

THE CAMBRIDGE HISTORY OF
THE ROMANCE LANGUAGES

*

VOLUME II

Contexts

What is the origin of the Romance languages and how did they evolve? When and how did they become different from Latin, and from each other? Volume II of *The Cambridge History of the Romance Languages* offers fresh and original reflections on the principal questions and issues in the comparative external histories of the Romance languages. It is organized around the two key themes of influences and institutions, exploring the fundamental influence of contact with, and borrowing from, other languages (including Latin), and the cultural and institutional forces at work in the establishment of standard languages and norms of correctness. A perfect complement to the first volume, this final volume offers an external history of the Romance languages combining data and theory to produce new and revealing perspectives on the shaping of the Romance languages.

MARTIN MAIDEN is Professor of the Romance Languages and Director of the Research Centre for Romance Linguistics at the University of Oxford, a Fellow of Trinity College, Oxford, and a Fellow of the British Academy.

JOHN CHARLES SMITH is Faculty Lecturer in French Linguistics and Deputy Director of the Research Centre for Romance Linguistics at the University of Oxford, and a Fellow of St Catherine's College, Oxford.

ADAM LEDGEWAY is Professor of Italian and Romance Linguistics at the University of Cambridge, and a Fellow of Downing College, Cambridge.

THE CAMBRIDGE
HISTORY OF
THE ROMANCE LANGUAGES

★

VOLUME II
Contexts

★

Edited by
MARTIN MAIDEN,
JOHN CHARLES SMITH
and
ADAM LEDGEWAY



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Published in the United States of America by Cambridge University Press, New York

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/9780521800730

© Cambridge University Press 2013

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 2013

Printed and bound in the United Kingdom by CPI Group Ltd, CRO 4YY

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

Library of Congress Cataloguing in Publication data

The Cambridge history of the Romance languages. Volume II, Contexts / edited by Martin Maiden, John Charles Smith and Adam Ledgeway.

pages cm

Includes bibliographical references and index.

ISBN 978-0-521-80073-0

- I. Romance languages – History. I. Maiden, Martin, 1957– editor of compilation. II. Smith, John Charles, 1950– editor of compilation III. Ledgeway, Adam, editor of compilation.

PC45.C246 2013

440.09–dc23

2012047375

ISBN 978-0-521-80073-0 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 HISTORY OF
THE ROMANCE LANGUAGES

*

VOLUME II

Contexts

What is the origin of the Romance languages and how did they evolve? When and how did they become different from Latin, and from each other? Volume II of *The Cambridge History of the Romance Languages* offers fresh and original reflections on the principal questions and issues in the comparative external histories of the Romance languages. It is organized around the two key themes of influences and institutions, exploring the fundamental influence of contact with, and borrowing from, other languages (including Latin), and the cultural and institutional forces at work in the establishment of standard languages and norms of correctness. A perfect complement to the first volume, this final volume offers an external history of the Romance languages combining data and theory to produce new and revealing perspectives on the shaping of the Romance languages.

MARTIN MAIDEN is Professor of the Romance Languages and Director of the Research Centre for Romance Linguistics at the University of Oxford, a Fellow of Trinity College, Oxford, and a Fellow of the British Academy.

JOHN CHARLES SMITH is Faculty Lecturer in French Linguistics and Deputy Director of the Research Centre for Romance Linguistics at the University of Oxford, and a Fellow of St Catherine's College, Oxford.

ADAM LEDGEWAY is Professor of Italian and Romance Linguistics at the University of Cambridge, and a Fellow of Downing College, Cambridge.

THE CAMBRIDGE
HISTORY OF
THE ROMANCE LANGUAGES

★

VOLUME II
Contexts

★

Edited by
MARTIN MAIDEN,
JOHN CHARLES SMITH
and
ADAM LEDGEWAY



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Published in the United States of America by Cambridge University Press, New York

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/9780521800730

© Cambridge University Press 2013

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 2013

Printed and bound in the United Kingdom by CPI Group Ltd, CRO 4YY

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

Library of Congress Cataloguing in Publication data

The Cambridge history of the Romance languages. Volume II, Contexts / edited by Martin Maiden, John Charles Smith and Adam Ledgeway.

pages cm

Includes bibliographical references and index.

ISBN 978-0-521-80073-0

- I. Romance languages – History. I. Maiden, Martin, 1957– editor of compilation. II. Smith, John Charles, 1950– editor of compilation III. Ledgeway, Adam, editor of compilation.

PC45.C246 2013

440.09–dc23

2012047375

ISBN 978-0-521-80073-0 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 HISTORY OF
THE ROMANCE LANGUAGES

*

VOLUME II

Contexts

What is the origin of the Romance languages and how did they evolve? When and how did they become different from Latin, and from each other? Volume II of *The Cambridge History of the Romance Languages* offers fresh and original reflections on the principal questions and issues in the comparative external histories of the Romance languages. It is organized around the two key themes of influences and institutions, exploring the fundamental influence of contact with, and borrowing from, other languages (including Latin), and the cultural and institutional forces at work in the establishment of standard languages and norms of correctness. A perfect complement to the first volume, this final volume offers an external history of the Romance languages combining data and theory to produce new and revealing perspectives on the shaping of the Romance languages.

MARTIN MAIDEN is Professor of the Romance Languages and Director of the Research Centre for Romance Linguistics at the University of Oxford, a Fellow of Trinity College, Oxford, and a Fellow of the British Academy.

JOHN CHARLES SMITH is Faculty Lecturer in French Linguistics and Deputy Director of the Research Centre for Romance Linguistics at the University of Oxford, and a Fellow of St Catherine's College, Oxford.

ADAM LEDGEWAY is Professor of Italian and Romance Linguistics at the University of Cambridge, and a Fellow of Downing College, Cambridge.

THE CAMBRIDGE
HISTORY OF
THE ROMANCE LANGUAGES

★

VOLUME II
Contexts

★

Edited by
MARTIN MAIDEN,
JOHN CHARLES SMITH
and
ADAM LEDGEWAY



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Published in the United States of America by Cambridge University Press, New York

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/9780521800730

© Cambridge University Press 2013

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 2013

Printed and bound in the United Kingdom by CPI Group Ltd, CRO 4YY

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

Library of Congress Cataloguing in Publication data

The Cambridge history of the Romance languages. Volume II, Contexts / edited by Martin Maiden, John Charles Smith and Adam Ledgeway.

pages cm

Includes bibliographical references and index.

ISBN 978-0-521-80073-0

- I. Romance languages – History. I. Maiden, Martin, 1957– editor of compilation. II. Smith, John Charles, 1950– editor of compilation III. Ledgeway, Adam, editor of compilation.

PC45.C246 2013

440.09–dc23

2012047375

ISBN 978-0-521-80073-0 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.

In memoriam

József Herman 1924–2005

Arnulf Stefenelli 1938–2002

Suzanne Fleischman 1948–2000

Contents

List of contributors page ix
List of abbreviations x

Introduction 1

1 · Latin and the making of the Romance languages 6

ALBERTO VARVARO

2 · The transition from Latin to the Romance languages 57

MICHEL BANNIARD

3 · Periodization 107

ROGER WRIGHT

4 · Evidence and sources 125

ROGER WRIGHT

5 · *Koinés* and *scriptae* 143

JOHANNES KABATEK

6 · Contact and borrowing 187

MARIUS SALA

7 · The Romance languages in the Renaissance and after 237

HELENA L. SANSON

8 · Geography and distribution of the Romance languages in Europe 283

ALVISE ANDREOSE AND LORENZO RENZI

Contents

9 · The sociology of the Romance languages 335

ALBERTO VARVARO

10 · Romance outside the Romània 361

MARI C. JONES AND CHRISTOPHER J. POUNTAIN

11 · Creoles 400

IRIS BACHMANN

References and bibliographical abbreviations 445

Index 508

Contributors

ALVISE ANDREOSE Ricercatore in Linguistica italiana, Facoltà di Lettere, Università Telematica e-Campus

IRIS BACHMANN Lecturer in Hispanic Linguistics, the University of Manchester

MICHEL BANNIARD Directeur d'Etudes, Chaire de Sociolinguistique diachronique romane, EPHE, Paris-Sorbonne

MARI C. JONES Reader in French Linguistics and Language Change, Department of French, University of Cambridge, and Fellow of Peterhouse

JOHANNES KABATEK Professor of Romance Linguistics, Romanisches Seminar, Universität Zürich

CHRISTOPHER J. POUNTAIN Professor of Spanish Linguistics, Queen Mary, University of London

LORENZO RENZI Professore Ordinario di Filologia Romanza, Università di Padova

MARIUS SALA Directorul Institutului de Lingvistică "Iorgu Iordan – Al. Rosetti" al Academiei Române, Bucharest

HELENA L. SANSON Senior Lecturer, Department of Italian, University of Cambridge and Fellow of Clare College

ALBERTO VARVARO Accademia Nazionale dei Lincei, Rome

ROGER WRIGHT Professor of Spanish, University of Liverpool

Abbreviations

Bibliographical abbreviations will be found under *References*, at the end of this volume.

*	unattested form or usage
**	ungrammatical or non-existent form or usage
=	cliticized to
1	1st person
2	2nd person
3	3rd person
ABL	ablative
ACC	accusative
AD	anno Domini
ANT	anterior
Arb.	Arabic
ArchCst.	archaic Castilian
ArchIt.	archaic Italian
Arg.	Aragonese
Art.	article
Ast.	Asturian
ATT	attribute
Bal. Cat.	Balearic Catalan
BC	Before Christ
BEN	benefactive
C	central
c.	circa
Cat.	Catalan
CC	circumstantial complement
cf.	compare
ch.	chapter
CIL	<i>Corpus Inscriptionum Latinarum</i>
COMP	complementizer
CONJ	conjunction
COP	copula
Crt.	Croatian
CSLat.	Classical spoken Latin

List of abbreviations

Cst.	Castilian
Ctb.	Cantabrian
DAT	dative
DET	determiner
DO	direct object
DOC	direct object case
E	east(ern)
e.g.	for example
Egd.	Engadinish
etc.	et cetera
EurPt.	European Portuguese
F	feminine
f.	and the following page
FPr.	Franco-Provençal
Fr.	French
Frl.	Friulian
Glc.	Galician
Grk.	Greek
Gsc.	Gascon
IMP	imperative
INF	infinitive
IO	indirect object
IOC	indirect object case
IPA	International Phonetic Alphabet
IPF	imperfect
IRo.	Istro-Romanian
It.	Italian
L ₁	first-language
L ₂	second-language
Lat.	Latin
lit.	literally
LOC	locative
LSLat.	late spoken Latin
LSLat ₁	late spoken Latin Phase 1: third to fifth centuries
LSLat ₂	late spoken Latin Phase 2: sixth and seventh centuries
M	masculine
ModRo.	modern Romanian
N	north(ern)
NEG	negator
NEUT	neuter
NOM	nominative
NP	noun phrase
O	old
Occ.	Occitan
OFr.	old French
OIt	old Italian

List of abbreviations

OSp	old Spanish
p.	page
PASS	passive
PEI	Prince Edward Island
PFV	perfective
Pic.	Picard
PL	plural
POSS	possessive
PP	prepositional phrase
pp.	pages
Prep	preposition
PRG	progressive
PRON	pronoun
ProtoCst.	proto-Castilian
ProtoFr.	proto-French
ProtoIt.	proto-Italian
ProtoRom.	proto-Romance
PRS	present
PRT	preterite
PRV	perfective
Prv.	Provençal
PST	past
Pt.	Portuguese
PTCP	participle
Ro.	Romanian
S	south(ern)
sg	singular
Sic.	Sicilian
Slv.	Slavonic
SOV	subject object verb
Sp.	Spanish
Srd.	Sardinian
StFr.	standard French
SVO	subject verb object
THEM	theme
TMA	tense, mood, aspect
Vto.	Venetan
W	west(ern)

Introduction

This *Cambridge History of the Romance Languages* stands on the shoulders of giants. A glance at the list of bibliographical references in these volumes should suffice to give some idea of the enormous body of descriptive and interpretative literature on the history of the Romance languages, both from the point of view of their structural evolution (the focus of the first volume, published in 2011) and with regard to the contexts in which they have emerged as distinct 'languages', and gained or lost speakers and territory, and come into contact with other languages (the main focus of this volume). This profusion of scholarship, adopting a multiplicity of approaches (synchronic, diachronic, microscopic, macroscopic) has more than once provided material for major, indeed monumental, comparative-historical synopses (e.g., Meyer-Lübke (1890–1902), Lausberg (1956–62), or the massively detailed and indispensable encyclopaedic works such as Holtus, Metzeltin and Schmitt (1988–2001) and Ernst, Gleßgen, Schmitt and Schweickard (2003–08)).

Much of the finest scholarship in Romance linguistics has, naturally enough, been conducted in Romance languages, or in German (the native language of some of the major founding figures of the discipline). One of our aims is to reach out to linguists who are not Romance specialists, and who may not know these languages. While the histories of some of the better-known major Romance languages (Italian, French, Spanish, Portuguese) have been treated in English, this work is certainly the first detailed comparative history of the Romance languages to appear in English.¹

The aim of *The Cambridge History of the Romance Languages* is not to compete with or supersede the works mentioned above, but to complement them, by

¹ There are, of course, some very useful smaller-scale works, such as Hall (1974), Elcock (1975), Harris (1978), Harris and Vincent (1988); also of interest is Posner and Green (1980–93).

presenting both to Romanists and to historical linguists at large the major and most exciting insights to emerge from the comparative-historical study of Romance. With this in mind, we have deliberately attempted in the presentation and discussion of the material of the two volumes to adopt a more inclusive approach which, while not alienating the traditional Romanist, bears in mind the practical limitations and needs of an interested non-specialist Romance readership (witness, for instance, the extensive translation of Romance and Latin examples), though in no case is this done at the expense of empirical and analytic detail.

It is our firm belief that the richly documented diachronic, diatopic, diastratic, diamesic and diaphasic variation exhibited by the Romance family offers an unparalleled wealth of linguistic data of interest not just to Romanists, but also to non-Romance-specialists. This perennially fertile and still under-utilized testing ground, we believe, has a central role to play in challenging linguistic orthodoxies and shaping and informing new ideas and perspectives about language change, structure and variation, and should therefore be at the forefront of linguistic research and accessible to the wider linguistic community.

The present work is not a 'history' of Romance languages in the traditional sense of a 'standard' reference manual ('vademecum') providing a comprehensive structural overview of individual 'languages' and/or traditional themes (e.g., 'Lexis', 'Vowels', 'Nominal Group', 'Tense, Aspect and Mood', 'Subordination', 'Substrate', 'Prehistory', etc.) on a chapter by chapter basis (cf., among others, Tagliavini (1972), Harris and Vincent (1988), Holtus, Metzeltin and Schmitt (1988–2001)), but, rather, is a collection of fresh and original reflections on what we deem to be the principal questions and issues in the comparative internal (Volume 1: Structures) and external (Volume 2: Contexts) histories of the Romance languages, informed by contemporary thinking in both Romance linguistics and general linguistic theory and organized according to novel chapter divisions, which reflect broader, overriding comparative concerns and themes (generally neglected or left untackled in standard works), rather than those which are narrowly focused on individual languages or developments. Inevitably, this will mean that certain aspects of the history of the Romance languages or individual members thereof – though admittedly very few, as a thorough reading of the following pages reveals – may not be exhaustively covered. Nevertheless, we are convinced that the merits of the individual chapter divisions adopted here far outweigh any potential lacunae (for which, in any event, there exist in virtually all cases other reliable treatments).

This work is organized around four key recurrent themes: *persistence*, *innovation*, *influences* and *institutions*. Thus, much of the first volume, dedicated to

thelinguistic ‘Structures’ of Romance, juxtaposes chapters or chapter sections dealing with issues of persistence on the one hand and innovation on the other in relation to the macroareas of phonology, morphology, morphosyntax, lexis, semantics and discourse-pragmatics. It goes without saying that the Romance languages are the modern continuers of Latin and therefore many aspects of structure persist from that language into Romance. It is not usual, however, for works on the Romance languages to concentrate on these factors of inheritance and continuity, since they – understandably – prefer to comment on what is new and different in Romance by comparison with Latin. By contrast, we believe that it is an important and original aspect of the present work that it accords persistence in Romance (and hence inheritance from Latin) a focus in its own right rather than treating it simply as the background to the study of the changes. At the same time, we devote considerable space to the patterns of innovation (including loss) that have taken place in the evolution of Romance.

Structural persistence and innovation within Romance cannot of course be studied in isolation from the influences and institutions with which the Romance languages and their speakers have variously come into contact at different periods in their history. For this reason, the authors of individual chapters in Volume 1 were encouraged to consider, as far as possible, structural persistence and innovation in relation to these influences and institutions and the extent to which they may have helped in arresting or delaying them on the one hand and shaping or accelerating them on the other. It is, however, in this second volume, dedicated to the ‘Contexts’ in which the Romance languages have evolved, that the central role assumed by *influences* and *institutions* is investigated, as well as their bearing on questions of persistence and innovation (cf. Bachmann’s discussion of the Romance creoles). It is well known that the Romance languages have been subject in varying degrees to the effects of outside influences. In addition to contact and borrowing (e.g., from Germanic, Arabic, Slavonic) and substrate effects (e.g., from Celtic), there is also the all-important role of Latin as a learned language of culture and education existing side by side and interacting with the evolving languages, as well as the role of contact and borrowing between Romance languages. When speaking of institutions, we have in mind both the role of institutions in the sense of specific organizations (the Church, academies, governments, etc.) in the creation of ‘standard’ languages and the prescription of norms of correctness, and also the language as an institution in society involved, among other things, in education, government policy, and cultural and literary movements.

Consequently, the focus throughout the two volumes is on an integration of the internal and external perspectives on the history of the Romance

languages, in part achieved through a multiauthor format which brings together the best of recent scholarship in the two traditions, and in part through careful editorial intervention and cross-referencing across chapters and volumes. *In particular, all cross-references have been introduced by the editors, and are not to be attributed to the authors.* Where the editors have added notes to individual chapters, these are indicated as such and followed by the initials MM, JCS or AL. Furthermore, all citations in languages other than English have been translated by the authors or the editors. However, as editors we have been keen to impose as few constraints on our contributors as possible in order to create an opportunity for international scholars of stature and intellectual vision to reflect on the principles and areas that have been influential in a particular subarea, and to reassess the situation.

It is necessary here to mention, albeit briefly, the rationale behind a number of our decisions in representing, and referring to, Latin. It is customary (though in no way a universally accepted practice) in many works on Latin and Romance to cite Latin forms in small capitals. Although we recognize that there are, of course, no linguistic grounds for this choice of typographic representation, inasmuch as Latin forms could just as legitimately appear in lower-case italics on a par with any other language, we have chosen to follow here the (more or less) established convention of employing small capitals for cited examples. While it is true that the ancient Romans did not use small capitals to represent their language, it is equally true that they did not use lower-case italics. However, we believe that the conventional practice of placing Latin forms in small capitals has the typographical advantage, especially in a work like ours, where reference to Latin forms is legion, of allowing immediate and efficient recognition of the two diachronic poles of our investigation, Latin (small capitals) and Romance (lower-case italics). Where we do depart, however, from current conventional practice is in our representation of the classical Latin high back vowel/glide [w], which is today usually represented as ‘v’ in syllable onsets (e.g., **VIVO** ‘I live’) and u in all other positions (e.g., **HABUIT** ‘he had’) or, according to another school of thought, as ‘V’ when it appears in upper case and ‘u’ when in lower case (e.g., *Viuo* ‘I live’). By contrast, we have preferred to adopt u (lower case) / U (upper case) in all positions (hence, **UIVO** and **HABUIT**) which makes the value of the grapheme more transparent in the discussion of Latin (morpho)phonology.² One further departure from current typographical conventions concerns

2 For a detailed discussion of Latin orthographic practices, see Wallace (2011).

our decision to cite all non-attested forms, whether reconstructed for Latin or any other language (but in all cases preceded by a single asterisk) in phonetic transcription (e.g., *vo'lere 'to want' replacing classical UELLE), and not in small capitals (e.g., *UOLERE), as is frequently the case in other works.

Finally, although we do not wish to enter here into a discussion of the value or the appropriateness of such labels as 'vulgar', 'late', 'spoken', 'literary' and many others in relation to Latin (for which we refer the reader to the chapters by Varvaro, Banniard, and Wright (chapter 3)), we are keen to point out that we do not consider Latin a monolithic variety, uniquely to be identified with the prescriptive norm passed down to us in the high literary and rhetorical models of the classical era. Rather, we take Latin, like any other natural language that has existed, to be a rich and varied polymorphous linguistic system which was subject, on both the diachronic and synchronic axes, to the same kinds of diatopic, diastratic, diamesic and diaphasic variation as its modern Romance descendants. We therefore deliberately avoid capitalized epithets in such syntagms as 'Vulgar Latin' or 'Late Latin', which unreasonably suggest an ill-founded linguistic and psychological demarcation between one supposed language, Classical Latin on the one hand, and an autonomous derivative, 'Vulgar Latin' or 'Late Latin', on the other. Rather, in the same way that linguists regularly append descriptive labels like 'modern', 'spoken', 'popular', 'dialectal', 'journalistic', 'literary', 'Latin-American' and such like to the modern Romance languages to refer to a particular 'variety' of that language (e.g., '(spoken) Barcelona Catalan', 'popular French', 'journalistic Italian', 'literary Romanian', 'Latin-American Spanish'; see Wright chapter 3, this volume, for further discussion), we have left it to the discretion of individual authors to indicate and identify, where necessary, the particular register, style or variety of Latin intended by means of an appropriate non-capitalized epithet or periphrasis, be it 'vulgar Latin', 'spoken Latin' or 'the Latin of North-West Africa'.

