 Tangut language, 202–3, 732
 Tanḥum Yerushalmi, 195, 796
 Tanikawa Kotosuga, 216, 342, 796
 Tāṇṭavarāyamutaliyār, 156
 Ṭapish Dihlavi, Mirzā Jān, 400, 407
 Taplin, George, 692, 796
 Tarahumara language, 600–1, 604, 733
 Tarascan language, 581, 582, 584, 585–6,
 588–90, 592, 733
 Targum. *See* Bible in Aramaic
 Tatar language, 244, 733
 Tauste, Francisco de, 559, 561, 562, 567, 796
 Taylor, Joseph, 400
 technical terms, 61, 98, 105, 282, 288, 304–5,
 307, 317, 322, 345, 355, 410–12, 427–9, 461,
 462, 464, 466, 470, 475, 489, 501, 523, 531,
 533, 538–9, 541, 545, 548, 576, 632
 Teichelmann, Christian Gottlob, 690–1, 796
 Telephus of Pergamum, 98, 797
 Telugu language, 390–1, 392–3, 394–5, 400–1,
 641, 656, 733
 Tench, Watkin, 683, 797
 Teng Shou-hsin, 337, 797

 Ten Raa, Eric, 687
 Tepehuan languages, 600, 733
 Terreros y Pando, Esteban de, 542, 797
 Teruel, Antonio, 660, 797
 Tew, Gwilym, 278, 797
 Texeira, Francisco, 650
 Teyvaccilaiyār, 82
 thematic ordering of entries. *See* macrostruc-
 ture, thematic
 Theodore of Tarsus, 279, 289
 Theon, 86, 797
 thesauruses, 2, 461
 Theurel, J. S., 361
 Thibault, André, 537, 797
 Thibert, Arthur, 601
 Thieberger, Nicholas, 686–8
 Thirrari language, 692, 733
 Thomas Magistros, 261–2, 265, 797
 Thomas, André Antoine, 534, 797
 Thomas, Mandy, 528
 Thoms, Peter Perring, 328
 Thoresby, Ralph, 513
 Thornton, Richard Hopwood, 518, 797
 Threlkeld, Lancelot, 689, 797
 Thucydides, 97, 262
 Tibetan language, 130, 137, 145–8, 203, 206,
 641, 733
 al-Tibrizī, 167
 Tietze, Andreas, 375, 376, 797
 Tiktin, Heimann Hariton, 548, 549, 797
 Tindale, Norman B., 685, 797
 Tittmann, Johann August Heinrich, 254
Tivākaram, 149–51, 152, 156, 644
 Tiwa language, 404, 733
 Tlingit language, 604, 733
 Tobler, Adolf, 536, 798
 Todd, Henry John, 406
Tolkāppiyam, 80–3, 148–9, 152
Tolkāppiyaṇ, 80, 798
 Tommaseo, Niccolo, 538, 539–40, 798
 tones, 113–15, 116, 117, 119, 205, 211, 219–21, 361,
 456, 459, 587, 611, 664, 677
Tóngwén bèikǎo, 117
 Tonocoté language, 567, 733
 topical ordering of entries. *See* macrostruc-
 ture, thematic
 topolects. *See* dialects and regionalisms,
 Chinese
 toponyms. *See* placenames
 Torero, Alfredo, 577
 Torres Rubio, Diego de, 555–6, 560, 561,
 564, 798
 Tours, François-Marie de, 639, 651, 798

- Trabot, Perez, 193, 284, 798
 Trần Trọng Dường, 360
 translation and interpretation. *See also* Bible
 in ancient Egypt, 41, 44, 45–7,
 85, 96
 in Anglo-Saxon England, 6
 in Arabic-speaking world, 231,
 415, 421
 in Australia, 693, 694–5, 696, 700
 in Chinese periphery, 203, 205, 214, 217,
 222, 357
 in colonial Africa, 481, 660, 670
 in early modern Europe, 243, 245, 307
 in Iran, 386
 in Latin Christendom, 276
 in medieval Egypt, 48–9
 in modern China, 323, 331
 in modern Europe, 419, 458, 539
 in pre-colonial Africa, 662
 in pre-modern China, 65, 127
 in South America, 563, 565
 in South Asia, 380, 402, 405, 410, 634,
 649, 656
 in Tibet, 130, 146
 in Turkey, 245, 372
 in Turkic-speaking world, 238
 in Vietnam, 363
 Trappists as lexicographers, 694–5
 Trávníček, František, 452, 798
 Trench, Richard Chenevix, 496
Trésor de la langue française, 535–6
 Trévoux, dictionary of, 534, 541
 Trigault, Nicolas, 326, 626, 798
 Trịnh Thị Ngọc Trúc, 357, 798
 Trumbull, James Hammond, 602, 603, 798
 Tshabe, S. L., 681
 Tsonga language, 481
 Tụ Đức, 359, 798
Tuhfat al-Hind, 381, 398
 Tupi language, 557–8, 567, 570, 733
 Tupi-Guarani languages, 733
 Turkic languages, 223, 226–30, 233, 238–40, 242,
 244, 366, 372, 392, 733
 Turkish language, 180, 223, 232, 234, 240–6,
 287, 366–77, 380, 381, 383–4, 392, 410–11,
 477, 547, 549, 733
 Turkmen language, 223, 734
 Tur-Sinai, Naftali, 433
 Tuscan language variety, 304, 538–9, 734
 TuTaTi (Mesopotamian sign list), 23, 24
 Tzeltal language, 582, 586, 734
 Tzotzil language, 6, 582, 734
 Tz'utujil language, 581, 734
 Ueda Kazutoshi, 344, 347, 798
 Ugaritic language, 15, 28, 734
Ugumu (Mesopotamian wordlist), 25, 26–7
 Ukrainian language, 453–4, 734
 Ullrich, Jan, 601
 Ulmann, Karl Christian, 458, 799
 Umbundu language, 660, 664–5, 680, 734
 Umpithamu language, 687, 734
 Underwood, Horace Grant, 351–2, 799
 unfinished dictionaries, 219, 263, 311, 387, 419,
 424, 516, 545, 546, 548, 551, 651
 Ungarinyin language, 697, 702–3, 734
 unpublished dictionaries, 194, 263–4, 291, 347,
 350, 386, 401, 444, 450, 455, 548, 563, 597,
 611, 634, 638, 684–7, 688, 690, 692–3,
 696–7, 701–2, 704
 Uppegger, Francis, 609, 611
Ura (Mesopotamian wordlist), 24–5, 31–3, 34
 Urbano, Alonso, 581, 595, 623, 799
 Urdu language, 6, 381, 382, 385, 396, 399, 400,
 407–13, 716
 Hindi-Urdu, 384
 Uri, Johannes, 383, 799
Uric col nikañtu, 149, 152
 usage notes, 88–9, 97, 101, 161, 421, 462, 468,
 487, 501, 503, 538; *see also* solecisms
 Ušakov, Dmitrij Nicolaevič, 453, 799
 Uto-Aztecan languages, 734
 Uyghur language, 203, 206, 734
 Vācaspati, 71–2, 799
 Vaiyapuri Piḷḷai, S., 157
 Valā, Aḥmad al-'Aziz al-Nā'itī, 387, 799
 Valdemārs, Krišjānis, 458, 799
 Valdivia, Luis de, 556, 561, 562–3, 567, 572, 799
 Valente, José Francisco, 665, 799
 Valentiano, Pietro Lupis, 241, 447, 799
 Valentijn, François, 638
 Vallot, P. G., 361
 Vāmanabhaṭṭa Bāṇa, 134, 138, 799
 Van Mol, Mark Leopold Cecile Maria, 417,
 423, 799
 Vararuci, 71
 Vārastah, Siyāl Kotī Mal, 379, 800
 variant forms
 Chinese, 52, 118, 120, 128, 320–1, 362
 English, 485, 491
 Greek, 88, 97
 Korean, 352
 Sanskrit, 136
 Scots, 514
 Sumerian, 33
 Turkic languages, 226