To conclude, we should like to remember here Suzanne Fleischman (1948–2000), József Herman (1924–2005) and Arnulf Stefenelli (1938–2002). The first-named died on the very day on which we wrote to her proposing that she might contribute to this work; the second before being able to complete the chapter we had invited him to write for this volume; the third shortly after contributing the chapter on 'Lexical Stability' to Volume I. They are sorely missed, but their legacy to Romance Linguistics lives on. We dedicate to them the present volume of the *Cambridge History of the Romance Languages*.

Latin and the making of the Romance languages¹

ALBERTO VARVARO

1. Latin: origins, characteristics and areal diffusion

Today the Romance languages are spoken over much of Europe (Iberian Peninsula, France, southern Belgium, western Switzerland, Italy and Romania), central and southern Latin America and Quebec, as well as in the former French, Portuguese and Spanish colonies of many parts of Africa and, to a lesser extent, Asia, where they enjoy official language status and function as the recognized languages of culture (see Andreose and Renzi, and Jones and Pountain, this volume, chapters 8 and 10). Their origins lie in a variety of Indo-European that was spoken from about the eighth century BC in a small area of the lower Tiber valley around Rome and the Alban hills. Although flanked to the north of the Tiber by Etruscan, a non-Indo-European language stretching as far as the southern bank of the River Arno, Latin was not isolated: to the north-east, east and south of the Latin-speaking area the closely related Oscan dialects were spoken, stretching as far as Campania and Lucania to the south, the Adriatic to the east and the territories of the closely related Umbrian-speaking tribes to the north. Other languages spoken to the north included the Indo-European varieties Picenian, along the Adriatic coast, and Celtic over an area stretching from Senigallia to the Alps (apart from the Indo-European varieties Venetic and Raetic to the east and north respectively), and a non-Indo-European tongue, Ligurian, spoken along the upper Tyrrhenian coast. Linguistic variation was just as great in the southern part of the peninsula, where, in the modern-day region of Puglia, Messapic, an Indo-European language, was

1 The ideas contained in the following pages were first developed for a lecture course I gave a number of years ago at the University of California, Berkeley, at the invitation of Yakov Malkiel – to whose memory I dedicate this chapter – and were written up for the first time during the summer of 2009. I would like to thank my friend and colleague Adam Ledgeway who saved me from a number of errors; all remaining errors are of course my own responsibility.

spoken, as well as Greek along the coasts following the early Greek colonizations, most notably at Taranto, which reached as far as the southern point of Calabria.²

The distribution of languages sketched above must not, however, be understood as a series of compact linguistic areas in which one relatively standardized language acted as a roof language for a number of local variants. Rather, the peninsula was made up of a patchwork of small tribes without any form of political, cultural (except in the archaeological sense of the word) or linguistic unity. Diatopic linguistic diversity then must have been enormous, as is still often the case today in poorly developed and sparsely populated areas of the world.

The success of Latin is a direct, if not immediate, consequence of the gradual expansion of Roman political rule. Thanks to the power of its political institutions, the might of its armies and its resolute tenacity, Rome succeeded in imposing its *imperium* over the entire peninsula and beyond, coming in time to dominate the Mediterranean and almost all of north-eastern Europe. Linguistic Latinization was only impeded in the East, where Greek, thanks to its greater cultural prestige, remained dominant. However, linguistic Latinization in the West was not the result of any deliberate linguistic policy. Quite the contrary. Permission to use Latin was initially granted to non-citizens of Rome only as an exceptional, much sought-after privilege. Rather, it was the prestige of the city and the superiority of its culture that led non-Romans, beginning with the upper classes, to adopt Latin.

In the course of the first centuries of the Middle Ages there emerged in this vast area, by then almost completely and homogeneously Latinized, a number of Romance vernaculars that had evolved from Latin, which continued to be employed as the language of culture and writing. From about the beginning of the tenth century AD, these lower varieties began to be used also in writing, giving rise to their own literary traditions which most probably continue an earlier oral tradition. Gradually the written Romance varieties underwent various forms of standardization, eventually yielding the Romance languages that we today associate with important literary cultures: Portuguese, Spanish, Catalan (and also Occitan in the Middle Ages), French, Italian and Romanian. Their subsequent expansion following the geographical explorations of the late Middle Ages and, in particular, the discovery of America, represents a complex historical process outside the scope of the present chapter (for which, see Jones and Pountain, this volume, chapter 10), which will only be able to

2 For the linguistic situation in pre-Roman Italy, see Prosdocimi (1978).

consider the intricate problems involved in the making of the Romance languages.

How then should this extraordinary episode in the making of the Romance languages be best understood and explained? Although richly documented, unlike many other similar episodes that must have occurred, this does not make it any easier to understand. Indeed, on this point Malkiel (1978:28) wrote: 'the working hypotheses proposed, in the course of almost two centuries of speculation and detailed research, on the differentiation of the Romance languages are historical conjectures, sometimes extremely original and sophisticated [...] but still just isolated hypotheses and not theories in the strong sense of the term traditionally attributed to it by logicians and mathematicians, as well as, more recently, by linguists accustomed, as they are, to abstract inquiry [...] and not in the weak sense of the term as it is usually employed in everyday spoken language.'

Today, as in the past, it is still very common to divide the process into two successive stages: the completion of Latinization of the Western Empire followed by fragmentation into the many Romance vernaculars, alleged to have begun, to borrow Wartburg's terminology, in 'a restricted area' from within 'the compact area' (Wartburg 1950:rf.). In my opinion, it must have involved a unique and very complex dynamic process, which was not necessarily unidirectional.

2. The making of the Empire

As a direct consequence of the gradual establishment of Rome as an imperial power, Latin first began to spread across the Italian peninsula and then beyond to the western Mediterranean, and finally also to the eastern Mediterranean and the rest of the European continent. For our purposes, it is important to have a precise understanding of the chronology of the developments involved in this lengthy process of expansion. The final stage in the Roman conquest of the Italian peninsula came to an end some time before 264 BC with the end of the war against Tarentum (modern Taranto) and King Pyrrhus of Epirus (280–275 BC). This brought about a complex structural network of very diverse bilateral relations between Rome and the individual local communities, which remained partially, but often largely, autonomous. However, from 338 BC numerous colonies of Roman or Latin citizens began to spring up at strategic points across the peninsula. Following the first Punic War (264–241 BC), Sardinia and most of Sicily also fell under Roman rule and the first provinces

directly governed by Roman officials were established. This provincial system was subsequently extended to all new territorial gains.

The dates for the establishment of the different provinces, about which we often only have approximate information, are the following: Sicily in 242 BC, Sardinia and Corsica in 227 BC, Hither and Further Spain in 197 BC, Macedonia in 148 BC (in turn followed by other eastern provinces), Africa (roughly coinciding with present-day Tunisia) following the Third Punic War (149–146 BC), Narbonese Gaul in 121 BC, Cisalpine Gaul by 89 BC, Numidia in 46 BC, the Three Gauls (Aquitania, Gallia Lugdunensis and Gallia Belgica) between 16 and 13 BC, and Noricum and Raetia after 15 BC. Pannonia was elevated to the status of province in AD 9–10, Mauretania in AD 42, Britannia in AD 43, Upper and Lower Germania in AD 90, and finally Dacia in AD 107. It was not until Diocletian that the provincial system was overhauled (AD 297) to include the Italian peninsula, establishing 101 different administrative divisions smaller in size than the original provinces.

The Roman Empire was predominantly governed, even in the provinces, through a system of indirect rule similar to that used centuries later by the British in India. Because in general 'the Romans fought the battles of the settled and normally pacific populations of Italy against the more roving and predatory ones, or the alien Celtic nomads' (Cary and Scullard 1975:103), they systematically found support in the upper classes of the populations which they appeared to protect. Given the prestige enjoyed by the Romans, they had no need to impose their culture and language: instead it was the subjugated populations, beginning with the upper classes, who sought to conform to the cultural and linguistic habits and practices of their rulers in order to obtain (reduced) Latin or full Roman citizenship, adapting both lifestyle and language in the process.

This system of governance, which afforded the indigenous population a not inconsiderable number of powers, whilst leaving the Romans in charge of foreign policy (to use modern terminology), the army and tax collection (albeit through intermediaries), ensured that contact between the indigenous populations and Romans was not uniform across society. The legions, which were originally recruited exclusively from Roman citizens and allies and, later, in the imperial period, by Italic peoples more generally, weighed heavily upon the Empire, their soldiers entering into all sorts of relations with all cross-sections of the local population. However, the presence of the army was concentrated in the border areas of the Empire and in a few unruly pockets within the interior. It is quite wrong to imagine that the presence of soldiers was generalized across the Empire or even comparable to the deployment of armies in the modern period.

We must not underestimate, however, the early emergence in Rome of a very entrepreneurial and well-developed merchant class, which often turned up in future provinces well before they were absorbed into the Empire. In this way, they prepared the ground for a sort of peaceful infiltration which, at the same time as disseminating some of the most typical goods and wares of Roman life, provided individuals from outside the Empire with some knowledge of Latin.

Nonetheless, a large part of the population of the Empire only had limited and indirect contact with the citizens of Rome proper. The cities are themselves a characteristic phenomenon of the Roman world and therefore often of recent foundation, replicating a single common model with strictly regimented political structures. Outside the cities, in the first two centuries of the Empire, peasants generally enjoyed contacts with the local market, in part controlled by Romans, and with the gentry made up of Romanized members of the indigenous population, the senatorial class or sometimes officials of the imperial tax system. Opportunities and reasons to learn Latin were therefore rather limited. Without doubt, the spread of eastern religions during the imperial period, and especially of Christianity, had a huge impact, including at a linguistic level. These religions, including Christianity, penetrated Greek (or at any rate eastern) circles and the Greek language, whereas their dissemination in the West was progressively coupled with the use of Latin. The persistence of Christian missionaries, who fought to eliminate paganism from some of the most remote areas of the western countryside, must have contributed greatly to the ultimate loss of most of the pre-Roman languages (cf. §4) and the generalization of the use of Latin.

3. The Schuchardt–Gröber hypothesis

A hypothesis particularly popular among Romance linguists, and for that reason also discussed here, is that advanced in 1866 by Hugo Schuchardt, who claimed that the different dates of Romanization of the individual provinces corresponded to different types of Latin exported to these same areas, whose characteristics were subsequently to surface in the Romance languages. Schuchardt was well aware of the linguistic complexity of the ancient world and of the diachronic, diatopic and diastratic variation which must have been present in the Roman Empire, even if he did observe that ‘uneducated Latin [. . .] effectively always appears [. . .] on the monuments of all areas as one and the same language’ (p. 92), perceptively concluding that ‘in the later period, at least, uneducated spelling was quite conservative’ (p. 93). In

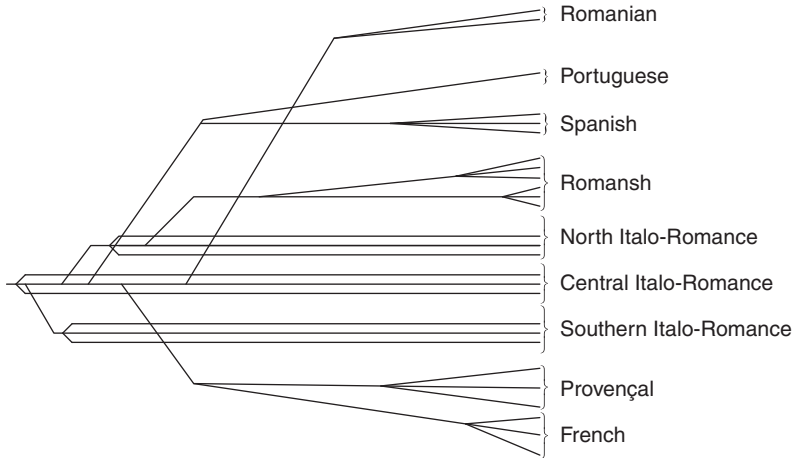


Figure 1.1 Schuchardt's (1866:82) Stammbaum Model of the Romance Languages

the wake of Darwin's then-recent theories, he believed it possible to draw a family tree of the Romance languages along the lines of A. Schleicher's model for Indo-European. Unlike his predecessor, however, Schuchardt traced a tree (see Figure 1.1) with a single evolutionary line from Latin to central Italian, from which the other Romance varieties branch off at different points in time according to the order in which the individual provinces were annexed to the Empire: southern Italo-Romance, northern Italo-Romance, Spanish, Portuguese, Provençal and French, Romansh, and Romanian. Clearly, then, for Schuchardt the Latin imported into these provinces represented subtly different evolutionary stages of the language (cf. Varvaro 1968:95f.).

This idea, still in its very vague conception, was taken up again about twenty years later by G. Gröber (1884), who wrote: 'the differentiation of the Romance languages therefore goes back to ancient times. It began at the time of the Romanization of the *first* province outside of Italy and repeated itself with the annexation of each of the new Latin-speaking areas, where on each occasion the language of the first Roman settlers formed the starting point for each of the Romance languages: the language had to withstand the arrival of new languages of later immigrants, be able to assimilate them and develop into a Romance language phonetically uninfluenced by them' (p. 213).

The hypothesis has to be praised for its attempt to historicize the relationships between languages, which the Darwinian-style trees represented in a highly simplified and abstract manner. However, the theory is largely flawed because it assumes that: (a) linguistic Romanization immediately followed

colonization; and (b) relations between individual provinces and the centre were so unstable, or even non-existent, as to prevent (or at least highly limit) linguistic innovations from moving beyond the area in which they had originally emerged.³

In short, it was an attempt, albeit a very crude one, at historicization. There can be no doubt that linguistic Romanization was an extremely long and complex process, hence ruling out (a) above. Nor is it realistic to consider that some areas of the Empire were so isolated that they would automatically have given rise to linguistic archaism. This presumed isolation cannot be shown to have existed either in Sardinia or in Lucania,⁴ to name just two oft-cited cases in the literature. While no one doubts that within the Empire there were both areas with strong and with more limited contacts with the outside world, we cannot legitimately speak of isolation.

In his analysis of presumed cases of regional archaism reported, for example, for the Iberian Peninsula, Adams (2007) observes how one must first demonstrate that such forms were already in use in the area in question and there alone, something which proves virtually impossible to demonstrate except in the case of a handful of lexical items. This observation gets straight to the heart of the problem, contrasting one method, based on existing documentary evidence for the period of the presumed first appearance of the archaism, with another quite different method, which naively relies on the present-day distribution of the suspected archaism, hence a reconstruction, without any consideration of how much of the early documentary evidence might have since been lost, nor how the survival of individual forms is often purely accidental. Thus, a reconstruction of the Latin linguistic situation based on the data of the Romance languages, which date from a period of at least five hundred years later, and in the case of modern dialect documentation from no less than 1500 years later, is a method unlikely to find favour with historians and should only be employed with utmost caution.

4. Latinization: chronology and methods

In effect, the Latinization of the western part of the Empire was an extraordinary, albeit slow, process which was repeated across all households within the same area, but which was never complete; witness the survival of Basque and

³ For a radically critical view of this hypothesis, see Adams (2007).

⁴ For Sardinia, see Blasco Ferrer (1984:16–52), and for Lucania, Varvaro (1983), as well as the arguments in Martino (1991), which ultimately have not caused me to change my mind on these matters since even isolation is a relative factor.

Albanian to the present day.⁵ However, the paucity and the often questionable nature of our sources is such that it is just as likely that there were, even within densely Latinized areas, several minority linguistic islands of non-Latin-speaking peoples, especially in those areas where Greek and Roman civilization had been less pervasive (Varvaro 2005:121).

Nonetheless, the quite extraordinary fact remains that a large part of Italy and the western provinces gradually gave up the various pre-Roman languages in favour of Latin. Yet, this does not mean that the process had been general since earliest times. Thanks to a few individual cases, we can get a glimpse of some of the dynamics involved to help us piece together the facts, although extreme caution must be exercised in not drawing too many wider generalizations from such cases. One method of observing the process involved in the movement from the indigenous to Roman culture, and hence presumably to Latin, is through the study of onomastics, where we can follow the situation within individual families, even if we must recognize that what we can actually observe is the adaptation, not of real linguistic usage, but of personal names which reflect the fashionable socio-cultural trends, not of those who bear the name, but of their parents (cf. Varvaro 2005).

The inscription which crowns the magnificent triumphal arch of Saintes (Charente-Maritime) informs us that the monument, which was dedicated to Tiberius, Drusus and Germanicus, was erected between AD 17 and 19 by the *praefectus fabrum* ('chief engineer') C. Iulius Rufus, who immediately afterwards became priest of Rome and of Augustus in Lyon, where he commissioned important works in the Amphitheatre of the Three Gauls, as attested in another inscription.⁶

Between them, the two inscriptions mention no less than six generations in Iulius Rufus' family, of which three refer to his recent ancestors, namely his great-grandfather, his grandfather and his father. His great-grandfather's name was EPOTSOROIDUS, clearly a Gaul with no signs whatsoever of Romanization. However, his grandfather bore the name C. IVLIVS GEDOMO, showing that he had indeed become a Roman citizen, most probably granted by Julius Caesar

5 I include here Albanian, although it is spoken today in an area which partially falls within the territory of the Eastern Empire. The Germanic varieties that survived within the Empire around the Rhine-Danube area were protected by their close proximity to the German interior, in the same way that the Celtic varieties of Roman Britain were bolstered by the Celtic varieties of Scotland and Ireland; a similar situation is observed for Berber, or its Libyan ancestor, in Africa.

6 The inscription from Saintes is to be found in *CIL* 13/1, n1036; the inscription from Lyon was discovered half a century ago and can be read in various sources, including Guey and Audin (1958) and Wuilleumier (1965:70, n217).

himself: only his *cognomen* ('family name') remained Celtic, although he is likely to have retained many original Celtic characteristics, beginning with his language. As for his father, we are told his name was C. IVLIVS OTVANEVNVS; he too, just like his own father before him, bore a Celtic *cognomen* indicating that he was born some time before his father had become a Roman citizen. Whatever the meaning of these three indigenous names, which I do not believe have ever been explained, we are presented with a family which, within the space of about sixty years, had become completely Roman, not just in name but also in behaviour and in terms of the public offices they held. Episodes like this, involving a change of name, happened time and time again across all provinces and in all periods.

In the mausoleum of El-Amrouni, to the north of Remada in proconsular Africa (about 200 km south of the island of Gerba in an area today part of Libya), a bilingual Latin-Punic inscription was discovered in 1894, dated to the first century AD. The Latin text, which is preceded by the Punic, reads as follows (Donner and Röllig 1966:122, nr 117; cf. Lidzbarski 1907:63f., nr 101):

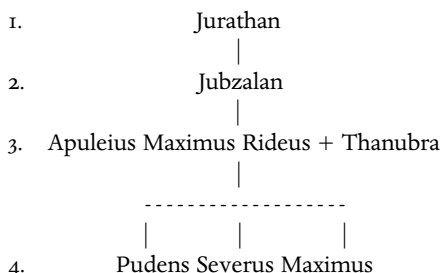
D(is) M(anibus) S(acrum)
 Q(uitus) APVLEVS MAXSSIMVS
 QVI ET RIDEVS VOCABA
 TVR IVZALE F(ilius) IVRATHE N(epos)
 VIX(it) AN(nis) LXXX THANVBRA
 CONIVNX ET PVDENS ET SE
 VERVS ET MAXSIMVS F(iliis)
 PIISSIMI P(atri) AMANTISSIMO S(ua) P(ecunia) F(ecerunt)⁷

I reproduce here in English translation Donner and Röllig's (1968:122) German translation of the Punic text: 'For the divine spirits of the dead of Apuleius Max[imus] Rideus, son of JWBZ' L' N', son of JWR'[T]' N, the Metabian, T'NBR, [his] wife, built (this mausoleum) for Pudens and Severus and Max[im]us [their] so[ns?].' The principal difference between the two texts is that in the Punic the mother is reported to build 'for, in the name of' her sons, whereas in the Latin she is said to do it 'with' them. Only the Latin text mentions the venerable age of the deceased, whereas the Punic alone refers to him by the nickname of HMTEBY, which apparently indicates his Metabian tribal origins, in short a family or clan name. The names in the two texts show some adaptation: for

7 Translation: To the shades of the dead. / Quintus Apuleius Maximus / who was also known as Rideus / son of Jubzalan, grandson of Jurathan / lived 90 years. Thanubra / his wife and Pudens and Se-/verus and Maximus his sons / most devoted built [this mausoleum] for their beloved father from their own money.

example, APULEIUS is rendered in Punic with the correct sounds, but appears in the vocative, the case form in which almost all Latin names appear in Neo-Punic inscriptions;⁸ the Latin permits the correct sounding of the indigenous names, but it too changes their endings.

At any rate, we can identify from the inscription four generations:



Now, generations (1) and (2) undoubtedly bear names of Berber origin.⁹ They were most probably Libyans who had been barely, if at all, Romanized, at least the first of them. In fact, the second of them, unlike his father, gave his son the name which is written in Latin as APVLEVS MAXSSIMVS, albeit with two irregularities: the second name bears a mere graphic error,¹⁰ whereas the first name shows the loss of the semivowel *i* (which in Latin was usually pronounced long), contrasting with the general tendency to reinforce the same sound by consonantalization, witness Italian *maggiore* and *peggio* from MAIORE(M) ‘better’ and PEIUS ‘worse’.