- Turkish, 246
 Vietnamese, 363
 Varo, Francisco, 616–17, 622, 623–4, 626, 627, 630, 800
 Varro Reatinus, 89
 Váša, Pavel, 452, 800
 Vefik Paşa, Ahmet, 370, 800
 Velasco, Juan Bautista de, 600, 800
 Velasco, Juan de, 556, 800
 Venda language, 481
 Ventallol, Magino, 626
 Verdugo, Andrés, 625
 Vergil, 96, 268–70, 510
 Verrius Flaccus, Marcus, 89–92, 95, 800
 versified dictionaries and wordlists, 3, 52, 80, 131, 147, 148, 150, 151–2, 197, 241, 278, 283, 288, 357, 359–60, 380, 390–1, 394–5, 398, 399, 400, 446, 643, 644
 Verstraete, Jean-Christophe, 687
Vessels and Garments (Mesopotamian wordlist), 17
 Veyssière de la Croze, Mathurin, 653
 Vico, Domingo de, 581, 800
 Vidal, Owen, 667
 Vietnamese language, 328, 356–65, 620, 622, 626, 627–8, 629, 632, 735
 Viguier, Pierre-François, 368–9, 800
 Vikramāditya, 71–2, 800
 Vilakazi, Benedict Wallet, 677, 680, 800
 Visayan languages, 618, 622, 623, 628–9, 735
 Viśvanātha, 138, 800
 Vivar, Pedro de, 618, 622, 625, 628–9, 800
 Vives, Juan Luis, 302
Vocabolario degli Accademici della Crusca, 304–5, 537–9, 540, 541
Vocabulário da lingua de lapam com a declaração em Português (1603), 214, 620–1, 624, 627, 631, 632
Vocabulario en lengua 4iche otlatecas, 583, 586, 593
Vocabulário na lingua brasílica (1621 or 1622), 557, 558, 568
Vocabularius breuiloquus, 281, 447
Vocabularius ex quo, 268, 282
Vocabularius incipiens Teutonicum, 282, 291
 Vodňanský, Jan, 446, 801
 Vodnik, Valentin, 450, 801
 Vogel, Claus, 72, 646
 von Schneidemesser, Luanne, 522
 von Soden, Wolfram, 11, 18
 Vopāla, 71
 Vossius, Gerardus Joannes, 295, 801
 Vrančić, Faust, 447–8
 Vullers, Johann August, 234, 384, 385, 801
 Vyāḍi, 71–2, 801
 Wadjuk language, 694, 735
 Wagner, M. L., 730
Wagyokuhen, 213, 342, 347
 Wahrig dictionary series, 471–2
 Wahrig, Gerhard, 471, 801
 Wahrig-Burfeind, Renate, 471
 Walker, John, 490, 801
 Walmajarri language, 703, 735
 Walther, Christoph Theodosius, 651
Wamyō ruiju shō, 209
 Wáng Ānshí, 122–4, 801
 Wáng Lì, 334, 801
 Wāng Yì, 325
 Wáng Yingdiàn, 117, 801
 Wang Yinglin, 359
 Wangkangurru language, 692, 735
 Warao language, 556, 735
 Warlpiri language, 703, 735
 Wartburg, Walther von, 535, 540, 801
 Washington, George, 491, 517
 Watson, George, 519
 Watson, Harry, 514
 Watson, William, 690
 Webster, Donald, 604
 Webster, Noah, 343, 399, 403, 490–1, 498, 517–18, 802; see also Merriam-Webster dictionaries
 Wehr, Hans, 417, 420–4, 425, 429, 802
 Wèi Jiàngōng, 334, 802
 Wèi Jiào, 123, 126, 802
 Wéi Yǐxīn, 335
 Weinbach, Liora, 442
 Wellig, Arnold Gottlieb, 458, 459, 802
 Welsh language, 278–9, 307, 735
 Wemmers, Jacobus, 299, 802
 Werner, Reinhold Otto, 544, 802
 Western Apache language, 606–13, 735
 Western Desert Language, 687, 697, 701–2, 735
 Whipple, Amiel Weeks, 608–10, 802
 White, J. B., 608–10
 Whitney, William Dwight, 497
 Whitworth, George Clifford, 527, 802
 Whomersley, Derek, 515
 Widdowson, John David Allison, 524, 802
 Wiegand, Hans Ernst, 1
 Wik Munkan language, 703, 735
 Wilkins, Charles, 385, 802
 Wilkins, John, 309
 Wilkinson, Endymion, 320
 Willems, Adriaen, 660, 802