However, it is not these features, which in any case are ultimately to be ascribed to the stonemason or, at most, those who commissioned the inscription, that point to the deceased as having lived in a transitional period. Rather, it is the fact that he bears a strange nickname [RIDEUS?] and a distinctly non-Roman clan name, as confirmed by his family tree. Apuleius Maximus was living in the middle of a period of Romanization, and in fact had given his sons impeccably Latin names (even if they all continued to be *peregrini* ‘foreigners’ rather than *cives* ‘citizens’), but his wife still had a clearly Numidian name and chose (whether alone or in conjunction with her ‘Romanized’ sons) to place on her husband’s mausoleum a bilingual inscription with the Punic text above

8 For an explanation, see Donner and Röllig (1968:122) and Röllig (1980:292); cf., however, Adams (2007:570f.) for the use of the vocative.

9 See Lidzbarski (1907:63f.). A ‘Jurathan’ figures as the father in the inscription of the *iuvetus* of Mactar, from central Tunisia, from the year AD 88 (cf. Charles-Picard 1959:91).

10 -xss- for -x- is attested elsewhere, including, for example, at Salona (Croatia) with MAXSSIMUNA for MAXIMINA (CIL 3, 8971).

that of the Latin. This clearly exemplifies the chronological discrepancy (here by one generation) in the Romanization of the two sexes (to which I shall return below) and the complexity of the situation that existed in provincial areas. Here the indigenous population undergoing a gradual process of Romanization seem to be Libyan, but when they decide to commission an inscription in a pre-Roman language, they chose Punic. Moreover, by the fourth generation we would not be at all aware of the Libyan ethnicity of the family, even if we had an epigraph with the full associated family history, since by this point the family seemed to all intents and purposes entirely Romanized.

In nearby Mactar (in present-day Tunisia, between Kairouan and El-Kef), a long inscription from AD 88 allows us to study the lineage of 69 *iuvēnes* ('young men'), aged between seventeen and forty-six, during the same century as El-Amrouni. Charles-Picard (1959:90) noted that among the parents we find twenty-three Punic names, eleven Libyan names and twelve of uncertain, but in any case 'African', origin, whereas their respective offspring bear twenty Punic names, just two Libyan names and eleven of uncertain origin, highlighting a significant fall in the use of Libyan but not Punic names. He did not observe any cases of fathers with 'Libyan' or 'African' names who gave their children Punic names: *Balsillec* son of *Zruma*, *Baricbal* = *Fortunatus* son of *Iasuctan*.¹¹ By the same token, Latin names go up from twenty-five to thirty-five. Exemplary in this respect are the cases of names such as ROGATUS F[ilius] D. ADDUN (n° 1), FORTUNATUS F. D. ARSACESE (n° 2; the name of the father is typically eastern) and SILUANUS F.D. MUZTHUMBAL (Rogatus, Datus; n° 3).

It would seem then that the trend is the same in both inscriptions and that the dating of the changes also coincides. In reality, however, the process proves more complicated: here six fathers with 'Latin' names have given their sons 'Libyan' or 'African' names, as in the case of IAILUAI son of BASSUS, the opposite of cases like TITUS son of IURAUCAN (n° 30). What seems to have happened then was that in a context where the causes behind Romanization were numerous (beginning with the very establishment of a *iuventus* 'youth group' devoted to Mars), but where everyone was still considered a *peregrinus* 'foreigner', different lifestyles and ways of speaking coexisted alongside each other in accordance with a very precisely defined hierarchy (Latin > Punic > Libyan), but with the possibility of individual choices in all directions.

Of course, what is important here is not so much the individual case but the collective trend towards Romanization, which appears to be winning through

¹¹ For the identification of the various names, I rely on Charles-Picard (1959:91).

in first-century Mactar (cf. Charles-Picard 1959:76, 149). However, the individual cases put us on our guard against writing off entirely or too hastily the minority local languages. At any rate, the dating of individual cases certainly does not allow us to infer any generalized dating for the phenomenon of acculturation at hand, inasmuch as the move away from indigenous to Roman culture within the family occurred at different times in different social classes and areas of the Empire.

In this connection, I mention here a slightly more recent case involving the career of Q. Lollius Urbicus. The army, together with the bureaucratic structures of the Empire, is of particular interest to the historical linguist because of the possibilities it offered for upward social mobility. Roman culture did not impose preconditions on the integration of non-Latin peoples into Roman culture. In this way it was possible for those in the provinces to achieve social success, first on an equal footing with those of Italic origin and eventually with all members of the Empire. Exemplary in this respect is the case of Trajan, a citizen from the provinces, albeit of Italic origin, who achieved the high office of Emperor in AD 98. Almost a century later, in 193, an African of Libyan origin, Septimius Severus, rose through the ranks of the army to become Emperor. Less than half a century later there followed a period during which soldiers from the most diverse of provinces, including half-Barbarians like Maximinus Thrax, were to ascend to the throne, for by this time it had become normal for the legions to acclaim and impose, as Septimius Severus had done, generals of recent and even quite superficial Roman heritage.

If the history of the Empire highlights just a handful of such striking and truly exceptional cases of elevation to the highest ranks of imperial office, the same cannot be said of the lower imperial ranks which were often occupied by recently Latinized peoples from the provinces, who had moved great distances from their original homeland to integrate not only into the now Roman civilization of their own provinces, but also to move within the cultural and social circles of the entire Empire. Significant in this respect is Millar's (1984:132) observation that 'both equestrian and senatorial careers generally show a marked absence of geographical specification [...] in this respect the Empire seems to have been an integrated system, which showed no tendency to regional compartmentalisation'.

Our knowledge of Quintus Lollius Urbicus is based on inscriptions and a few other sources (Petersen 1970:87f.), but above all through the inscription placed under the statue dedicated to him by his fellow citizens of Castellum Tidditanorum or Tiddis, a small village 16 km to the north-west of Cirta (the

present-day Algerian city of Constantine), then part of the province of Numidia. From this inscription we learn that Q. Lollius Urbicus had begun his career as a *quattuorvir*, an official responsible for the upkeep of the roads, before moving on to become *tribunus* (commander) of the XXII Primigenia Legion, *quaestor urbanus* (magistrate in charge of Roman tax collection), *legatus* (assistant) to the governor of Africa, *tribunus plebis* (magistrate for the masses), *praetor* (magistrate), and *legatus* (commander) of the X Gemina Legion; in 133–35 he accompanied the Emperor Hadrian as his envoy on a campaign in Judaea, then became a fœtal priest (guardian of the peace), *consul suffectus* (substitute consul, one of two chief governing magistrates) by the year 138, and governor of Lower Germania, stationed in Cologne. Later, between 139 and 142, he became governor of Britain, where he succeeded in moving the border north of Hadrian's Wall to what would later become known as the Antonine Wall; he perhaps subsequently became governor of Africa and, finally, governor of Rome some time after AD 150.

The conclusion to be drawn from this is an important one: 'It need hardly be pointed out that at no other period of history could the second or third son of a Berber landowner from a very small town in the interior enjoy a career which took him to Asia, Judaea, the Danube (where the 10th Gemina was stationed), the lower Rhine and Britain, culminating in a position of great power and honour in the capital of the Empire to which all these regions belonged' (Wells 1984:247).

The examples that I have cited, chosen quite randomly and without doubt open to criticism on a number of counts, demonstrate that the process of Latinization must have been long and complex. Roman imperial society appears to us as a melting pot, though not in the same way the term is used in relation to modern society to refer to the mass assimilation of, at times, enormous numbers of immigrants settled in foreign lands far away from their roots. In the ancient Empire, by contrast, we are dealing with a much more complicated process, consisting, above all, in the acculturation of large numbers of compact groups of indigenous peoples who remained in their homeland where they had always lived.

Moreover, Latinization represents just one aspect of a historical process which a recent book has termed *Becoming Roman* (Woolf 1998). I mention Woolf's work on Gaul here because it expertly illustrates how *becoming Roman* implied living through a period of heightened tension between forces of cultural convergence and resistance (on which see in particular Bénabou 1976). The aim was not so much to assimilate to the other inhabitants of the Empire as to 'acquir[e] a position in the complex of structured differences in

which the Roman power resided' (Woolf 1998:243). This system of differences was to a certain extent independent of geographical location within the Empire, insofar as it was not to be found in one place rather than another. Instead it represented 'the set of manners, tastes, sensibilities and ideals' (p. 241) that were to be found in all areas where this ideal was shared. The existence of this 'symbolic center' (Woolf 1998) was intimately tied to the existence of the Roman political community, as was the assimilation process, increasingly spreading across and through geographical and social space (from top down, from the cities to the countryside, and with differences from one locality to another), giving rise to a process involving language which lasted for centuries.

Although Woolf's thesis is convincing, it deliberately leaves out a factor of fundamental importance: Christianity. If the thesis turns out to be correct, as would appear to be the case for certain aspects of his arguments, we would expect Latinization to have been thrown into crisis, or at least deeply affected, by the end of the Empire. But instead the process continued uninterrupted: with the exception of Britain and Africa, almost all remains of the pre-Roman languages disappeared and the old imperial territories became, if not homogeneously Latin-speaking, almost entirely Romance-speaking. The fact remains that, from the fourth century onwards, there arose alongside *becoming Roman* a parallel model, which in time was to prevail, that of *becoming Christian*, of which Latin formed an essential part, at least in the West. Despite leading to (or perhaps at the very least facilitating) the end of the Western Empire, this model also ensured the future survival of the Roman tongue.

In discussing the eastern Pyrenees some years ago, Juan Corominas expressed some views which I believe still hold true today, and which can be applied, with the appropriate changes in emphasis, to all cases of contact between Latin and the pre-Roman languages:

The problem is that the process of Romanization is viewed too simplistically; we seem to believe that everybody began speaking Latin or Romance overnight, giving up the pre-Roman language for good: the first generation would normally have spoken Basque whereas their children's generation would have abruptly switched to Latin. In reality, however, the facts must have followed a similar course to that which can be observed in the recent history of the Basque-Navarrese country.

Both languages existed side by side for centuries with many generations of bilingual speakers [...]

In these areas [Upper Aragon and Pallars], the languages are constrained, not by geographical, but by social boundaries: noblemen, the clergy and

the upper classes spoke Romance since early times, whereas peasants, tenant farmers and herdsmen remained faithful to Basque for generations. Even when the peasants living in the villages adopted Romance, the herdsmen and the farm workers remained steadfastly attached to the language of their ancestors; also the towns and villages, which were quickly Romanized, were surrounded by Basque, which was initially still spoken even in the suburbs, but finally only in the mountain villages or in small hamlets and farms. Nonetheless, most people, both in the hamlets and the villages, were forced to still understand Basque, which they no longer spoke within the home, remaining more or less perfectly bilingual for many generations to come.

(Corominas 1965:119f.)

We must emphasize, however, that such a long period of coexistence of the two languages did not mean that individuals or groups of speakers remained inert when faced with the threat of the loss of their original identity (cf. Varvaro 2005:117–33, on which much of the present discussion is based). In his *Apologia*, written during the reign of Antoninus Pius (AD 138–61), Apuleius, himself an African from Madaura, denies using magic to gain the attentions of the widow Pudentilla of Oea (= Tripoli). In actual fact, he had met the widow through her elder son, Pontianus, an old student friend from Athens, who had encouraged their marriage. When Pontianus died, Sicinius Emilianus, the brother of Pudentilla's first husband, befriended her younger son, his young nephew Sicinius Pudens, making an ally of him in his attack on Apuleius and encouraging him in all his vices. For his part, Apuleius paints a decidedly negative picture of the boy and his habits, including his speech: 'he only speaks Punic, with the exception of a few Greek words passed down to him from his mother; as for Latin, he neither has the desire nor the ability to speak it' (98,8, Vallette 1924:117). There can be no doubt then that in second-century Oea it was not unusual for Punic rather than Latin to be used among wealthy families; if anything, the prestige language would have been Greek. Of course, Apuleius condemns this in the same way he condemns the young boy's behaviour, although his perspective is that of a rhetorician defending himself before a Roman court. Sicinius Pudens, by contrast, has 'neither the desire nor the ability' to speak Latin: the young man's conscious refusal to assimilate linguistically, which makes him a figure of contempt in the eyes of his stepfather and all those who have accepted the process of acculturation, provides us with valuable evidence of the existence of non-conformists.

Clearly it is not always easy to distinguish between cases of anticonformist reaction through refusal to assimilate and simple inability through long-standing allegiance to tradition. The latter, which is also a precondition of

the former as well as a frequent driving force behind man's tendency towards self-assertion, must have been stronger among women.

Although Pudentilla was open-minded enough to have taught her son a little Greek, the example of Septimius Severus' sister at the end of that very same second century, also an African from the city of Leptis Magna, was not an isolated case. Her knowledge of Latin was reported to be so bad that it made the emperor blush, and he promptly sent her home.¹²

And yet Septimius Severus himself had retained his African accent.¹³ The slow pace of assimilation among women, which can readily be assumed on the basis of what we know about the social structures of the ancient world, is moreover confirmed by other evidence. This leads us to think that bilingualism, an inherent part of each stage of language change, was often prolonged and reinforced in the Empire on account of the fact that the indigenous language remained the children's first language even after the adult males of the family had adopted Latin.

5. Latin and the other languages of the Empire

Following the expansions of the Republican Era and the early years of the Imperial Age (cf. §2), Latin found itself alongside numerous languages of many diverse linguistic affiliations, necessarily giving rise to extensive bilingualism. Just how such bilingualism spread throughout the Empire and how it then progressively disappeared are, however, much more difficult questions to answer. We would be entirely mistaken to imagine that the Roman provinces underwent a process of rapid monolingualization. In actual fact, they were never monolingual and the dominant picture of Latin as a unitary language spoken over a homogeneous area proves entirely inaccurate for all periods. The reality was, rather, quite different.

Above, I referred to the overwhelming linguistic heterogeneity of the Italian peninsula at the time of its annexation to Rome. The situation was not any different outside Italy. Even in Sicily, the oldest of the Roman provinces, Greek was spoken in the eastern and southern coastal colonies from at least Catania to Agrigento and Selinunte, Punic in the western Carthaginian settlements (including Marsala, Mozia and Palermo), Elymian, a language about which we know very little, around Segesta and Erice, Sicel – apparently related to

12 Ael. Spart. Sev. 15,7, in *Script. Hist. Aug.*, ed. Hohl *et al.* (1965:i,148); for the English translation, see Magie (1967: 407).

13 'His voice was clear, but retained an African accent even to his old age', Ael. Spart. Sev.19,9, Magie (1967: 419).

Latin – in the interior, and Oscan in Messina where it had recently displaced Greek. The result, as can be seen, is a patchwork of languages that was repeated, albeit with different varieties but in a similar fashion, across all the provinces.

Here is not the place to list the western pre-Roman languages that would later be replaced by Latin and Romance. In many cases we know little more than their names and, perhaps, a few words that have come down to us through Latin sources. The most important of these in terms of distribution and because they have left us with some textual evidence, generally (but not only) inscriptions dating from after the time of Roman annexation and the dissemination of a culture written in the Latin alphabet, are Lusitanian, Celtiberian and Iberian in the Iberian Peninsula, Gaulish and Ligurian in present-day France, Germanic in the Rhineland, British in Britannia, Celtic around the Upper Danube, Illyrian in Dalmatia, Sard in Sardinia, and, of course, Greek in the ancient colonies along the Mediterranean coasts from at least Ampurias to Marseille and Nice and as far south as Naples and Calabria. Today we are fortunate in having Adams's (2003) excellent study which offers both a very rich overview and detailed analysis of bilingualism in the Roman world, paying particular attention to the Italic languages (Oscan, Umbrian, Venetic, Messapic), Etruscan, Gaulish, Punic, Libyan-Berber, the languages of the Iberian Peninsula, Germanic, and – in the eastern part of the Empire – Aramaic, Hebrew, Egyptian, Getic and Sarmatian, and Thracian.

At the time of the fall of the Western Empire (AD 476), not all of these languages had been replaced by Latin and some of them never were. Basque, a non-Indo-European language, continues until the present day a pre-Roman language, although which one is still a very much-debated issue. The same is true of Albanian, an Indo-European language. Breton, a Celtic language, appears to have been imported into French Brittany from Britannia in the last years of the Imperial period or perhaps even later. Modern-day Berber most certainly continues a pre-Roman language, probably Libyan, which was widely spoken south of the imperial border. It is not easy to establish when the languages which were unable to withstand the impact of Latin actually died out. At most, all we can say is when our last texts or records for these languages date from. For example, there is no trace of the languages of the Iberian Peninsula after the second century AD (with the exception of the ancestor of Basque, which might have been spoken in ancient Aquitania; see Trask 1997), whereas Punic was spoken until at least the beginning of the fifth century AD and Gaulish for another century still (cf. Varvaro 2004:97).

Although not exploring the finer details of ancient bilingualism, scholars of the Romance languages have long speculated about the possible role of substrate phenomena in the formation of the Romance languages (see also Sala, this volume, chapter 6). A substrate is defined as a language which, after a period of bilingualism, is replaced by another language in which it has left some recognizable trace. We must be sure, however, to make a distinction here. It is beyond doubt, and readily demonstrable on the basis of reliable sources in many cases, that the languages once spoken by the local populations which progressively came under Roman rule left many traces and also lexical loans in Latin, which subsequently passed into the Romance languages. It is quite a different matter, though, to claim that the substrate influenced the phonetic, phonological, morphological or syntactic structures of Latin. This has been the greatest area of speculation among scholars, although the results are negligible and are anything but reliable.

Substrate theory goes back to G. I. Ascoli, who in 1882 formulated a set of methodological principles. Yet the idea that proof of substrate influence can be gleaned from the modern distribution of a phenomenon coinciding with the areal distribution of a purported substrate language fails to recognize that such equivalences are always approximate. Not only can we never know the exact extent of the area covered by the presumed substrate language, but it is also misleading to think that we are dealing with continuous and compact areas. Without doubt the biggest problem, though, is that no account whatsoever is taken of the fact, at times fully demonstrable, that the modern distribution of the phenomenon is not the same as in medieval times. To take just one example, the change *-MB-*, *-ND-* to *-mm-*, *-nn-* in southern Italy was once considered a feature of indisputable Oscan origin, in part because of the supposed coextensiveness of the two linguistic areas involved, but the diffusion of the phenomenon can demonstrably be retraced to a restricted area of central-southern Italy during the early medieval period (Varvaro 2004:180–98). Even if the phenomenon, which is nothing more than a straightforward case of assimilation, had been caused by the substrate (though see Adams 2007:406–21 for convincing counterarguments), it would have been such only in the original area and the (relative) correlation between the modern area and the ancient area of Oscan speech would prove entirely irrelevant to the question. Similar considerations apply in the case of cacuminalization of *-LL-*, long attributed to a Mediterranean substrate, which not only displays an extremely uneven distribution (cf. Rohlfs 1966:§§234–35), but which on the whole cannot be traced to earlier than the end of the medieval period or the start of the modern period (cf. Caracausi 1986:122–55, esp. p. 143).

Although the effects of the substrate prove very difficult to verify, except in the case of lexis, this does not automatically imply, however, that the many and very different languages that once coexisted alongside Latin in Roman times did not have any effect on the latter. At the same time, this implies the rather banal, though often forgotten, fact that most Latin-speaking families had at some point in their history changed language, with all the consequent implications this had for speakers. Neither should it be forgotten that multilingualism in the Empire ensured that everyone was aware of the heterogeneity of language and the differing expressive possibilities offered by each language.

6. Linguistic variation in the late Empire, and substandard Latin

In accordance with a tradition that goes back to at least the humanist period, it has long been the norm to consider Latin as referring solely to the language of writers, both great and minor, of Roman literature, a practice that goes hand in hand with that of equating Latin with the established models of the literary language as instilled by grammarians. As has been observed for some time, this naive conception is not supported by the texts themselves. Yet Schuchardt's (1866–68) insightful study, while perfectly illustrating the early variation present within the language, had the undesirable consequence of popularizing, probably well beyond the author's own original intentions, the concept of 'Vulgar Latin', which, in my opinion, has greatly harmed the development of research in this area ever since. His *Vokalismus*, like all of his later works on vulgar Latin, is given over entirely to reporting deviations from the classical norm, irrespective of period of attestation (from Plautus to the later authors), region or text type. This assorted mass of evidence has been portrayed in terms of a misleading synchrony, syntopy and symphasy to create a non-existent system, certainly distinct from that of the literary norm but above all viewed as an alternative system existing within a sort of diglossic situation. The only exceptions to this rather simplified and deliberately paradoxical picture that I have presented here are those monograph studies that limit themselves to the study of a body of quite homogeneous texts (such as the Pompeii inscriptions, which all come from the same area and same period (AD 79 or just before), the letters of Claudius Terentianus (cf. the recent edition by Strassi 2008) or the *Albertini Tablets*, although the results of these analyses are, in turn, portrayed as belonging to one and the same system, as if Pompeii in the final years before AD 79, Egypt under Trajan

and southern Numidia of AD 484–96 can be considered similar cases ultimately to be treated on a par with each other.