- Williams, Roger, 602, 802
 Williams, Samuel Wells, 330, 802
 Williamson, John Poage, 603, 803
 Wilson, Horace Hayman, 403
 Wilson, Nigel, 260
 Wiltens, Caspar, 642, 646, 803
 Winer, Lise, 525–6
 Winter, Marc, 320
 Wiradjuri language, 685, 687, 690, 735
 Witsen, Nicolaas, 652, 736, 803
 Wittenberg, Martin, 440
 Woenne, Susan Tod, 687
 Wolmby, Topsy, 704
 Wolof language, 665–6, 736
 women
 as readers of dictionaries, 209, 215
 language of, 73, 396, 407, 627, 677, 679
 Wood (Mesopotamian wordlist), 17
Woordenboek der Nederlandsche Taal, 475, 478–9, 493
Word List C (Mesopotamian wordlist), 17, 19
 Wordsworth, Christopher, 512, 803
 Worms, Ernest Ailred, 698, 700, 704, 803
 Worrorra language, 696–7, 736
 Worrorran languages, 699, 702, 736
 Wright, Joseph, 510, 512, 513, 520, 803
 Wright, Thomas, 512, 803
 writing media
 bamboo slips, 53
 clay tablets, 6, 12, 13–14, 16, 19, 29, 30
 ostraca, 36
 palm-leaf, 6, 149, 156, 639, 650
 paper, 6, 247, 267, 390, 394, 395–6, 650, 652, 653
 papyrus, 36–7, 38, 42–5, 84–7, 92–6
 parchment, 247, 267, 271
 wax tablets, 271
 writing systems
 Amharic, 671
 Arabic, 200, 231, 240–1, 242, 244, 245–6, 367, 373, 374, 393–4, 407, 409, 416, 419, 422, 423, 426, 648, 669, 671
 Aramaic, 232
 Chinese, 51–66, 109–29, 202–22, 340–65, 616, 619–20, 637
 Chữ Nôm, 356–65
 Coptic, 44, 46
 cuneiform, 11–35, 40
 Cyrillic, 350, 387, 446, 450
 demotic, 40, 42–3, 44, 46
 Deseret, 604
 Glagolitic, 445
 Grantha, 646
 Greek, 39, 42, 44, 445
 Hebrew, 192, 200, 235, 284, 286, 288, 440, 483
 hieratic, 39, 40, 42–3, 44, 46
 hieroglyphs, 38, 39, 40, 41, 42, 43, 46, 50, 680
 Indic, 138, 142, 212–13, 393–5, 406–7, 409–10, 648
 International Phonetic Alphabet, 685
 Japanese syllabaries, 207, 209–10, 341, 619–20
 Jurchen, 205
 Korean, 202, 219–20, 349–51, 352, 354
 Mandarin Phonetic Symbols, 321, 325
 Old Coptic, 44
 Old Turkic, 228
 ‘Phags-pa, 205, 219
 phonetic (for Australian languages), 696
 Roman, 237, 242–4, 246, 287, 317, 321, 325–7, 331–4, 338, 351, 352, 358, 361–4, 367, 369, 373, 374, 387, 406, 422, 446, 456, 457–8, 483, 584, 587, 615–16, 620, 621, 626, 648–50, 667, 669, 671, 679–80, 683, 698
 South Asian, 639, 648
 Standard Alphabet of Lepsius, 671, 672–3, 678, 680
 Tamil, 393, 620, 648, 650
 Tangut, 202
 Telugu, 396
 Tibetan, 204
 Uyghur-Mongol, 203, 205
 Wú Jǐngróng, 336, 803
 Wú Yù, 116, 803
 Wú language variety, 319, 330, 736
Wújīng wénzì, 111, 118
 Wunambal language, 697