This portrayal is clearly absurd. It is patently obvious that Latin, a language spoken for such a long time, over such an extensive area and by individuals who on the whole had learnt it as an L2, was not lacking in substandard diatopic, diachronic and diastratic variation. Yet, the study of these substandard phenomena proves extremely difficult, since they must have been typical above all of the spoken language, whereas we only have access to written records, which, by their very nature, are of little use in research of this type. First, we obviously have no texts written by the uneducated masses. Second, in many cases, including, for example, inscriptions or ancient manuscripts, we are dealing with texts written by professional scribes and hence subject to greater scrutiny than texts written by individuals for personal reasons. Third, the texts which have come down to us through the manuscript tradition, whether literary or technical in nature, are often not only the work of professional scribes but were also often copied decades or centuries after the original.

Despite these severe limitations, the study of substandard Latin is still to some extent possible, and today we are fortunate in having Adams's (2007) admirable synthesis which attempts to integrate the results within a framework of diatopic variation. His conclusion, to which I fully subscribe, is that '[w]e should get away from the idea that Latin was monolithic until a very late date, when some catastrophic event caused it to "split up"' (Adams 2007:725).

There are however some direct testimonies of a metalinguistic awareness of such variation.¹⁴ One of the most important is that of St Augustine, who in the *De ordine*, composed near Como in AD 386, writes to his mother: 'Should I say that you will easily attain a state of language free from faults of expression and pronunciation, I would certainly lie. For I myself, upon whom there has been a great compulsion to learn these things thoroughly, am still criticized by the Italians in the matter of many sounds within words, and they in their turn are criticized by me in the matter of sound. It is one thing to be secure in one's training, another in one's birth' (translation from Adams 2007:193f.).

Although Latin clearly must have displayed regional variation, as confirmed by the large number of metalinguistic testimonies to this effect, Latin textual scholars have always been aware of the problem that texts from before the

¹⁴ Adams (2007:685–87) has collected and studied an impressive body of such examples with his survey of over fifty textual references between the second century BC and the fifth century AD.

fourth century AD cannot be localized to a specific area of the Empire on the basis of their linguistic features.¹⁵ By contrast, one only has to think of how many early Italo-, Gallo- or Ibero-Romance texts can be readily localized in the same way. Adams (2007:726) sums up his findings by way of the following list: ‘signs of differences between the north and south of Italy in the late period (in the translation of Oribasius), lexical variations within Gaul (and particularly terms restricted to the south), the Latin of the Gaulish countryside contrasted with that of Aquitaine, and several micro-communities with local usages’, but is forced to admit on the same page that ‘we cannot map dialects in the manner of traditional dialectologists’. This amounts to saying that there is no recognizable homogeneity in such variation, be it in relation to the areas of the future Romance varieties or otherwise.

It is only in the case of lexis that we are able to spot diatopic differences and verify these with corresponding Romance forms. I give here just one example of the many highlighted by Adams. In the middle of the fifth century, Eucherius, the Bishop of Lyon, wrote: ‘there are those who think that owls, popularly called *cauannos*, are nocturnal birds named after the cry they produce.’ His testimony is confirmed by the note of the Berne scholia on Vergil, where we read: ‘*ululae*: birds named after their cry, the diminutive of which [word] is *uluccus*, as the Italians say; this bird the Gauls name *cauannus*.’ Here there is a clear distinction between a general term (*ULULA*), the local Gaulish term (*CAUANNUS*) and the local Italic term (*ULUCCUS*). These same distinctions are found in the corresponding Romance forms (cf. Adams 2007:251f.): *CAUANNUS*, a Celtic word, continues in Gallo-Romance, as well as in eastern Emilian and the Ladin spoken in Cadore (*REW*, 1787; *FEW* 2/1, 548 e 550), whereas *ULUCCUS* survives in northern and central Italian dialects (*REW* 9038a).

However, this and many similar lexical cases, often less transparent than this particular one, are of no use in constructing a systematic picture of the regional varieties of late Latin. At the substandard level, Latin appears to be rich in variation, yet it is not systematic, seemingly not producing dialects and failing to prefigure in any way subsequent Romance variation. But just how should this lack of correspondence be interpreted? Variation in late Latin is undoubtedly relevant for what was to come later, inasmuch as it undermined the cohesion of the Latin-speaking area. To a certain extent it represents an

15 It must be said that Adams’s efforts in this respect have not produced, to my mind, very impressive results. The first texts which can be unproblematically localized to a specific area are the *Mulomedicina Chironis* and the *Peregrinatio Aetheriae*.

essential, though not sufficient, precondition since there is no easily definable continuity between the regional phenomena of late Latin and the different Romance varieties. For example, the linguistic features which can be used to localize a later Latin text, say from Charlemagne onwards, are completely different from those found in earlier texts. Romance variation does not appear therefore to be the direct continuation of late substandard Latin.

7. A very sub-substandard level?

Here we ask not only whether there existed variation within standard and, in particular, substandard Latin across the Empire, and in what it consisted, but also whether such variation can be considered the precursor of the Romance languages. The results of Adams's in-depth investigation provide an affirmative answer to the first question, but a negative answer to the second. Substandard phenomena, especially within the field of lexis, follow a pattern of diatopic distribution which becomes increasingly pronounced towards the end of the Empire and especially in the period immediately thereafter, but the areas in which they are distributed are generally entirely unrelated to those of the later Romance varieties.

The hypothesis, which I advance here, that in the late Empire there was a linguistic level even below that of substandard Latin, which with some trepidation I hazard to call sub-substandard Latin,¹⁶ is based on the observation, which in itself is not that original, that there are a certain number of common features shared by all (or almost all) Romance varieties which are documented since the earliest written documents but which, however, fail to surface in any (or almost any) earlier Latin text. The existence of these shared features across all Romance varieties cannot therefore be due to their swift dissemination during the course of the ninth and tenth centuries, when contact between the different Romance varieties was certainly not of an order to justify such a wide dissemination.

The existence of such features has been known for some time, as a result of which it has been hypothesized that the relevant shared innovations must go back to an earlier Latin stage. Furthermore, this hypothesis is in full accord with the teachings of comparative linguistics: if two or more languages share an innovative feature that has not been transmitted from one of these to the other(s), it must be retraced to a common ancestor. It was on this basis that the

¹⁶ Whatever designation we use, I fear, as happened in the case of 'vulgar Latin' and 'pre-Romance', it will be transformed into a fictitious autonomous system.

concept of proto-Romance was conceived (cf. the lucid summary in Dardel 1996a and references therein).

Although spoken Latin obviously lies outside our direct field of observation, Dardel (1996a:91) claims that ‘this difficulty can, to a certain extent, be overcome through the reconstruction of the parent language of the Romance languages on the basis of the Romance languages themselves, with the help of the genetic comparative method’. This is widely acknowledged to yield an abstract result: ‘The historical parent language is a concrete datum, although not attested. Proto-Romance is an abstraction, derived from the Romance languages by means of an extrapolation which goes back several centuries in time’ (p. 91).

It strikes me as highly unsatisfactory to invoke and apply exclusively and rigidly the comparative method, originally defined in relation to distant and hypothetical Indo-European and only capable of yielding abstract results, for a historical period in which we are fortunate enough to have a relatively rich textual record. To adequately deal with such a historical period we need to approach it as historians, whatever the difficulties we might encounter. On the other hand, recourse to abstractions should only be made when all other avenues of enquiry are closed off to us. While it is not our aim here to question the validity and utility of the comparative method, one major problem that it poses is its inescapable assumption that the many forms under comparison must all be traced back to a single original form. This view, however, is a dangerous fiction which is bound to mislead: the period of the formation of the Romance languages was undoubtedly characterized by enormous diatopic and diphasic variation, far from displaying a unitary system across the whole area of the future *România*.

At the same time, I fully acknowledge that it seems a priori difficult to believe that, in a period of general upheaval in the communicative structures of western European society, there could have existed, despite failing to leave any trace whatsoever in the textual record, a variety which, like all those reconstructed through the comparative method, was characterized by a relatively homogeneous structure. In short, what we are dealing with here is an exercise in abstract linguistics, *more geometrico demonstratum*.

I begin by considering what traces there are in Latin for lenition, a phonetic development which proves extremely important not only because it distinguishes the Romance languages from the parent language, but also because it differentiates between two large Romance areas. The process consists in the weakening of consonants in intervocalic position or when followed by a liquid, resulting in the voicing of original simple voiceless obstruents and in

the fricativization or complete loss of original voiced obstruents.¹⁷ Ultimately, lenition is a case of assimilation, which could in theory arise in all areas, but which in the Romance languages displays a compact distribution in an area to the north and west of a line running across the Italian peninsula from La Spezia in the west to Rimini in the east.

Cases of lenition are not unknown in Latin texts. Already in AD 37–39 we find, outside the future leniting Romance area, attestations of TRIDICI and TRIDIGI (from TRITICUM ‘wheat’) in the Murecine Tablets from near Pompeii (Wolf and Crook 1989:15.3.1, 16.3.5). At Pompeii itself, hence in the year AD 79, TRIDICUM and TRIDICI are once again attested, as well as the proper names AG(G)GRATIS ‘Acratus’, ARPOGRA ‘Arpocra’, PAGATUS ‘Pacatus’, VIRIODAL ‘Uiriotal’ (Väänänen 1959:53f.). Perhaps from the year 217, near Aquincum in Pannonia, we find EXTRICADO < EXTRICATUS ‘disentangled’ (CIL 3, 3620), and from the third century in Mérida (Extremadura, Spain) IMUDAVIT < IMMUTAVIT ‘changed.3SG’ (CIL 2, 462). In some relatively ancient papyri originating from outside the subsequent leniting Romance area we find AUDEM < AUTEM ‘however’, PECADO, PECCADIS < PECCAT- ‘sin, error’, and perhaps also EXPEDIVI < EXPETIUI ‘I desired’ (Väänänen 1982:113). As for the original voiced obstruents, we find AVSTO for AUGUSTO (CIL 2, 5728) in Spain, AVSTA (CIL 8, 9877) in Africa in AD 452, MAESTRE for MAGISTER ‘master’ (CIL 3, 14730) in Dalmatia, CALCOSTEIS for CHALCOSTEGIS ‘bronze roof beams’ in the *Appendix Probi* (of disputed African or Roman origin), and VINTI < UIGINTI ‘twenty’ and TRIENTA < TRIGINTA ‘thirty’ in Gaulish inscriptions (Väänänen 1982:113–15). In the fourth century, Filastrius uses PUDORE for PUTORE ‘stench’ (Richter 1934:155); at Terni in 491 there occurs DEB(OSITA) IN PACE for DEPOSITA IN PACE ‘laid to rest in peace’ (CIL 2, 4339), followed a year later by TERRE MODUS for TERRAEMOTUS ‘earthquake’, and then in the *Itala* manuscripts of the New Testament we find QUODANNIS for QUOTANNIS ‘each year’ (Richter 1934:155). In the *Formulae Andecavenses* (from seventh-century Gaul) there occur THEUDERIGO and THEODORIGO alongside THEODORICO (MGH, *Formulae*, p. 2, ll. 22, 26–28), as well as the hypercorrect spelling IUCUM for IUGUM ‘yoke’. In Rome we find an undated AGOLITUS (Richter 1934:135) for ACOLITUS ‘acolyte’, and still in Italy in the Edictum Rhotari (AD 643) the forms FOGUM and FOGOLAREM occur for FOCUM/FOCULAREM ‘hearth’ (MGH, *Legum* IV, p. 34, l. 3).¹⁸

17 Cf. Lausberg (1956–62:§§347–405) and Loporcaro (volume I, chapter 3, §2.2). I will not consider here the reduction of geminates to single consonants, which appears to be a later phenomenon.

18 Some of the material presented here comes from Richter 1934 (§§108, 155, 160), a source which, however, must be used with some caution, taking care to eliminate the less

However, it is not until the Merovingian charters (625–750) that the lenited forms begin to occur in abundance (cf. Viellard 1927). In just the XII Royal Diploma, dated to between 657 and 673, we find *MATRIGOLARIS* (for *MATRICULARIS* (ABL-DAT) ‘man on the list of the poor’), *PODIBAT* (for *POTERAT* ‘could.3SG’), *MATRIGOLARIE* (for *MATRICULARIE* ‘woman on the list of the congregation of widows’), *GRADANTI* (for *GRATANTI* ‘rejoicing’), *MOVILE* (for *MOBILE* ‘moveable’) and *PRINCIBEBUS* (for *PRINCIPIBUS* (ABL-DAT) ‘leaders’). Cases of lenition are, however, attested before this date (Viellard 1927:45, 52), including, not surprisingly, in the Chronicle of Fredegar where we find *negare* (for *NECARE* ‘to put to death’; iii. 19; p. 100, 28).

As in other cases, including, for instance, the retention or loss of word-final -s, in the midst of the imperial period the phenomenon is barely documented and its distribution does not correspond to that subsequently found in Romance. It is only after the fall of the Empire that the attestations become more frequent and begin to cluster in the later Romance leniting area.

Let us now move on to consider a case which strikes me as typical, involving a grammatical category of the utmost importance. Latin, unlike Greek,¹⁹ lacked the category of article, which, by contrast, is found in all Romance varieties. In all such varieties the definite article is etymologically derived from forms of the Latin demonstrative, generally *ILLE* ‘that’, but in some cases also *IPSE* ‘-self, same’;²⁰ apart from Romanian, where the article (at least today) is postposed, while in all other Romance varieties it is preposed.

It is highly unlikely that the definite article represents an innovation which began in just one Romance variety and subsequently spread to the other languages. Besides, to my knowledge, nobody has ever advanced such a hypothesis. While admittedly in some of our earliest Romance texts, such as the *Strasbourg Oaths* and the *Placiti cassinesi* (see Pountain, volume I, chapter 13:§2.2.1.1), the definite article is not attested, although reflexes of *ILLE* are found in the clitic system, this is probably a consequence of the juridical nature of the texts, where the use of the article was felt to be too vulgar and hence avoided. Without any claims to exhaustiveness, the earliest occurrences of the

convincing examples and those which are attested in much later copies of the original manuscripts. Often different editions of the same text do not correspond, precisely because they are based on different manuscripts. Stotz (1996:§§159, 184, 204) also cites several forms, though not their sources, thereby rendering it impossible to verify their authenticity.

19 In referring to the Greek model of the article, I am thinking here not of learned influence, but rather the language of all those, both free men and above all slaves, who came to Latin through Greek.

20 Sornicola (2008) correctly excludes any value associated with that of the article in the early medieval use of *IPSE*.

article, both in texts of a predominantly Latin and Romance nature, include the following (care has been taken to cover the main geographical areas, despite the chronological discontinuity this inevitably involves):

710	Montmarcq (Oise)	'illo teleneu . . . ad illo marcado' and 'caret de illo the[l]en[j]io, de illo mercatho' (Sabatini 1996:90) 'the tax . . . to the market' and 'is without the tax, the market'
post 750	France	'(et) ipssa cuppa frangantla tota ad illo botiliario frangant lo cabo at illo scanciono tollant lis potionis' (Avalle 1983:27) '(and) let them break the whole drinking-cup, let them break the head of the wine steward and let them take the drinks from the cup-bearer'
774	Verona	'uno casale et lo campo foras in fundi Pupiliano' (Arnaldi and Smiraglia 2001:1349) 'a farm house and the field outside on [the] Pupiliano estate'
779	Lucca	'in rio qui dicitur la Cercle . . . in la Cercle' (Sabatini 1996:80) 'in [the] river which is called the Cercle . . . in the Cercle'
804	Valpuesta (Burgos)	'de alia parte de illo molare' (Gifford and Hodcroft 1966:1.8) 'on [the] other side of the flat-topped mountain'
831	Liébana (Cantabria)	'casa et horreum et illa binia novella' (Gifford-Hodcroft 1966:3.11) 'house and barn and the new vineyard'
ante 850	Rome	'non dicere ille secrita a bboce' (Sabatini 1996:173–217) 'don't say the secrets [= prayers said by the priest concluding the offertory of the mass] aloud'
c. 860	<i>Eulalie</i>	'li Deo inimi, les mals conseilliers, la polle, li rex pagiens, la domnizelle, la mort' (Henry 1953:3) 'the enemies of God, the evil counsellors, the girl, the pagan king, the young lady, the death'
post 872	Anastasio Chr.	'illos trecentos servos' (Arnaldi and Smiraglia 2001:226) 'the three hundred servants'
886	Teramo	'terra de illi filii Rassuloni' (Arnaldi and Smiraglia 2001:226) 'land of the Rassuloni sons'
943	San Juan de la Peña	'denante illo abate' (Menéndez Pidal 1956:333) 'before the abbot'
950	Padua	'terra casalina iuxta la calcaria' (Arnaldi and Smiraglia 2001:1337) 'domestic land next to the lime kiln'
958	Aragón	'homines de illas billas' (Gifford-Hodcroft 1966:87.6) 'men from the villages'
967	Oña (Burgos)	'cella Sancti Vincenti de la mata' (Menéndez Pidal 1956:337) 'chapel of Saint Vincent of the shrub'
971	Ibeas de Juarre	'qui est a la fonte de Nafarruri' (Menéndez Pidal 1956:337) 'which is at the source of Nafarruri'
978	Covarrubias	'de la Torquiella . . . usque uenit al Servo et del Servo usque . . .' (Menéndez Pidal 1956:337) 'from the Torquiella . . . until he comes to the Servo and from the Servo as far as . . .'

988	Nocera	'pecie quod vocamur "a la Fusara"' (de Bartholomaeis 1899–1901:267) 'fields which we call "the Fusara"'
992	Gaeta ?	'ribo della pillara' (de Bartholomaeis 1902–1905:27) 'stream of the pillars'
ante c. 1030	Tolosa (Spain)	'non tolran lo castel de Dormian Atoni' (Débax 2002:69) 'they will not take the castle of Dormian Atoni'
c. 1050	Umbria	'da lu battismu' (Monaci 1955:6.4) 'from the baptism'
1080–85	Logudoro	'su toloneu' (Monaci 1955:4.7) ²¹ 'the tax'
c. 1100	Rome	'fili de le pute' (Monaci 1955:7.1) ²² 'sons of bitches' (lit. sons of the whores)

In unambiguously Italo-Romance literary texts such as the twelfth-century *Ritmo Laurenziano*, the article appears from the very first line (here and below highlighted in bold): *Salva lo vescovo senato* (Formentin 2007:33) 'Bless the judicious bishop'. Similarly, in the earliest known Catalan text, the *Homilies d'Organyà*, the article appears twice in the opening line, a translation of Ecclesiastes 1, 2: *Totes les coses del segle son uanitats* (Miret i Sans 1915:39) 'All the things of the world are vanities' ('Vanity of vanities, all is vanity'). *OnCat.* (6: 457–60) cites many toponyms which incorporate forms of the IPSE article, for example *Sescloses* (< 'the mountain passs') (AD 891), *Sesrovires* (< 'the oak groves') (AD 1091), *Sacosta* (< 'the slope') (AD 1148), *Çagruya* (< 'the crane (bird)') (AD 1167).

I now turn to a third case, that of the forms of the future. Latin had a synthetic future with such forms as CANTABO 'I shall sing', DELEBO 'I shall destroy', UENDAM 'I shall sell' and AUDIAM 'I shall hear', of which the Romance languages only preserve a few isolated residues (e.g., the verb 'be' in early French: *ier, iers, ert, ermes, ertes, ierent* < *ERO, ERIS, ERIT, ERIMUS, ERITIS, ERUNT*; Lausberg 1956–62:§918). There was also a periphrastic future formed from the active future participle and the verb 'be' (e.g., CANTATURUS SUM 'I am about to sing'), which failed to leave any trace in Romance. The Romance forms which came to replace the Latin future all involve a periphrasis with the infinitive: UOLO CANTARE lit. 'I-want sing.INF' (Romanian), DEBEO CANTARE lit. 'I-must sing.INF' (Sardinian), UENIO AD CANTARE lit. 'I-come to sing.INF' (Surselvan) and, in particular, CANTARE HABEO lit. 'sing.INF I-have', which in only a few

²¹ For other Sardinian examples, see Blasco Ferrer (1984:85).

²² Conspicuous is the complete absence of documented examples from Gaul, present-day France. Although Vielliard (1927:183) lists about ten examples of *ILLE* in its article function from as early as AD 670–71, they all involve contexts in which *ILLE* would prove entirely normal in the classical language. For that reason, they have not been included here.

early northern Italian dialects, some southern Italian dialects, Sardinian, Portuguese and occasionally in Romanian presents the opposite order, sometimes with an intervening preposition (AD 'to', DE 'of', *da* 'from'), e.g., southern Italian *aggò a cantà*, Pt. *hei-de cantar*.

CANTARE HABEO thus continues in Portuguese, Spanish, Catalan, French, Occitan, Engadinian and Italian. In all these varieties, HABEO has given rise to an inflectional ending, although in Portuguese, as well as in old Castilian and old Occitan, the infinitive and the inflection can still be separated by a clitic pronoun (e.g., OSp. *darme has el libro* lit. 'give=me you-have the book') such that the two original components preserve a degree of autonomy.

Now, although the HABEO + infinitive periphrasis proves relatively common already in Latin, we still have to ascertain whether in these uses HABEO (or DEBEO or UOLO) preserves its full original semantic value or whether it has already been completely grammaticalized and hence semantically bleached. Furthermore, it is also necessary to find documentation of forms in which both parts of the original periphrasis are already intimately fused together.²³ And indeed they already appear fused together in the earliest Romance texts. For example, already in the *Strasbourg Oaths* we find *salvarai* 'I shall save' and *prindrai* 'I will take' (alongside older *ert* < ERIT 'will.3SG be'); *non tolran lo castel* 'they shall not take the castle' in the Toulouse Oath from c. 1030 (Débax 2002:69); *jrás* 'you will go', *farás* 'you will do' and *tornarás* 'you will return' in the San Millán Glosses (Menéndez Pidal 1956:361); *farai* 'you will do' in the Umbrian Confession Formula; *monsteroll[o]* 'I will show it' in the *Ritmo Laurenziano*; *parterimu* 'we will leave', *adrenderimo* 'we shall give back', *atverimu* 'we shall have' in the Fabriano charter from 1186; and *amerò* 'I will love', *dirò* 'I will say', *farò* 'I will do', *scanerò* 'I will kill', *serò* 'I will be' (for the Italo-Romance data, cf. Monaci 1955 *passim*).

A scrupulous examination of the Latin attestations does not reveal incontrovertible proof of the existence of the new future.²⁴ Adams (2007) often touches upon this problem, noting with a liberal dose of caution that 'habeo + infinitive, the forerunner of the Romance future, seems particularly common in African texts [...] its frequency in Africa could possibly show [...] that it

23 Salonijs (1920:283 and n) claims that the Latin examples are rare and adds some further examples from the medieval Latin Alexander Romance from the Naples area. For the UOLO + CANTARE periphrasis, cf. Adams (2007:730).

24 Several scholars have explored this avenue, including, among others, Salonijs (1920:282–90), Mørland (1932:144f.) and Adams (1991). The last highlights the importance of the position of HABEO with respect to the infinitive, noting how only when the latter precedes HABEO do we see the first signs of the Romance future.

started in Africa [...] But even that is doubtful. [...] The syntagm infinitive + *habeo* ends up as the Romance future, it is true, but it is not always straightforwardly a future in Latin texts. [...] it is far from being a grammaticalised future' (pp. 729–30). Equally prudent in his assessment is Herman (1998b), pointing out how HABEO + infinitive has 'a meaning which, although inevitably referring to an action situated in the future with respect to the event time, also incorporates a "modal" value, at least in addition to its purely temporal value; in Tertullian, however, whole sequences of examples with a clearly temporal value can be found [...] whereas other texts, of a popular nature, do not show any trace whatsoever of the use of this temporal periphrasis. [...] there are no grounds to believe that it functioned as an established replacement for the ancient future since the time of the Empire' (p. 19). For Herman, these observations, together with the fact that the forms of the new future are not uniform across Romance, are proof that the innovation occurred after the fall of the Empire, establishing itself most quickly in France and more slowly in Spain.

These latter conclusions are, in my opinion, unconvincing. Herman acknowledges that the passage from the Chronicle of Fredegar ii.62 (p. 85, 32), where one reads 'ET ILLE RESPONDEBAT: "NON DABO". IUSTINIANUS DICEBAT: "DARAS"' ('And he replied: "I will not give [them to you]". Justinian said: "You will"'), in which an initial DABO 'I will give' is followed by a DARAS (< DARE 'give.INF' + HABES 'you have'; Fleischman 1982:68), provides proof that there were already grammaticalized forms of the new future type by the second half of the seventh century. Yet the sentence is attributed to the Emperor Justinian (sixth century) and is alleged to have been pronounced in Nisibis in Mesopotamia, rather strange circumstances for an early example of the Romance future, to say the least. Yet in 1977 another example, from about a century earlier, was discovered on the inscription of a tomb in the Merovingian cemetery of Ledoix-Serrigny (Côte-d'Or), where we read: LANDELINUS FICIT | NUMEN | QUI ILLA POSSIDERAVIT VIVA | VSQUI ANNUS MILI IN D(EO) (Stimm 1977) 'Landelinus made [this representation of ?] God; may he who shall possess it (POSSIDERAVIT < POSSEDERE + HABET) live for a thousands years in the Lord'.

This attestation from central France from the second half of the sixth century is too early to indicate that the grammaticalization of the new future occurred after the fall of the Empire and is too isolated (even if the interpretation of Fredegar's *daras* is correct, it would, in any case, have been used in the East) for it not to be considered something of a taboo: these new synthetic forms were considered to be extremely vulgar, as is, not coincidentally, the language of Landelinus' tombstone.

How are we then to explain what we have just observed in relation to consonantal weakening, the definite article and the future? We can immediately dismiss the hypothesis of a dissemination of these innovations in the early Romance period, and the hypothesis of a proto-Romance variety strikes me as historically implausible since, above all, it faces the difficulty of explaining the alternation between *ILLE* and *IPSE* forms of the article and the competition between alternative forms of the new future. Instead, I believe it necessary to assume that, before the eighth century, there was in Latin a low substandard linguistic level in which, in the western area, intervocalic consonants tended to undergo lenition, and in which the demonstratives had undergone a functional shift towards the function of definite article, and new forms of the future had emerged. There can be no doubt that the use of the definite article was fiercely repressed to such a point that it never appeared in writing, not even in the writings of the uneducated, until the eighth century, whereas the new future appears just once in each of the sixth and seventh centuries. In the spoken language these innovations must have been widespread in all areas, albeit not necessarily used consistently (and certainly not associated with high usage). By the time that the reference model of written Latin (which did not allow lenition, articles or the new synthetic future type) had distanced itself from spoken Latin and had begun to display differences across the different areas, in some (or all) of these areas the article and the new futures had already been generalized.

The phenomena which can be attributed to this low level of substandard language, and which prove relevant to the making of the Romance languages, are numerous. Here it is not possible to examine them all one by one, but a list, albeit not exhaustive, would include: the change in stress and concomitant restructuring of the vowel systems; spontaneous and metaphonic diphthongization; palatalization; the loss of the nominal case system and, in part, that of the neuter; and the formation of new passive, periphrastic past and conditional paradigms.²⁵ Not all these phenomena can be considered on an equal footing. For instance, evidence of the weakening of the case system is not uncommon in our written texts, whereas diphthongization is barely attested. Of course, the substandard certainly would have also included phenomena which prove totally irrelevant for the formation of the future Romance varieties.

By way of a final consideration regarding the article, it is interesting to note that in medieval documents from Cava dei Tirreni and Gaeta in southern Italy, but also in other areas, we often come across the formula *'LOCUM* [or some

25 For detailed discussion of all these phenomena, see volume I.

other specification] QUI ART. + X DICTUR' ('place which ART. + X is called').²⁶ It would not seem foolhardy to interpret this formula as the expression of an awareness of a dual linguistic level, namely Latin on the one hand without articles and substandard, but by now Romance, with articles on the other. In short, this is a case of switching between two linguistic levels which are by now quite distinct.

8. Was there ever a catastrophe?

Assuming that there was a linguistic level below that of substandard Latin, capable of explaining the shared innovations of the Romance varieties, we still have to explain how the change in the structure of the communicative system came about which brought to light this lowest sociolinguistic level. The earliest explanation, which goes back at least to the humanist period, sees the 'corruption' and 'barbarization' of Latin as a result of the Germanic incursions.

Today, nobody believes in such a theory, not least because we now have a much better understanding of the not always dramatic conditions under which the Germanic populations settled, over a long period of time, across the Empire (cf. Wolfram 1998). However, the conviction that these invasions provided the necessary, though insufficient, prerequisite for the making of the Romance languages has still not disappeared today. I cite here just one example, which I and most other scholars still believe to be accurate: 'The Germanic invaders created the political conditions throughout the *România* which permitted linguistic fragmentation, reinforced the autonomous development of the ancient regions formerly under Roman rule, which now fell outside of the sphere of linguistic influence that was Rome, and enabled the formation of the Romance languages' (Pfister 1978:70).²⁷ Against the background of this premise, I shall now try to demonstrate that the historical process involved was extremely complex.

26 Above I have only cited the example '*pecie quod vocatur "a la Fusara"*' (*fusara* refers to the pit used for macerating hemp), but the formula (or its variants) proves very frequent: '*ubi "a lu Valneu" dicitur*' ('[a place] known as the "Bath"'), '*Johannes qui dicitur "de l'ancilla Dei"*' ('Johan who is said to be "of the handmaid of God"'), '*ubi "a lu Labellu" dicitur*' ('[a place] known as "the Wash Tub"'), '*ubi "a la Statua" dicitur*' ('[a place] known as the "Statue"'), '*ubi "a lu Pratu"*' ('[a place] known as the Meadow'), etc. (all examples taken from de Bartholomaeis). Identical formulas are now documented and studied in Carles (2011:294–99) for the Auvergne region.

27 Here I gloss over the problem of the effects of the Germanic superstrate, as highlighted in particular by Gamillscheg (1934–36) and Wartburg (1950). While there can be no doubt about the influence of Germanic on lexis, in other areas the effects of the Germanic invasions must have been above all indirect, as will be argued below.

Above, we witnessed the possibilities for social mobility open to high officials during the time of Trajan and Hadrian through the example of Q. Lollius Urbicus. His career progression, as we noted, was not in any way out of the ordinary. It was quite normal for emperors, holders of civil and military offices, legions and, let us not forget, merchants and missionaries of different eastern religions, especially Christianity, to travel from one end of the Empire to the other. The rich, especially members of the senatorial class, owned lands in different and often distant provinces of the Empire, which they would often visit. Moreover, they were always represented on the ground by their local delegates. It must not be forgotten that mobility decreased as one went down the social scale, with, for example, peasant farmers travelling little, if at all. However, they enjoyed vertical relations which sooner or later would lead back to individuals in contact with the provincial centres and, directly or indirectly, with Rome or the later centres of the Empire. In short, people frequently crossed the length and breadth of this open space that was the Empire, although some areas must have seen considerably less movement than others. Nonetheless, there are clearly no grounds for speaking of isolated or distant areas, except in a relative sense.

During the imperial period (and perhaps even earlier) there were also concrete and visible objects to remind everyone of the actual size of the Roman world that could be crossed by all. These were the epigraphic itineraries, of which we now only have a few remains.²⁸ The most important text of this type is what remains of the three sides of such an itinerary from Autun (Saône-et-Loire), which can be read in *CIL* 17/2, n° 490. One of the three sides must once have contained all the road stations in the journey from Rome to Autun, with their associated distances; all that remains today is the section BONONIA – MUTINA – PARMA. Side (b), which is the best preserved of all three, lists the stations and distances from AUTESSIODURO, namely Auxerre (Yonne), to INTARANUM, today Entrains-sur-Nohain, in Nièvre. Side (c) appears to have once contained the route from Autun to the Rhine, through Chalon-sur-Saône, Langres and Toul.

In Tongres, in Belgium, there are the remains of a similar document (*CIL* 17/2, n° 675), which listed the various road stations of at least three different routes, all starting from Tongres: one to Strasbourg (there only remains the part from Cologne to Worms), another to Boulogne-sur-Mer, and the last leading towards Amiens and Arras.

²⁸ My attention was drawn to this by my recent reading of Coulon (2009; cf. pp. 122–24 'Les indicateurs routiers').

Inscriptions of this type, which were clearly visible to all and probably not uncommon at road intersections, would have provided a very concrete idea of the reality of the Empire and its size. If today we have very few remains of such monuments, this is probably a reflection of the fact that at a certain point they were no longer of use to anybody. Everyone would therefore have been aware of the possibilities of travelling the length and breadth of this vast area that was the Empire, even those who in practice had no realistic chance of ever travelling across its territories.

There are numerous other examples which could be cited. Here, I confine myself to examining some historiographic texts and the oecumene implicit in each of these. I begin with Ammianus Marcellinus, the author of the important late fourth-century *Res gestae* in thirty-one books, of which only books XIV–XXXI have come down to us. Ammianus was born, between 330 and 335, in Antioch in Syria, and was a native speaker of Greek, but had learnt Latin, the official language of the military, in the army. Nonetheless, he wrote Latin with native proficiency. Assigned to Ursicinus, the head of the Eastern Army, Ammianus followed him to Gaul and then to Mesopotamia before returning to live in Antioch and travelling to Egypt and Greece. In 380 he finally settled in Rome, where he probably died around 395. His work was supposed to be a continuation of the histories of Tacitus from the accession of Nerva until his own time, including the catastrophe of the Battle of Adrianopole in 378.

Here, what interests me is the magnitude of the world in which the historian moved, which must have been the same for all educated members of the senatorial class and the readers of the *Res gestae*. In the pages of Ammianus' writings Rome continues, of course, to occupy a central position, although it was no longer the actual seat of imperial power. Indeed, the facts and vicissitudes of no other city are recounted with such care and attention to detail as those of Rome. It is not surprising then that he should often write about the imperial residences or that the East should constantly be at the forefront of his thoughts, as is to be expected given the continual and serious threat from Persia. He also speaks at length of the Rhine-Danube border and is clearly as familiar with the latter as he is with the eastern border. If we are left somewhat in the dark with regard to Africa and the Iberian Peninsula, this is simply because in the fourth century these were quite peaceful areas. The same is true of Britain. It is quite clear for Ammianus that cities such as ARELATE 'Arles', CARTHAGO 'Carthage', LONDINIO 'London', MAGONTIACUM 'Mainz' and SIRMIO (in Pannonia) are all considered of equal standing, namely centres of Roman life situated in different regions of the same oecumene. No area of the Empire, whether in the East or the West, escapes his attention, and

he appears to be at home in all of them. Yet it must be added that Ammianus' oecumene goes beyond the confines of the Empire, especially in the East, where his interests extend as far as India.

It is worth repeating here that Ammianus Marcellinus lived in the second half of the fourth century AD and was therefore essentially a contemporary of St Augustine, with whom he shared a similar view of the world. Gregory of Tours, by contrast, lived from c. 540 to 594, during the second half of the sixth century AD, some two hundred years after Ammianus. In 573 he was ordained as Bishop of Tours, after which he wrote his *Historiae Francorum*,²⁹ which, following an initial summary of the history of the world until c. AD 400, quickly moves on to a personal account of the episodes of the author's own time. We should immediately point out that Gregory's universal history proves rather different from the earlier accounts of the, still pagan, Roman historians such as Ammianus. In fact, Gregory confines himself to summarizing the story of the Bible, in which the history of Rome barely figures. For instance, Caesar and Augustus are mentioned in 1, 18, and of the former we are merely told that 'he became absolute ruler of the entire Empire', whereas of the latter we are reminded that Lugdunum (Lyon) was founded under his rule and, of course, that Christ was born.

Despite Gregory's background – he was descended from a senatorial family from Auvergne, one of the most profoundly Romanized regions of Roman Gaul – his outlook on the world was greatly different from that of Ammianus. Although the pages of his history are full of references to regions, cities and other minor localities in Gaul, which would not have figured in the accounts of previous historians, he appears to be familiar with very little outside the world of the Merovingian sovereigns. Although reference is often made to Rome and Constantinople, they always remain external points of reference. Moreover, all such references are never anything more than generic in nature, typically cited in relation to the various movements of the ruling families in their matrimonial alliances.

Let us consider a few examples. At his country mansion in Besslingen, situated in the midst of the Ardennes forest (today in Luxembourg), King Childebert called a meeting during which Queen Brunhilda lamented the fact that her daughter Ingunde had been taken away to Africa and that nobody had come to her assistance (viii.21). Whereas information about Besslingen is quite detailed, references to Africa, on the other hand, are vague, presenting it as a distant land. Gregory attended to Queen Ingoberga on her deathbed to

29 I cite here from Krusch and Levison (1951).

hear her will; a few months later she died 'leaving her only daughter, whom the son of a certain King in Kent had married' (ix.26). Gregory does not even know the name of this king of the nearby, but apparently remote, Kent. Saint Martin of Tours 'was from Savaria, a town in Pannonia' (x.31). The saints, in their time, moved around the Empire, but nothing is now known about the areas from where they came: Savaria (today Szombathely, in Hungary), had first been destroyed by Attila's Huns and then by an earthquake in the middle of the fifth century. Gregory fails to notice errors which would have been unthinkable in Ammianus: 'After which, however, the great cities of Antioch of Egypt and Apamea of Syria were captured by the Persians' (iv.40). For him, Babylon refers firstly to the biblical city and secondly, still on the basis of biblical sources, to the fortunately quite distinct city today known as Cairo: 'On its bank [of the Nile] there lies not the Babylon of which we spoke above, but another city of the same name in which Joseph built [...] some grain houses' (i.10). The whole of the East appears as a distant, almost fairy-tale land, where terrible adventures happen: in 591 'from beyond the sea there arrived in Tours a bishop by the name of Symon. This nobleman announced to us the destruction of the city of Antioch, and claimed to have been taken from Armenia to Persia where he had been held prisoner' (x.24; cf. Krusch and Levison 1951). Many similar examples abound in the work of Gregory. While Gregory is admittedly not writing a history of Rome, but a history of the Franks, it cannot be denied, and is not without significance, that his view of the world is infinitely more limited than that of Ammianus two centuries before him. The world, and perception and awareness of it, has shrunk extraordinarily.

The reader might well ask what all this has to do with the linguistic history that we are trying to reconstruct here, but I believe that this was the beginning of the real catastrophe which changed the Euro-Mediterranean world after AD 400. The Germanic invasions were not in themselves catastrophic, as the humanistic historiographic tradition would have us believe. The economic, demographic and social upheavals were not always as dramatic as could have been feared. In much of the Western Empire, with the exception of the marginal areas of the Rhine-Danube and Britain, the local populations, now entirely Romanized or in the process of being Romanized, had absorbed the new arrivals relatively quickly. However, the collapse of the complex imperial structures, which had not been adequately replaced by the emerging structures of the Roman Church, had radically changed these peoples' outlook on life. Their world was no longer synonymous with the Euro-Mediterranean area, and not even with the large administrative areas of the late Empire, but,

rather, with their local region or diocese and, in the case of the ruling classes, with the barbarian kingdom.

This drastic limitation was inevitably accompanied by a radical change in the prestige relations which had previously governed the way language was used. As long as the area of reference had been the Empire, normative linguistic usage had been based on the idealized model of the imperial court, whether in practice that was Rome, Milan, Trier or some other centre. The senatorial class embraced this model and, in one way or another, would transmit it, as a model, to the lower social classes where it would reach as far as the peasant farmers of the most isolated areas of the region. There was thus a common norm, albeit with all the regional variations discussed in section 6 above and with all the usual violations that we expect to find with all such models. Yet now, in this much-shrunk world, the immediate field of reference is at most the see of the bishop or the itinerant court of the king, in any case a non-Roman world whose language had very little in common with the model of the ancient Empire embodied in the literature of the classical period. The bishop was increasingly a local figure, with hardly any knowledge or awareness of how people spoke in neighbouring areas.

Many years ago Antonio Tovar (1964:129) wrote some words which still seem illuminating to me today: 'The shift from the Roman provincial capitals to the fields, the decadence of Tarraco and Nîmes and the emergence of the medieval capitals like Paris and León signified that people who were not saturated with Roman culture took the lead in the development of the language, and that rural tradition, hidden for centuries but keeping much of the linguistic habits of the ancestors and affecting Latin through generations of bilingualism, won the upper hand.'

This was the real catastrophe, not because it brought into linguistic usage new and unknown words, as in the case of a number of original Germanic lexical items, but because it destroyed the structure of the whole system, changing or even turning on their head prestige relations and reducing the previous norm to a literary model. This made it possible for different, regionally based spoken norms to emerge that were completely divorced from the normative model of the earlier Empire.

This process of 'the shrinking world' that I have tried to describe here is not only, nor especially, cognitive or psychological. It has a very concrete parallel (which might even be its cause) in the changes that happened in the economy and, in particular, in the world of commerce. Archaeological research in recent decades has brought to light evidence to demonstrate that in the period

of the late Empire it had become possible to produce and distribute with considerable efficiency not only luxury goods for the élite classes, but also large quantities of high-quality goods for the lower-middle classes, including those living in the more remote areas of the Empire. Indeed, it is even possible to reconstruct both the international and regional distribution networks used in the commerce of individual products.

The British archaeologist and historian Bryan Ward-Perkins (2005:104) convincingly maintains that '[i]n the post-Roman West, almost all this material sophistication disappeared. Specialised production and all but the most local distribution became rare, unless for luxury goods; and the impressive range and quality of high-quality functional goods, which had characterised the Roman period, vanished, or, at the very least, were drastically reduced, the middle and lower markets [. . .] seem to have almost entirely disappeared.' In many areas, the collapse of the Empire weakened the local economy to levels lower than those seen even before the Roman conquest. This did not bring about a recession, with the usual reductions in production volumes while still maintaining the overall system. Instead, there followed 'a remarkable qualitative change, with the disappearance of entire industries and commercial networks' (p. 117). Here we are most interested in the end of the interregional networks. In this respect, Ward-Perkins cites the example of Noricum in the time of St Severinus, as richly documented in the writings of his biographer Eugippius: 'Even local exchange had apparently been made impossible. Unsurprisingly, the import of goods into Noricum from afar had also become very difficult' (p. 135). Of course, the timing and speed with which this change took place would have been different in different provinces of the West, but the overall trend was the same everywhere.