 Xavier, Angela Barreto, 654
 Xeberó language, 556, 736
 Xenophon, 97
 Xhosa language, 676, 678–80, 681, 736
Xiàndài Hànyǔ cídiǎn, 333–4, 336
Xīnhuà zìdiǎn, 333–4
 Xú Kǎi, 119, 803
 Xǔ Shèn, 53, 59–63, 66, 803
 Xú Xuàn, 119, 803

 Yādavaprakāśa, 134, 137, 138, 140, 141, 803
 Yakut language, 230, 736
 Yamada Tadao, 345, 803
 Yandruwandha language, 692, 736
 Yang, Huiling, 327
 Yáng Xióng, 57–8, 126, 329, 803
 Yao language, 670, 736
 Yapugway, Nicolás, 563

- Yaqui language, 600, 736
 Yäska, 69–71, 804
 Yassic language, 287, 736
 Yawarrawarrka language, 692, 736
 Yawuru language, 699, 736
 al-Yazīdī, 164, 804
 Yedidia of Rimini, 185, 804
 Yehuda Hasha^cari da Porta, 197
 Yellin, David, 434, 438
 Yi Heeseung, 354
 Yiddish language, 439, 477, 482–3, 736
 Yīqīējīng yīnyì, 128, 218
 Yorta Yorta language, 687, 736
 Yoruba language, 666–8, 680, 736
 Young, Robert, 602, 609, 611
 Yuè language variety, 127, 319, 330, 737
 Yule, Henry, 527, 804
 Yulparija language, 704, 737
 Yūpiān, 59, 62–5, 208, 213
 Yup'ik language, 604, 737
 Yuwaalaraay language, 687, 737
 al-Zabīdī, 178, 181, 419, 804
 al-Zagğāğī, 168, 804
 al-Zamahšarī, 159, 177, 180, 239, 804

 Zamboni, Alberto, 714
 Zapotec languages, 561, 580, 581, 584, 585,
 591–2, 595, 737

 Zaubnmüller, Wolfram, 1
 Zawzanī, Qāzī Ḥusayn, 232
 Zdanowicz, Aleksander Maria, 451, 804
 Zeisberger, David, 603, 804
 Zemánek, Petr, 417
 Zenker, Julius Theodor, 372, 651, 804
 Zezuru language, 675, 737
 Zgusta, Ladislav, 1
 Zhang Lo tsā ba Dge ba, 146
 Zhāng Xuān, 125, 804
 Zhào Yíguāng, 120, 804
 Zhèngzī tōng, 121
 Zhōngyuán yīnyùn, 115, 117, 220
 Zhōu Déqīng, 115, 805
 Zhū Móuwěi, 125, 805
 Zhū Xī, 116, 125
 Zibell, Wilfried, 604
 Ziegenbalg, Bartholomäus, 640, 644, 650, 652,
 653, 805
 Zihui, 65, 120–2, 207, 214, 217
 Zingarelli, Nicola, 539, 805
 Ẓiyā' al-Dīn Aḥmad Kḥān, Munshī, 399
 Zonaras, lexicon attributed to, 254–5,
 263, 265
 Zulu language, 675, 676–8, 680–1, 737
 Zúñiga Marroquín, Dionisio de, 582
 Županov, Ines G., 654
 Zyzanij, Lavrentij Ivanovyč, 448, 805