Moreover, Ward-Perkins perceptively observes that 'because the ancient economy was in fact a complicated and interlocked system, its very sophistication rendered it fragile and less adaptable to change' (2005:136). The same must have happened with the system of communication. Had the Empire been subject to the diatopic and diastratic homogeneity that is misleadingly taken for granted, then the system of communication might certainly have shown greater resistance in reaction to the political, social and economic crisis of the post-Roman period. Yet the actual system of communication was extremely complex: although it had a common and high reference point, namely the literary language of culture, it was structured into different levels of social, regional and even local variation, splintering, in some cases, into what were remains of the pre-Roman languages. The way the system worked and the dynamicity it showed towards ever-increasing integration

presupposed that contact between different groups remained very much active and perceived indicators of prestige constant. With the collapse of these latter two, the system could no longer function, except perhaps at the more élite levels of society: in the same way that the circulation of luxury goods survived, so there survived an ever-dwindling number of the educated classes who were still able to use normative Latin and were familiar with its literature. The popular masses, on the other hand, could only draw on local products and the most uneducated of linguistic models.

9. New identities

I now return to our linguistic history. It seems to me that, for most scholars, the collapse of the complex system of communication of the late imperial period implies the triumph of local identity and individualism. According to this view, the different linguistic communities of western Europe underwent an extraordinary process of atomization, in effect the catalyst for the gradual formation of new groupings which, in the fullness of time, would become the Romance languages.

This idea is central to the work of Roger Wright, to which I shall return below, but it also seems to me to be implicit in James Adams's work, where, after accepting the concept of a Romance linguistic continuum and observing – quite rightly – that 'the naming of dialects [...] may reflect extra-linguistic, non-linguistic, realities' (Adams 2007:723), he cites Dante's *De vulgari eloquentia* ('On Eloquence in the Vernacular') and his examination of fourteen Italian dialectal varieties identified on the basis of regional divisions, concluding that 'we have to wait until about 1302 for someone (Dante) to attempt to count³⁰ the Italian vernaculars' (p. 726).

This view does not seem to me to be entirely accurate. Half a century before Dante, commentators on the story of Babel already listed Romance and non-Romance varieties (cf. Lusignan 1986) and, more than two centuries before Dante, Goscelin of Saint-Bertin had been able to provide a reasonable overview of western European communities on the basis of their language rather than through anything else. He identified ten groups within what we call today Gallo-Romance: Francians, Allobrogiens, Auvergnats, Berrichons, Normans, Manceaux, Angevins, Poitevins, Gascons and Catalans (cf. Varvaro 2007). For some time it has been noted that the French dialectal divisions correspond to the areas of the pre-Roman and Roman *civitates* ('districts'),

³⁰ Here I would add 'and to name'.

which subsequently became the episcopal sees of the early Middle Ages (Morf 1911). This state of affairs gave rise to a long-lasting sense of identity, ensuring that people continued to feel, for example, Auvergnats for centuries long after the Alvernii had ceased to exist and the population of the region had undergone all manner of changes. Regional identities are confirmed time and time again in Latin historiography of the early Middle Ages, when it was normal to speak of Lombards (in the broad sense of the term then still current, not in its modern narrow sense) and the Leonese were not the inhabitants of just León but of the entire region.

Although our documentation is, of course, limited and patchy, and cannot always be guaranteed to refer to homogeneous linguistic areas, I do not know of any counterevidence: between the seventh and eleventh centuries these linguistic communities, like all the other communities, shrank in size but never disappeared. Individuals were aware of belonging to groups which expressed themselves more or less in the same way and were therefore different from other groups which spoke differently, and often, so it was perceived, quite bizarrely. The satirization of different vernaculars, for example of Anglo-Norman with respect to Francien, is known to us through later texts (cf. Varvaro 2007), but from the silence, which for some centuries conceals such comparisons, I believe we can only draw one conclusion: the problem of linguistic communication had no political or cultural importance and perhaps was of no concern to anyone. This does not seem consistent with an overall picture of extreme fragmentation, which would have proved very dysfunctional.

The existence of a Romance dialectal continuum appears, albeit with some limitations, realistic, but at the same time it does not in any way rule out the existence of splits, showing little, if any, mutual intelligibility. To take the example of the Iberian Peninsula, there is no doubt that linguistic variation from Salamanca to Zaragoza shows all the characteristics of a continuum; but from Zaragoza to Lleida the situation is not at all the same: between the heart of Aragon and the eastern Catalan-speaking part of the region the degree of variation involved is much greater.³¹ When speakers recognize that they speak the same variety, this is not because boundaries between different varieties are sharply delineated, but is the consequence of a common recognition of a more complex identity than that based on linguistic data alone.

31 The boundary between Aragonese and Catalan, I note in passing, does not coincide at all with that separating the Kingdom of Aragon from the Principality of Catalonia, despite claims (for example, Wright 2003:352) that the dialectal boundaries involved can be traced back to original political boundaries and the imposition of prestige norms by those in power. Claims of this type often find no confirmation on the ground.

The fact that the continuum does not map onto a constant and even rate of variation and that there had always been quite a widespread awareness of individual identity helps us to explain a particular situation which posed a problem for nineteenth- and twentieth-century philologists: although the provenance of most of the earliest Romance texts can be identified by the presence of local regional features, they also display a number of supra-regional linguistic features. One only has to think of the language of the *Strasbourg Oaths* (see also Wright, this volume, chapters 3 and 4, Kabatek, this volume, chapter 5) and the hotly debated and still unresolved issue of the origin of the Gallo-Romance variety used in them. Similarly, it has been observed that the first word of the *Placiti cassinesi* (again, see also Kabatek, this volume, chapter 5), namely *sao* 'I know', does not perhaps represent the expected local outcome, namely *saccio* (< *SAPIO*). However, the problem is considerably more general than this, and has led many editors to the quite improbable conclusion that most texts were written in areas straddling linguistic borders.

While studying the earliest document from Liège more than half a century ago, Louis Remacle (1948) came up with the concept of *scripta* precisely to account for the presence of a more or less large number of non-local features in this as in all other documents from the *langue d'oïl* area of northern France. The term *scripta* has since spread to other, albeit inappropriate, uses, but was originally coined to refer to the mixed linguistic nature of written texts. While the concept is a valuable one, it perhaps has the disadvantage of giving the impression that partially similar phenomena are not found in spoken language, where there are no tendencies to minimize differences with other varieties by adopting characteristics of these. Of course, written language is by its very nature more reflective and the tendency to accommodate is greater. There are therefore good grounds to think that in the early Middle Ages the varieties spoken in Romance-speaking Europe (just like those in Germanic-speaking Europe) were felt to be readily reducible to a limited number, whose area of diffusion was relatively small, but even so not limited to a single village, and which allowed the possibility of adopting non-local features.

10. Logographic reading and new graphies

A preliminary and fundamental aspect of understanding the making of the Romance languages and one that we cannot ignore concerns the question of how the early documentation is to be interpreted. More specifically, how

should our texts be read? Above, I referred to cases where particular Romance developments are barely or not all documented, implying that I do not accept the hypothesis, which for a number of decades has enjoyed some popularity, that the spelling of written Latin in the early Middle Ages was logographic and not alphabetic.³² According to this view, written Latin texts before the time of Charlemagne are said to be Latin only in their spelling, a graphic representation reputed to have concealed the fact that they were actually read as Romance. Matters apparently changed, though, when 'Germanic scholars established, in the Carolingian realms, for the first time in any Romance-speaking area, an official spoken standard, in addition to the old written one' (Wright 2003:347). Nonetheless, this claim is entirely without substantiation. According to historiographical sources, the Carolingian reform dealt with the correctness of written Latin and the form of writing itself (witness the introduction of the Carolingian minuscule), but had nothing to say, at any point, about the spoken language; neither is it true that *Germanic scholars* were involved in establishing a spoken norm (of the three most famous scholars involved in the reform, Alcuin of York, Guido of Pisa and Paul the Deacon, only the first was 'Germanic' by virtue of being Anglo-Saxon),³³ nor are there any grounds for speaking of an *official spoken standard*.

Accordingly, then, this hypothesized additional Carolingian reform supposedly 'invented' medieval Latin and, by doing away with the logography which allowed what had previously been written in the same way to be read according to different varieties,³⁴ was apparently the reason why the different Romance orthographies were 'invented' only at a much later date as the result of conscious and deliberate efforts.

In actual fact, however, no writing system, Romance or otherwise, in normal use is able to capture pronunciation exactly, otherwise there would have been no need to develop the International Phonetic Alphabet (IPA). Of course, there must have been some degree of partial logography, some level of separation between writing and reading, in the Latin texts between the sixth

32 The main supporters of this hypothesis, defended with great tenacity although not necessarily with the same arguments, are Roger Wright (1982; 1994; 2003; cf. also this volume, chapter 3) and Michel Banniard (1992a; also this volume, chapter 2).

33 Paul the Deacon was from a noble Longobard family of Friuli, but was brought up in Pavia and lived as a monk in Montecassino.

34 For Wright these varieties are 'different' only to a certain extent, since he speaks of 'the existence of an essentially monolingual Early Romance (Late Latin) speech community until at least the ninth century: Since they themselves do not seem to have made any systematic distinctions between the speech habits either of different areas or of different social groups, we may well be confusing the issue if we make such distinctions ourselves for that period now' (Wright 2003:344).

and eighth centuries. For example, it is perfectly possible that an intervocalic voiceless consonant was frequently read as if voiced or a tonic vowel read as a diphthong without causing any problems. A page of Gregory of Tours would certainly have been read differently by somebody brought up and living on the banks of the Loire from somebody born in the vicinity of the monastery of Montecassino (where codex A1 in Krusch's stemma is kept). Perhaps the differences then were greater than they are today between, say, how *Le Monde* is read in Paris, Marseille or Brussels. It is quite another thing, though, to imagine that Gregory of Tours was read in the *langue d'oïl* in Tours and in some Italo-Romance form in Montecassino, and that his language was not universally recognized as Latin but, rather, as some form of the vernacular.

Undoubtedly there are texts where one can legitimately question whether they are written in Latin or the vernacular, for which the term *circa romançum* ('Latin almost Romance') has been reintroduced, somewhat inadvisedly (in my opinion), in the course of the last fifty years.³⁵ While there is admittedly a spectrum of variation between texts written by the educated and texts produced by those barely able to write, it is unrealistic to deny that this spectrum also had an upper range associated with the language of the educated.

The logographic hypothesis (which is nothing more than just that, namely a hypothesis) allows scholars to read into medieval Latin texts whatever they consider most appropriate,³⁶ but leaves unresolved many primary problems. First, it is obvious that the hypothesis that Latin was read as Romance cannot be maintained for those areas which never became Romance-speaking, such that a monk from Fulda must have read Latin differently from his fellow brethren in St Martin of Tours. The former is said to have employed an alphabetic reading system, while the latter used a logographic system. But how is it that nobody ever commented on such a difference, which would have made communication extremely difficult between two people who actually spoke Latin very differently? To my knowledge, nobody has ever thought to explain what a logographic reading of, say, Vergil, would produce, or what would have become of Christian hymnography, which was largely conditioned by musical considerations. It is easy enough to understand how PEDE 'foot' might have been read *pie*de or *piè*, but I fail to see how the synthetic LAUDATUR 'praise.3SG.PRS.PASS' ('it is praised') could be read as the analytic

35 This term comes from thirteenth-century documents from the Iberian Peninsula (cf. Menéndez Pidal 1956:454–60) and was reintroduced by Avalle (1983).

36 Sometimes to the point of extreme ingenuity, as in the case of Emiliano (2003a:294) who goes so far as to transcribe the first Latin document of the kingdom of Portugal in IPA, to be understood, of course, as Portuguese.

è lodato 'is.3SG.PRS praised.PTCP' in one place and as *est loué* 'is.3SG.PRS praised.PTCP' in another, or how a Latin future such as *AMABO* 'I will love' could correspond to distinct Romance future forms such as *amerò* or *aimerai*.

Wright has always thought that the differences between Latin and Romance, as well as those among the individual Romance languages, are essentially phonetic, even to the point of making such surprising claims as 'the spelling was the only problem that required the elaboration of new [Romance] written systems, since the representation of vernacular morphology and syntax required no such thought, for that was already there in their minds, and all of the details of normal word-order and syntactic construction could, if desired, be transcribed more or less direct in whatever orthographic form was deemed to be required' (Wright 2003:350). It is for this reason that Wright repeatedly bemoans the 'regrettable and avoidable political phenomenon' (p. 353, and *passim*) which brought about the end of a unitary logographic writing system which, according to him, even today would still offer many advantages.³⁷

The logographic hypothesis forces us to associate the Carolingian reform, which had always been understood as a moral, cultural, scholastic, grammatical and graphic restoration,³⁸ with values and aims which strike me as highly improbable. We have seen that it is not true that those who carried it forward were all men from non-Romance areas and were therefore used to the alphabetic reading of Latin; consequently, Alcuin's part in the reform has necessarily been exaggerated, with his *De orthographia*, which is never directly cited, being seen as an innovative work, which it surely was not.³⁹ Nor to my knowledge has there ever been given any serious consideration to the problem of what happened in the areas outside the Carolingian Empire such as San Millán de la Cogolla or Montecassino, where even the paleographic reform arrived much later.

In the same way that the 'invention' of medieval Latin seems highly improbable, the Romance languages cannot be said to have come into being through the 'invention' of a writing system. Writing is a technique which is learnt in

37 We read, for example: 'I still cannot help feeling that on the whole things would have gone better in Medieval Europe if the Carolingian reform had never happened, and the whole Romance speaking world had still preserved, even perhaps until the present day, only the old international written standard spellings of words, with all the attendant advantages, Late Latin, in short' (Wright 2003:357).

38 Cf. Brunhölzl (1975:243–315). I find it strange that no mention is made of the writing reform (the Carolingian minuscule), which is much better studied and easier to study, including its diffusion, which was a rather slow process (cf. Bischoff 1990).

39 The text can be read in Bruni's (1997) edition (cf. Brunhölzl 1975:273).

relation to a particular language; undoubtedly during the Middle Ages, and for many centuries afterwards, it was Latin which was taught and learnt, and not the vernacular languages. Those who use a writing system for a language for which the system was not originally intended can only operate by adapting it accordingly. This is still the situation today for Italians who, having learnt to write Italian, try to put down in writing a dialect with no written tradition.

There is no reason to believe that medieval practices were any different, and indeed this is exactly how the writing of Old English or Old High German emerged. Initially, writing in these languages was certainly motivated by political considerations of a cultural, and not a linguistic, nature. Exactly the same situation happens later with the Romance languages. When the Oxford codex of the *Chanson de Roland* was written, the revolutionary act was not so much the choice of a particular orthography, but rather the decision to treat a vernacular poem as if it were a Latin text and to put it in writing.

But, in any case, why would it have been necessary to 'invent' a Romance orthography? Individual vernacular words and phrases had long been written down before the appearance of systematic documentary or literary texts.⁴⁰ Although we still do not have an exhaustive history of educated and semi-educated spelling between the seventh and twelfth centuries, it would surely show that these cases are not the result of deliberate political planning nor a chaotic series of individual initiatives. Here too there exist traditions, inasmuch as the concepts of orthographic chaos and writing tradition are not irreconcilable.

But let us now look at some concrete examples. One of the problems facing those writing Romance words or texts is how to render palatal consonants, which were absent from Latin. In what follows, we observe through some limited and quite random surveys the solutions adopted for just two palatal phonemes.

In the western area of the Iberian Peninsula (Galician and Portuguese; cf. Kabatek, this volume, chapter 5) the palatal nasal *ɲ* and lateral *ʎ* are rendered by the following graphemes (Monjour 1995:701):

<i>ʎ</i>	<i>l</i>	<i>moler</i> 'woman, wife'	<i>ɲ</i>	<i>ni</i>	<i>tenio</i> 'I hold'
	<i>li</i>	<i>molier</i> 'woman, wife'		<i>nn</i>	<i>uinna</i> 'vine'
	<i>ll</i>	<i>moller</i> 'woman, wife'		<i>gn</i>	<i>signo</i> 'sign'
	<i>lly</i>	<i>Jullyo</i> 'July'		<i>nh</i>	<i>vinha</i> 'vine'

⁴⁰ The inventory and interpretation of examples collected in the vast and extremely useful volume by Selig (1993) and other contributors is judged 'worryingly old-fashioned' by Wright (2003:348). However, 'fashionable' is not always synonymous with 'sound'.

lh	<i>concelho</i> 'advice'	n	<i>senor</i> 'lord, gentleman'
llh	<i>vellho</i> 'old'	gnh	<i>pegnhorar</i> 'to attache'
lhe	<i>alhear</i> 'to find'	h	<i>uiha</i> 'vine'
		ny	<i>pecunyas</i> 'money'

The solutions in the Castilian area are equally heterogeneous (Menéndez Pidal 1956:52f.):

λ	li	<i>filio</i> 'son'	j	ni	<i>uinias</i> 'vines'
	lg	<i>Gilgelmus</i> 'William'		in	<i>uergoina</i> 'shame'
	gl	<i>maglolo</i> 'vine shoot'		ng	<i>Lecengana</i> 'Leciñana'
	lig	<i>meligor</i> 'better'		nig	<i>senigor</i> 'lord, gentleman'
	ll	<i>basallo</i> 'servant'		nn	<i>anno</i> 'year'
	l	<i>decolaren</i> 'they declare'		n	<i>Eneco</i> 'Iñeco'
	ill	<i>taillatu</i> 'cut.PTCP'		inn	<i>Noinnez</i> 'Nuñez'

For the Occitan-speaking area, it will suffice to note that in the glossary of Appel's (1930) anthology, containing mostly lyric texts which have often been regularized by their editors, the outcome of MULIERE(M) 'women, wife' is variously written as *molher*, *moylier*, *moillier*, *moller*, *moler*, *muller*, the outcome of MELIORE 'better' as *melhor*, *meillor*, *melor*, *mellor*, *milhor*, *milor*, and the outcome of SENIORE(M) 'lord, gentleman' variously appears as *senior*, *senior*, *signor*, *seinhor*, *seinnhor*, *engor*, *sennor*, *senor*. In documents from the Languedoc area we find the following spellings (Grafström 1958:209–14):⁴¹

λ	l	<i>melor</i> 'better'	j	n	<i>sener</i> 'lord, gentleman'
	il	<i>meils</i> 'better'		in	<i>seinor</i> 'lord, gentleman'
	li	<i>filia</i> 'daughter'		ni	<i>Dornia</i> 'Dourgne'
	ll	<i>moller</i> 'woman, wife'		ne	<i>vinea</i> 'vineyard'
	ill	<i>moiller</i> 'woman, wife'		nn	<i>vinna</i> 'vineyard'
	illi	<i>meilluramentz</i> 'improvements'		inn	<i>seinnors</i> 'lords, gentlemen'
	lg	<i>cosseig</i> 'advice'		ng	<i>ordeng</i> 'device'
	lh	<i>filha</i> 'daughter'		ing	<i>peings</i> 'pledges'
				ngn	<i>Alvergne</i> 'Auvergne'
				ingn	<i>luingna</i> 'distance'
				nh	<i>Rossinhol</i> 'nightingale'
				gn	<i>segnoria</i> 'lordship'
				ign	<i>seigner</i> 'lord, gentleman'
				hn	<i>Vinhes</i> 'Vines' (toponym)
				ihn	<i>destreihner</i> 'to compel'

⁴¹ See also Carles (2011:495), who studies the equally rich orthographic variation in Romance lexemes contained in Latin documents from the Auvergne region.

For France, rather than cite the forms in use during the twelfth and thirteenth centuries, it will suffice to recall here that in his 1470 transcription of the fourth book of the *Chroniques* of Jean Froissart, the copyist Buisset freely alternates between several graphemes in rendering the same word, including, for example, *nouveaux*, *nouveaulx*, *nouviaux* or *palais*, *palaix*, *pallais*.

For the Italian outcomes of these palatals, a quick glance at Monaci (1955:596f.) reveals the following variation:

Λ	li	<i>filioli</i> 'sons'	ɲ	ni	<i>senior</i> 'lord, gentleman'
	ll	<i>fillo</i> 'son'		nni	<i>giunnio</i> 'June'
	lli	<i>mollie</i> 'woman, wife'		ngni	<i>compangnia</i> 'company'
	lg	<i>molge</i> 'woman, wife'		ngn	<i>Bolonga</i> 'Bologna'
	lh	<i>filholi</i> 'sons'		ngi	<i>Spangia</i> 'Spain'
	lhy	<i>bactalhye</i> 'battles'		ign	<i>seignor</i> 'lord, gentleman'
	lg̃l	<i>milgl̃ore</i> 'better'		gni	<i>giugnio</i> 'June'
	lg̃li	<i>velg̃lio</i> 'old'		gn	<i>signo</i> 'sign'
	gl	<i>mogle</i> 'woman, wife'			
	gli	<i>mugliere</i> 'woman, wife'			

This extraordinary polymorphy is not, however, limited to just a few graphemes. For example, in southern Italy during the Middle Ages the grapheme <ch> had for some time two distinct values, that of the voiceless velar stop [k] (cf. It. *che* 'that') and that of the voiceless palatal affricate [tʃ] as in *chi* 'there' (cf. It. *ci*), as well as a possible palatal value [kj] as in It. *chiave* 'key'.

If the 'invention' of Romance spelling (and through this the invention of the Romance languages) had really started from the top, then things would have been very different. The stabilization of an orthography for the individual languages was a process that lasted centuries and which required the intervention of genuine legislators, grammarians and even typographers. As still happens today, the desire to distinguish oneself from one's neighbours also played an important role in this process (for example, if the Castilians write *ñ*, then the Portuguese opt for *nh* and the Catalans for *ny*). But for centuries there was much experimentation, which gave rise to relatively long-lasting and competing traditions. It may be that in Castile the very active and prestigious school of scribes founded by Alfonso X made it possible for the court's preferred spellings to become established more quickly. However, even if we confine ourselves to the spelling used in the three remaining testimonies of Juan Ruiz's fourteenth-century *Libro de buen amor*, we are left with no doubt that the emergence of a norm was still very far off.

II. Conclusions

In conclusion, I fear therefore that we must abandon hypotheses that offer the attraction of novelty but nothing of realism, and return to the patient and meticulous, old-fashioned style analyses of texts within their specific context and in relation to their enormous variety, whilst tempering our interpretations of the data with a healthy dose of caution.

Those data, whose interpretation has been definitively settled, also tell us something else. The logographic hypothesis forces us, especially as regards morphology, to assume that many of the Romance innovations happened *after* the adoption of non-logographic spelling, since they are barely or not all reflected in previous writing. In short, the Romance future, conditional or passive are claimed to have emerged after the supposed Carolingian reform. Consequently, future forms such as *serai* 'I shall be' and *prindrai* 'I shall take' that we find in the *Strasbourg Oaths* must, according to this hypothesis, have become 'normal' only in the fifty years before 842. While this single example highlights the paradoxical nature of the logographic hypothesis, it also implies that all Romance developments and diatopic variation, which already appear to be well established since the earliest texts, must be compressed within a very limited time frame. Once again the role and place of phonetics is overestimated: we are forced to assume that while the relevant phonetic developments took place over more than half a millennium, the rest of the linguistic system was, by contrast, subject to an incomprehensible (inasmuch as unjustified and unrealistic) process of accelerated and late development.

What emerges from the texts is that the crisis which broke the bond between the spoken languages of the Western Empire and the Latin norm happened long before the advent of Charlemagne, namely during and immediately after the collapse of the Western Empire. With daily life now increasingly confined to more limited areas, including both long-established areas, such as the Roman *civitates* ('districts') subsequently transformed into dioceses, and newly established areas such as Burgundy and subsequently Normandy, local varieties far removed from the Roman norm soon acquired a new prestige. For some time these varieties were limited to the spoken language, whereas the written language, whose importance had considerably diminished in early medieval society and become the sole prerogative of the clergy, was still based on some form of Latin, be that good or bad Latin. But here and there, despite all the taboos, the spoken language does occasionally come through and, of course, more so in the writing of those who are less apt to stigmatize the spoken language. While the Carolingian reform remains

extremely important for cultural and religious issues, in the same way that the Carolingian Empire was for political history, there is nothing to suggest that it was important for the vernacular, which was finally and for the first time accorded some practical uses in the Council of Tours of 813.

After Charlemagne, society changed again, with the establishment of local powers and communities wanting to have a voice, and a readily understandable voice at that, on matters concerning them. It was the demands of everyday life – and not the political agenda of the powerful, who for centuries had no interest in language (cf. von Moos 2008) – that led to the use, in different types of text, of the spoken language. The ultimate emergence of the Romance languages (and their respective literatures) represents an extraordinary phenomenon in the history of European communities, which cannot be explained away in terms of the supposed political decisions of a few individuals.

Our ability to understand these centuries is considerably undermined by the difficulty of correctly understanding the enormous process of linguistic history during this same period. Many of us think of it, sometimes consciously but often not, in terms of a linear development from left to right, which up to a certain point we associate with Latin and thereafter with Romance. In essence, this image is implicit in the title of Lot's (1931) famous study 'A quelle époque a-t-on cessé de parler latin?' ('When did Latin stop being spoken?'). Framed in this way, the problem cannot be resolved, if for no other reason than because Latin remained a spoken language for many centuries after the Middle Ages, up until almost modern times.

Almost as naive, and perhaps even more misleading, is the frequently and widely accepted idea that spoken Latin is a socially lower variety of written Classical Latin, which in the early centuries of the Middle Ages underwent a gradual process of decay until the Carolingian reform, which brought Latin back to its original correctness and broke the bond between the spoken languages and Latin, thereby forcing the former to become languages in their own right (see Figure 1.2).

In fact it is on the basis of this hypothesis, which is generally not even expressed, that texts from the seventh to the ninth centuries (and even later) are read, imputing all instances of aberration from the classical norm to the vernacular. On the other hand, all early Romance texts are, of course, considered to be documents representing spoken usage. As a result there is a tendency to forget that *all* our texts, whether in Latin or Romance, are written texts, whose relationship with the spoken language of the same area or the same period needs to be defined on a case by case basis, and even then the

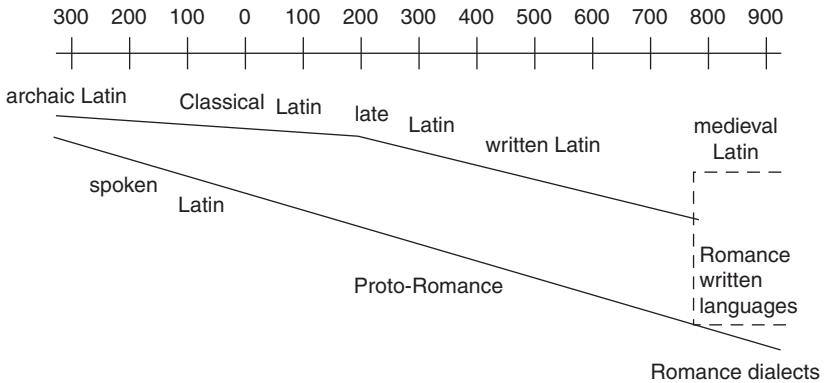


Figure 1.2 Schematic representation based on Kiesler (2006:12)

results are not necessarily unambiguous. Although Latin and vernacular are felt to be different languages, it is instructive to remember that they were both spoken languages, even if the number of speakers of the former was rapidly decreasing.

Like all the world's languages, Latin had always been, as was still the case around the year AD 600, a complex linguistic system made up of different levels of usage, ranging from the most prestigious literary norm to the lowest registers of the substandard. As the means of expression of a great culture and of a confident and self-aware civilization, the literary norm had remained compact and stable for a long time, but from AD 400 it too began to display regional variation. In the lower and spoken registers diatopic variation must have been quite considerable since the earliest times, and certainly the changing size of the speaker's world would have compounded this. This progressive diversification operated therefore both horizontally (between increasingly less similar linguistic varieties) and vertically, inasmuch as Latin remained a spoken language in clerical and learned spheres and was now also differentiated according to education and usage.

However, the system no longer had an unambiguous point of reference, either politically within the imperial structure or socially within the ancient senatorial class. Its very complexity pushed it towards breaking point, which it finally reached, albeit probably at different times in different areas. Gradually, individual communities became aware that they spoke something different from the Latin used by the learned and the Church. There thus emerged two parallel diaphasic systems: Latin, which ranged from the classical language based on the norms of earlier authors, and in any case still imitated by the

writers of the time, to the substandard language of the least educated monks; and Romance, which was different from region to region and was, at first, not employed for written and literary uses but gradually emerged as a written language by a process of trial and error, starting with the direct representation of spoken language and other contexts in which it proved appropriate and necessary for those who could write to put down on paper the form of these varieties. Only a modest percentage of the population of a given region would have continued to use the Latin system, which had remained more homogeneous between one region and another on account of its traditional points of reference, namely schooling, classical literature and the Church; this limited number of people also knew (and used when necessary) the other system, and it was certainly from these true bilinguals that the original push would have come to adopt the vernacular languages also for written uses.

The enormous inventory collated some years ago by Frank and Hartmann (1997) allows us to observe the relentless development of Romance textual production in the early centuries.⁴² It must never be forgotten, of course, that what is contained in this inventory represents just a part of all that was written and that pure chance will often have played a significant role in the possible transmission or loss of texts that have come down to us today, especially in the case of those texts where the subject matter and the level of language were more modest. The fact remains though, that broadly and relatively speaking the inventory of what still exists today is probably not much different from what would have been available in the Middle Ages. The differences and particular details that we note today between different areas, to the extent that they reflect earlier distinctions, must be taken to imply that the process by which the tradition of writing in the Romance vernaculars arose must have varied considerably from area to area.

From a very early date there emerges in the *langue d'oïl* area a large number of literary texts, at first religious and then also secular in nature, which find no parallel in quantity or quality with other areas and which soon also begin to exert an enormous influence outside the Romance-speaking area.⁴³ On the other hand, the use of the vernacular in non-literary texts emerges later, beginning to gain ground slowly from 1220.

In Occitania, by contrast, legal and documentary texts already appear from an early date and also in considerable number, as is also the case in

42 For a thorough documentation and study of Romance inscriptions prior to 1275, see now also Petrucci (2010).

43 The apparent precocity and volume of Anglo-Norman production is probably also the result of the better preservation of the relevant codices.

Sardinia (but not in Corsica). But whereas Sardinia has very few literary texts, southern Gaul quickly gave rise to a highly prestigious and widely disseminated school of lyric poetry, as well as literary texts of other types. The literary prestige of southern Gaul was such that Catalonia, for example, appeared to be nothing more than a literary appendage to the Occitan-speaking area, but here both documentary and literary prose was just as slow to emerge and spread as the use of the vernacular in legal texts.

The Italian peninsula presents, in its earliest phase (before 1200) and in the centre-north, a large number of heterogeneous texts, including literary texts. Subsequently, a strong lyric poetry tradition emerged along the Sicily-Tuscany axis, and immediately afterwards prose underwent an extraordinary expansion, especially within Tuscany, giving rise to a very rich body of Italo-Romance texts.

In the Castilian-speaking area, the vernacular continues to infiltrate Latin documents for a long time, but only becomes established as an autonomous code in both prose and poetry, and in both practical and literary texts, from 1200. Neither is the emergence of the vernacular any more precocious in Portugal.

Nowhere are we left with the impression that the process is instigated from the top through politicians or intellectuals. This hypothesis might have some credence in southern Gaul and in Sardinia, given the types of texts first produced in these areas, although we are dealing here with the two regions which were least subject to strong central powers. During these early centuries, Sardinia was unaffected by trends coming from the north, and Occitania was a land notoriously divided in terms of its power structures.

The study of these processes still requires in-depth investigation. We need to return to our texts, verify, as far as is possible, the percentage of those phenomena which have been lost, explore in detail the conditions surrounding each individual text and how each textual tradition arose. The European world from the tenth to the fourteenth centuries was extremely fragmented and what happened during that period cannot be readily reduced to simple processes. A fundamental fact such as the establishment of the cultural supremacy of France (understood here as the area between the Loire and the Meuse, to which we can add Norman England) followed the Crusades, when local writing traditions had, in part, already emerged. In Romance-speaking Europe, as in Germanic-speaking areas, the myth of the unified Empire will remain alive for many years to come, but the reality will remain one of division and differentiation.

The transition from Latin to the Romance languages

MICHEL BANNIARD

I. The issues

The title of this chapter alone raises a welter of issues which merit detailed consideration; at the same time, a contribution of this nature cannot hope to do justice to all the complexities involved. I have therefore chosen to focus in what follows on a particular problem – how to establish a chronology of the development of Latin into the Romance languages using the methodology of historical sociolinguistics (we might speak more accurately in this connection of ‘retrospective’ sociolinguistics), whilst at the same time attempting to model this change in a historically informed way.

To speak of a transition from Latin to the Romance languages involves certain assumptions. First, the distinction between singular and plural implies a shift from linguistic unity to linguistic diversity. We shall need to elaborate on this point below, since the problem to be addressed has two facets – the shift from one language to a different language, and the shift from one language to several languages. The history of Romance linguistics shows that these two issues have often been confused, so that diversification, often referred to as ‘fragmentation’, has frequently been seen as a cause of change in itself. This failure to distinguish between what should be two different areas of enquiry has regularly led to change being described in terms of external factors (‘barbarization’), and a consequent reluctance to seek out possible internal factors. The external hypothesis has proved all the more attractive because speaking of a parent language in the singular poses the conceptual problem of accounting for linguistic change within a unitary language. Since some might say that unity is in principle invariant, philologists have traditionally sought to address the problem by postulating a parallel language which represents disunity: so-called ‘Vulgar Latin’.

The term ‘transition’ also implies the existence of a relatively broad diachronic border zone where the common spoken language that goes in is Latin

and the one that comes out is Romance. It implies that the diachronic isoglosses do not bunch, but are more or less regularly spaced in time between the beginning and end of this zone. The terminology therefore implies a linear model of change. But this is clearly an a priori assumption; there is nothing to suggest that language change takes place progressively, at a steady pace, without hiccups, and without speeding up or slowing down. When evaluating work in this area, we should be aware that the fact that scholars have traditionally used linear models does not of itself make those models any more valid. There is no reason to believe that the modelling of language change is any more advanced, cognitively speaking, in the year 2013 than the modelling of human language itself; despite the tremendous progress made in the second half of the twentieth century, a satisfactory description of the neurological bases of speech is still a long way off.

Now that these preliminary issues are out of the way, we can focus the discussion that follows on the frequently formulated question: 'Who speaks what to whom, where, and in what circumstances?', adding the essential diachronic element 'When?'. This question has four complementary aspects to it: (1) geographical: Do the answers differ according to the region involved?; (2) chronological: Is there significant linguistic discontinuity according to the period we are dealing with?; (3) social: Should the history of the language take account of social stratification?; and (4) cultural: essentially the same question as the previous one, although answered from a different perspective.

2. Diachronic sociolinguistic models

2.1. *Problems and methods*

2.1.1. Synchrony and diachrony in sociolinguistics

Since this *Cambridge History of the Romance Languages* aims to offer a new approach to the problems of Romance linguistics, I make no apology for using research methods which are relatively novel, yet at the same time sufficiently established to have already yielded significant results (Banniard 1975; 1991b; 1992a; 1992b; 1992c; 1996b; 1999b; 2005c; Herman 1967; 1988; 1993; 1996b; 1999; Richter 1976; 1983; 1985; 1994a; 1994b; 2005; Stotz 1996–2004; van Uytanghe 1976; 1977; 1987; 1989; 1991; 1994; 2000; 2005; 2008; Wright 1982; 1991; 1995a; 1995b; 1997; 1998; 2000; 2003). I am referring to historical (or retrospective) sociolinguistics. In the abstract, this discipline aims to apply the findings of synchronic sociolinguistics to diachronic linguistics. This approach has been adopted by some contemporary historical linguists, although, in the

continuing absence of a genuinely interdisciplinary framework, they have had to invent their own ad hoc working methods. Certain features of synchronic sociolinguistics can be identified which have a direct bearing on diachronic studies (Calvet 1999; Fishman 1971; Labov 1972a; 1972b; 1994; Trudgill 1991; Weinreich 1953).

- (a) Synchronic sociolinguistics generally involves assembling a corpus, most commonly from a database of oral speech; in this respect, its methods are not dissimilar to those of traditional dialectology.
- (b) It then attempts to classify these data according to universal diatopic, diastratic and diaphasic models (van Deyck *et al.* 2004; 2005), combining speakers' subjective impressions and their objective reality (of which they themselves may be unaware).
- (c) Once a certain number of results have been arrived at, work can begin on modelling processes of dynamic variation in language; this work may lead to interesting conclusions, which account for paradoxes and suggest solutions to hitherto unsolved problems.

These procedures clearly cannot be carried over, without modification, to the diachronic study of Latin/Romance. The main issues which we have to deal with are the following:

- (a) Any corpus is limited by the fact that the documents to which we have access are written and deal with a restricted range of subject matter. Because we are dealing with a closed set of written materials, there is, alas, enormous scope for filling in the gaps by creative acts of the imagination. In this way, hypotheses are frequently advanced which appear to be descriptions of an earlier reality, but which turn out on closer inspection to be nothing more than modern scholars projecting their own mindset back into the past.
- (b) Synchronic sociolinguistics is a relatively young discipline, and is therefore still free to define itself to a large extent; however, historical sociolinguistics connects to an intellectual tradition going back well over 150 years (if we take the work of Diez as the starting-point of 'scientific' Romance philology). Although it builds on the findings of traditional philology, it also takes issue with them, with the result that one of the great achievements of social dialectology and sociolinguistics has been to unmask the frequent a priori presuppositions involved in many conceptions of language, and, in so doing, to raise important epistemological questions about the methods of traditional historical Romance philology

(Banniard 1999a; 2000b; 2006a; Herman 1992; 1996b). This process, set in train by what we might call the 'European Sociolinguistic School', has led to new models being put forward which have contributed substantially to our knowledge of the circumstances in which the Romance languages came into being, and, in so doing, have led to profound changes to traditional philological models (Koch 1995; 2008; Koch and Oesterreicher 2001).

2.1.2. The rationale of historical sociolinguistics

What, then, are the aims, methods and principles of historical sociolinguistics? For obvious reasons, this brief discussion will be restricted to the Latin West (Banniard 1992a:10–63).

- (1) The aim is to establish a chronology of the transition between Latin and the Romance languages.
- (2) The method consists of eliciting data from native speakers about their linguistic intuitions.
- (3) The basic principle is to give careful consideration to each and every factor which might influence the answers to these questions and our interpretation of them.

What are the detailed implications of these statements?

- (1) The first statement presupposes that a chronological basis can be established for the transition between Latin and Romance; in other words, that there are one or more significant thresholds before which Latin is a living language and after which it is a dead language, or, to put it the other way round, before which Romance is a virtual language and after which it is an actual language. From a purely linguistic point of view, this type of analysis has two consequences:
 - (a) we must agree on what 'living language' and 'dead language' actually mean;
 - (b) we must be able to distinguish between Latin and Romance typologically.

At least two paradoxes emerge at this point. The first concerns our instinctive feeling that, in order to die, a language first has to grow old. But this notion is an artefact of metaphor, and, unless external factors are at work, it makes little sense, diachronically speaking – a language is always young, since each generation of speakers brings it into being anew when they learn it. Latin never grew old. In other words, speech

has a continuous existence, and yet it is constantly being modified in ways which may lead to speakers changing not only the system that underlies it, but also the very language they speak. Therein lies the principal difficulty of this question. It is so difficult to formulate an answer that most attempts to do so have proposed external solutions; that is, they have claimed that the linguistic innovation took place elsewhere and only subsequently came to affect the language under investigation.

The second paradox, closely linked to the first, concerns the distinction between Latin and Romance. We cannot fail to notice striking differences if we compare a page of the *Aeneid* (c. 20 BC) and a page of the *Chanson de Roland* (often dated to c. AD 1100). However, certain caveats are called for. First, an impressionistic feeling of difference is no substitute for a proper linguistic analysis. Second, there is more to this comparison than just the time-span involved, because comparing a page of the *Chanson de Roland* with a page of Hugo's *Légende des siècles* (1859–83) will also produce the feeling that the two texts have little in common linguistically. Finally, if we look more closely, we shall find ways in which the languages of the two texts actually resemble one another. These resemblances may be few and far between; but they are there nonetheless. So the question is not a simple one, and Procrustean typology sheds little light on the diachronic issues involved, at least for the linguist struggling to come to terms with the unforgiving mass of textual detail. The problems grow worse if we compare late Latin and proto-Romance. It is at this point that it becomes absolutely imperative to adopt a sufficiently large number of explicit criteria. In this way, the diachronic linguist is in a similar position to the dialectologist puzzling over the problem of establishing geographical boundaries between dialects (Banniard 1980a; Brun-Trigaud 1990). Latin, yes – but which Latin? Or, rather, which Latins?

- (2) Eliciting data from native speakers of the past about their linguistic intuitions involves lengthy and difficult procedures, which I have described elsewhere. Rather than rehearsing the arguments yet again, I shall outline some of the problems which beset this method of enquiry, in the view of practitioners and sceptics alike.
 - (a) First, it is obvious that a crucial limitation of the *testimonia* garnered from a range of centuries, regions and authors is that they are produced by individuals who had some ability to write, and that such material therefore embodies a different culture and mentality from purely oral evidence. However, fundamentalists who object that the written and spoken languages are entirely separate (notably

anthropologists such as Goody 1977 and Ong 1982) are flying in the face of scientific findings concerning the relationship between the two codes. The written language is no more than a cultural variant or extension of the spoken language, as can be demonstrated by a *reductio ad absurdum*: no natural written language has ever come into existence except as an emanation from a pre-existing spoken language. True, access to the written code requires specific training which brings particular areas of the brain into play (Changeux 1983; Pinker 1994), but the cognitive processes involved are not so different from the acquisition of other expressive modalities, such as painting. Of course, whether or not social life is underpinned by writing has huge consequences for the evolution of a society (Lévi-Strauss 1962), but we should be wary of overestimating the significance of this distinction (Miller and Fernandez-Vest 2006). Adopting a Manichaeian approach to this problem leads us to a convenient but simplistic dualism: 'vulgar' = living = unwritten vs. literary = dead = written.

- (b) The second objection concerns the reliability of evidence from literate speakers of Latin. Latin literature defines itself in terms of linguistic difference, and those who can read and write are aware of prestigious literary models. Does this mean that evidence from literate speakers is inadmissible when we study variation in Latin, whether synchronic or diachronic? Many scholars would make this claim, and their position appears at first sight to have the merit of intellectual rigour. But these same linguists who, in their haste to pin down the authentic spoken language, tend to set aside data derived explicitly or implicitly from literate speakers are perfectly happy to rely on this type of evidence when discussing the 'chaos [sic] of the late Latin period'. And they have no problem in accepting uncritically the doom-laden pronouncements of grammarians (those natural pessimists). Once we actually attempt to chart sociolinguistic variation in spoken Latin on the basis of detailed textual evidence, we quickly reach the position where the data are so multifarious that it is difficult to paint a coherent picture (Reichenkron 1965; Müller 2001; Adams 2007). Finally, the notion that literate usage obscures the true spoken language is often the result of confusing genuine distinctions of style with imaginary (or, rather, aprioristic) differences between languages. There is a parallel here if we compare the language of Proust (1871–1922) with that of the letters written home from the front by soldiers serving in the First World

War: we are dealing with a single language, but one which has multiple instantiations (cf. also Wright, this volume, chapter 4).

- (c) The third objection is a quantitative one. It revolves around the fact that at least 90 percent of Latin-speaking society lies beyond the reach of any documentation, whether direct or indirect (Herman 1983; 1998b); it is made up of individuals we cannot observe (although written usage may not in fact be as nugatory in this respect as it appears; see Feugère 2004), and about whom any hypothesis can be put forward, because none can be gainsaid. Even so, in every area of the Western Empire where the Latin heritage was not obliterated by external factors (Slavonic, Germanic and Arab migrations), a Romance language is still spoken. This means that the language which developed in these regions developed from Latin, and must imply that Romanization was general, even as regards the 90 percent of the population which cannot be observed directly. Of course, some people will object that Romanization consisted in the spread of Romance, not the spread of Latin. On this point, we have to appeal to Occam's razor. If we know nothing of 90 percent of the population, then why assume that they spoke Romance rather than Latin? The simple answer is that this assumption has provided a neat solution to the problem of explaining the development of Latin into Romance by externalizing the phenomenon – it is a sort of diatopic counterpart of the diastatic account. It is simply an admission of defeat, a reluctance to explain linguistic change by internal factors, and at the same time it lures scholars into a sort of parallel universe, where any hypothesis is valid. In the absence of any evidence for the existence of a horizontal divide between the 10 percent of speakers to whom we do have access and the 90 percent to whom we do not, it is more intellectually reputable to admit the existence of a geographical continuum, without second-guessing the many variations and vacillations which must have existed within it. The countries of Latin America (Auroux *et al.* 1998) provide an insight into what the linguistic situation must have been at the height of the Roman Empire. Dialect surveys reveal substantial variation within the Portuguese- and Spanish-speaking communities, but there are no dialects showing features of any other type of Romance language.¹ Any foreign element derives

1 For the marginal case of Italian influence on the Spanish of Buenos Aires, see Jones and Pountain, this volume, chapter 10: §5.2.3.2.2.

from languages which were already there before the conquest and the heterogeneity of first-generation learners exposed to a new culture. Once several generations had acquired the new language, these speakers became part of the continuum.

- (3) In addressing the third point, we must consider issues which are more culture-bound.
 - (a) Both Latin philology and Romance philology have a very particular view of language change. Nineteenth-century scholars created notions of Latinity which were ethical rather than linguistic. Latinists, since at least the fifteenth century, have stubbornly maintained that there exists an unbridgeable chasm between the good Latin of the literate, which is accessible through standard texts, and the bad Latin of those who are illiterate or semi-literate, which is accessible through texts of inferior quality, possibly classical, but preferably 'late', and ideally medieval (Chomarat 1982). So, from the late Middle Ages onwards, a cultural divide was elevated to the status of linguistic dichotomy, and Latin philologists of the nineteenth century merely enshrined this social distinction in intellectual discourse. For their part, nineteenth-century Romance scholars, especially the German-speaking pioneers of the discipline, were imbued with the Romantic notion of a pure and exalted initial state which could only change for the worse (Fourquet 1980a; 1980b). Even a brief glance at a manual of Romance philology will reveal the striking persistence of a terminology which combines the outraged purism of Latinists confronted with deviation from the norm and of Romanists who can only perceive language change as linguistic degradation (Avalle 1965; Schiaffini 1959; Väänänen 1967). One might reflect on the social underpinning of this world view in the late nineteenth century, and on its persistence into the twenty-first; it is cultural exceptionalism born of historical accident. For my part, I prefer to reclaim the history of change in Latin for general linguistics, and to do so using as neutral a terminology as possible (Banniard 2005a).
 - (b) A neutral viewpoint is likewise lacking in studies of all the fields which impinge on these linguistic issues. The historical and cultural context of the period of transition between Latin and Romance was in the past uniformly discussed in terms which had contemporary resonance (Gibbonian, and then positivist), and implied a harsh stance towards the civilization of what was described as the 'Decadent Empire', the 'Decline and Fall of the Roman Empire', and so on. These gloomy

analyses were usually bound up with a morbid fascination for all things 'barbarian'. Romanticism and positivism colluded to bequeath to the twentieth century a vision of these 'Dark Ages' which is worthy of Dante's Hell. It was only in the latter half of the twentieth century that this perspective began to alter; by the last quarter of the century, this had led to a sea-change in our view of the period. Rather than allowing the pendulum to swing towards mindless exaltation of this era, the European Historical School has come to terms with its true nature and dynamic, to the extent that we now refer to it using the more neutral term 'Late Antiquity'; and although the barbarians still exist, modern scholarship is happier to do business with them (Pohl 1999; Le Jan 2006). So the symbiosis of 'barbarian times' and 'barbarian language' which was so beloved of philologists is now ruled out, because the first expression is no longer accepted usage. Unfortunately, this new view of the cultural context has not yet permeated the scholarship of many Latinists and perhaps the majority of Romanists (Zamboni 1998; 1999), who still lovingly retail the apocalyptic vision prevalent in the early twentieth century (Carrié and Rousselle 1999; Dumézil 2005; Geary 1988; Lewitt-Gibbon 2001; Toubert 2004). It is clear that the basic question of sociolinguistics – 'Who is speaking?' – cannot be answered unless we make every effort to understand the time, place, society and mindsets involved.

2.2. *Sampling the data*

2.2.1. A false assumption about the fifth century

These somewhat abstract methodological considerations may be illustrated with reference to recent discussions about significant documents, discussions which have involved specialists in a variety of disciplines. One example concerns the linguistic situation in Roman Africa in the fifth century and the identity of the common spoken language of this province. A sociolinguistically informed study by Vincent (2001) does not shy away from the problem of vertical communication, and judiciously refuses to rule out the possibility that illiterate speakers could understand Latin (unlike Lot 1931, who dismissed this suggestion out of hand). However, as Vincent does not believe that this automatically makes such speakers speakers of Latin, he revives the idea that there was 'two-speed' vertical communication, so to speak: one using *sermo humilis* (Auerbach 1958), intelligible to all, and one using *sermo altus*, which could be understood only by the literate minority. Taking as an example of the former a page of the *Peregrinatio* and of the latter a page of

the *City of God*, he quite properly points out the differences between these two texts at the levels of organization and discourse, only to draw the conclusion that these differences between the texts mean that they are written in two different languages. By so doing, he is able to agree with those scholars (let us call them ‘traditionalists’) who place the end of Latin and the beginning of Romance as early as the fourth century – if not before.

However, this type of argument takes liberties with the data by failing to give sufficient weight to literary and sociolinguistic factors. To begin with, to be accepted as valid in any particular instance, this heuristic should be tried and tested on different periods of language history. If we apply the same sorts of argument to the spoken Latin of the Classical period by comparing, say, Lucan to more rudimentary texts such as fables, riddles, cookery books or soldiers’ letters (Adams 1977), we shall reach the same conclusion: first-century Latin was no longer a living language (some scholars have drawn the same conclusion on the basis of phonetic evidence; see Bonfante 1999). Likewise, if we compare the *Pensées* of Pascal (1623–62) with a farce from the same period, we shall conclude that the illiterate hay-pitchers of the Port-au-Foin, whom the poet François de Malherbe (1555–1628) famously cited as a target audience, would have had great difficulty understanding Pascal. But does this mean postulating that two languages existed side by side in seventeenth-century Paris? It soon becomes clear that the reasoning of many Romanists leads us badly astray, based as it is on an a priori assumption – that Latin was always split into a learned language and a ‘vulgar’ language, separated by a chasm.

It remains for us to account more fully for the linguistic status of these works of late Latinity, to make an accurate assessment of their communicative status, and to decipher what they can tell us about the development of Latin into Romance. St Augustine’s *City of God* represents the summit of patristic Latin and is by any definition a masterpiece of Latin literature. Its aims and its audience are clearly defined: it is directed at the Christian élite and addresses their concern at the setbacks encountered by the Empire and the attacks of pagan intellectuals who are keen to point out that divine providence had been unable to protect Rome (Brown 2001). St Augustine is dealing with scholarly issues in historical theology, and the difficulty of these topics entails a corresponding complexity of language. By way of comparison: once they take flight, a Kant, a Heidegger or a Sartre can be impossible to understand for anyone unfamiliar with their thought and language, even learned scholars; but we can draw no conclusions at all from this fact about the everyday language spoken by their contemporaries.

Having cleared up this misapprehension, we now turn to the more strictly linguistic issues. To start with, it will help to compare the language of the *City of God* with that of St Augustine's own sermons. There is not a shadow of doubt that when St Augustine preached, he was understood by everyone present (Banniard 1992a:65–104; Dolbeau 1996). I will not rehearse here the evidence and arguments concerning this point, which I have set out at length in earlier work; suffice it to say that this fact is absolutely obvious to anyone who is familiar with Late Antiquity and Patristics, and, indeed, the exceptional personality of St Augustine himself. Any problems of understanding are due not to linguistic incompatibility, but to the complexity of the argument, the intricacies of theology, and possibly some trivial points of style. When St Augustine expresses himself simply, it is not because his audience speaks proto-Romance; it is because their lack of intellectual sophistication calls for patient explanation on his part. A charismatic American evangelist in our own time would doubtless do likewise.

In any case, fluctuations in linguistic register are characteristic of both the sermons and the *City of God*. When he is speaking to the faithful, St Augustine's thoughts may be more or less complex, according to circumstances, whether this complexity is sustained through the whole of a sermon or merely present in parts of it (Banniard 1998a). And his language mirrors these fluctuations, so that some sermons incorporate passages of relatively great syntactic complexity, and are certainly in late Latin of very high quality. But nothing suggests that the audience to which this complex language was directed was incapable of understanding it – at least, not for linguistic reasons. When misunderstandings occur (and there are many examples in the sermons), they never seem to involve the existence of two parallel languages. There are positive *testimonia* to this effect, both direct and indirect; it is unreasonable to sweep them grandly to one side and instead put forward negative hypotheses for which there is no sound documentary evidence.

Moreover, although this type of linguistic fluctuation is less pronounced in the *City of God*, it is not difficult to come up with examples. Alongside long and complex sentences we regularly find passages in a style which is more staccato, more direct, more colloquial, more immediately accessible, and practically identical to certain passages of the sermons: “*Unus panis, unum corpus multi sumus.*” *Qui ergo est in eius corporis unitate, id est in christianorum compage membrorum, cuius corporis sacramentum fideles communicantes de altari sumere consuerunt, ipse uere dicendus est manducare corpus Chisti et bibere sanguinem Christi* (Avg., *Civ. Dei*, 21, 25) [“We, who are many, are one bread and one body.” Therefore, anyone who is in the unity of that body (that is, within the membership of Christianity), the body whose sacrament the faithful

communicants have been accustomed to take from the altar, is truly to be said to eat the body of Christ and drink the blood of Christ']. A great writer like St Augustine, with his training in the arts of rhetoric (Marrou 1958) and his expertise as a communicator, was obviously going to respect the ancient rule of *uariatio*, which enjoined an orator (and the *City of God* is, amongst other things, an extended piece of advocacy) constantly to shift stylistic level. There is nothing to suggest that the passages in question could not be satisfactorily communicated to the mass of believers simply by being read aloud.

Here it is appropriate to return to the language of Ætheria (Löfstedt 1911; Väänänen 1967), and compare it with that of St Augustine's sermons. Naturally, we find differences between them, although these cannot really be ascribed to distinctions of language *stricto sensu*, but rather to distinctions of discourse or style. In any case, it is not obvious why the *Peregrinatio* should be taken as more typical of the linguistic situation in the Western Empire of the fifth century than St Augustine's preaching. Once again, we must bear in mind the work of dialectologists, sociolinguists and cultural historians who work on periods where data can be verified 'on the ground', so to speak. They all agree on stressing the huge disparity between written and spoken language; the large number of strata which make up the continuum of a spoken language and the emergence of structures which are unexpected but which can be explained after the event in terms of spontaneous speech have been the subject of countless studies. There really is no reason to suppose that Latin-speaking Roman Africa in the fourth and fifth centuries was any different (Blanche-Benveniste 1997; Labov 1972b). To claim that differences of style and of communication amount to different languages is to be guilty of a category error, and suggests that the linguistic situation of this period had an exceptional status, an assumption which is wholly unjustified.

Indeed, at the beginning of the fifth century, Latin presents a communicative continuum which is not very different from what one would find in other sociolinguistic contexts, ancient and modern. Rather than taking contrasts which stem from rules of rhetoric, use of communicative devices, and fluctuation in ordinary speech, and elevating them to the status of distinctions between different languages, we should simply accept that there was a linguistic continuum corresponding to the cultural one, and that the two were intimately linked.

2.2.2. A genuine example of linguistic insecurity in the eighth century

With the passage of time and the often drastic external changes in communicative conditions which it brings, come internal changes to the *testimonia*.

These are amenable to interpretation, so long as the study of sources is undertaken with sufficient care, eschewing hasty and Procrustean readings (Banniard 1992a:50; 1992c). In the eighth century, a modest author on the borders of Berry and Touraine (modern Saint-Cyran-du-Jambot) sets out to rewrite a *Life of Saint Sigiramus*, an abbot who had flourished in the previous century (edited by B. Krusch, in *MGH, SRM*, t. 4, p. 603 sqq. *Vita Sigiramni abbatis Longoretensis*). The monk who was given this task is not in the orbit of the Carolingian Renaissance; he was no doubt working in the first half of the century, in a language and culture which were still Merovingian. An analysis of his language shows clearly that it was uninfluenced by the *reformatio in melius*, not even in the basic matter of spelling, whose ‘improvement’ began under Pepin (Pei 1932).

Nonetheless, our author begins with the *topoi* of the time. Given his intellectual limitations, he is going to write ‘not in a lofty style [*non quidem sermone coturno*], but rather ‘in a simple one [*simpliciter cupio proloqui*], hoping that the saint will grant him ‘free-flowing eloquence’ [*largifluam facundiam*]. St Jerome’s celebrated description of the style of Hilary of Poitiers here puts in a reappearance, so to speak, a subtle hint to those who have commissioned the text that the writer is not without culture. Next comes a very personal statement about a specific communicative situation: the author seeks to justify his rewriting of the *Life* by explaining that the monks of Saint Sigiramus have requested him to do so. There was already an older version of the *Life*; but, explains the author, ‘its whole syntax is difficult to understand [*omnis constructio eius ad intelligendum confusa*], because ‘neither words nor syllables are in keeping with grammatical tradition [*prout se habet auctoritas litterarum, tam uerbis quam sillabis*]. The author has the original manuscript to hand as he writes; it is in a deplorable condition, he says, for which he blames both the copyist and the reviser, who have produced a defective edition, in which the text is ‘so deformed as to be absurd [*haec . . . nimis esset absurda ualdeque deprauata*].’

Despite this harsh criticism of the form of the earlier *Life*, the author is more circumspect when it comes to its substance. He warrants that he has respected tradition (including oral tradition) (*non quod me aliquid quod foret inopinatum referre, uel quod rationabiliter insitum erat minuere decreuissem . . . nullatenus me in hoc opusculo temerarium deputet*), and justifies any apparent liberties by reference to no less a precursor than St Jerome, who, with his usual asperity, argued that it was pointless to produce an impeccable manuscript if it was going to be vitiated when copied or read aloud (*ut beatus ait Hieronymus, nil profuit emendare uel corrigere, nisi in conscribendo aut recitando, quod rationabiliter*

insitum est, omnimodis conseruetur). Finally, he underplays his claims by using litotes to define his role: he aims to have produced a version which is ‘not wholly without refinement [*efficaciter, prout potui, non impolitam reliqui*]’.

Since the manuscript used by the editor dates from the eleventh century, it is difficult to make any judgement concerning the grammatical competence of its author. The text contains a handful of insignificant Merovingianisms which in no way detract from its clarity. As the original *Life* has not come down to us, we cannot form an impression of how bad a state it was really in or the extent of the corrections which have been made to it. But, even so, this document is an important source for us, if only because it constitutes a sort of ‘missing link’ between Gregory of Tours and Alcuin. The former deplores his limited knowledge of grammar (even though this is a conscious affectation, it does not invalidate his *testimonium*) (Banniard 2001b; Beumann 1964), but earnestly entreats his intellectual successors to treat the manuscript text of his *Libri Historiarum* with scrupulous respect and not to alter it in the slightest way (Greg. Tur., *HL*, 10, 18, p. 536). The latter takes exactly the opposite point of view: he has no hesitation in rewriting the Merovingian *Lives* in order to correct their language; and his *emendatio* is at times so far-reaching as to affect the substance of the text itself. Whilst the second biographer of Sigiramus proclaims his scrupulous respect for the original, Alcuin will completely rework the narrative to make it more noble (Banniard 1992a:378; 1993b). This *testimonium* is part of a lengthy history in which textual *topoi*, however laden with meaning, are not allowed to stand in the way of significant evolution. The author of the new *Life of Sigiramus* corrects the language of the original *Life*, thereby taking a different position from that of Gregory, but is much more prudent than Alcuin in avoiding any changes to the content.

He clearly cannot have plucked this middle way out of thin air; his preface is therefore of great value to us and validates the sociolinguistic method of enquiry. All this said, does this text tell us anything more about vertical communication in eighth-century Merovingian Gaul? It can, if we subject it to a close reading and make comparisons with the two authors just mentioned, Gregory and Alcuin. First of all, the author criticizes the incorrect language of the original manuscript not because it offends against the rules of grammar, but because it is communicatively inefficient. It is no coincidence that he cites St Jerome and refers to the text being read aloud (*recitando*). Reading the original *Life* aloud seems to have become so much of a problem that vertical communication was compromised. We cannot establish whether the work was intended to be read *intra muros* for the benefit of the monastic

community or whether it was also meant to be declaimed *in ecclesia* for the benefit of the faithful as a whole – although, in a sense, this consideration is unimportant, as some of the monks would themselves have been illiterate. In any case the whole historical context of the Merovingian *Lives* argues in favour of the second hypothesis. Be that as it may, whether the target of communication was the restricted audience of monks or the more general audience of lay people, the author's statements contradict those made two centuries earlier by Gregory, who maintained that 'incorrect' written Latin made vertical communication easier. The proximity of Saint-Cyran to Tours rules out any diatopic interpretation of these discrepant views. Rather, they should be construed diachronically: between the sixth and eighth centuries the linguistic situation has changed to the extent that the members of the speech community now feel linguistically insecure. Ordinary spoken language can no longer readily compensate for the pitfalls engendered by the defective Latin of the copy of a *Life* intended for vertical communication.

3. Chronology of communication and chronology of speech

3.1. *Stages in the functioning of vertical communication*

Now that we are confident of our answers to these preliminary methodological questions, we can begin to examine the results obtained by this method when applied to the transition from Latin to the Romance languages. We shall return to our initial sociolinguistic question, but frame it in a slightly different way: How long was Latin able to serve as a language of general communication? However, we need to consider other issues, as well. If we can actually come up with a concrete answer to the question, we shall be able to put more faith in it if it can handle nuances on two levels of analysis: synchrony (fluctuations linked to contextual factors such as the difficulty of the topic, the prowess of the orator or the reader, linguistic variation) and diachrony (careful overall consideration of how the passage of time may disrupt communication). Neither synchrony nor diachrony consists simply of binary oppositions: any reasonably complex communicative exchange involves some loss of information, yet the existence of holes in the fabric of communication does not necessarily imply the existence of heterogeneous linguistic systems; likewise, no significant diachronic development can take place without showing symptoms beforehand – cracks in the communicative structure, as it were – whose extent in space and time will enable us to tell that a major sociolinguistic shift is on its way.

3.I.I. The vitality of vertical communication in the third and fourth centuries

Thirty years' worth of work informed by the principles of historical sociolinguistics has enabled us to produce a genuine history of vertical communication between Latin speakers in the West. According to the evidence available from every region, during the fourth and fifth centuries vertical communication had a vitality of function comparable to what we would expect to find in any modern-day language (Banniard 1992a: ch. 1; 2001b, 2000c; Olivar 1991; van Uytanghe 1976; 1994; 2000; Wright 1982).

Nonetheless, as is to be expected, communication is more or less effective according to various considerations which we may group together under three headings.

- (1) *Considerations having to do with external factors operating in Late Antiquity: the society, culture, and mindset of the period.* These factors can be negative (huge social inequality; high percentage of illiteracy; gulf between urban and rural areas) or positive (the unifying effect of the institutions of Empire; continuous road and sea traffic; a school system which is efficient, in spite of its limitations; the prestige of written monuments) (Carrié and Rousselle 1999).
- (2) *Internal considerations having to do with the spread of Christianity.* The Christian faith is based on a written tradition, and its wholesale adoption in the West took place through the medium of Latin. By 400, the fusion between the old pagan linguistic and cultural traditions and the more recent Christian ones had made considerable progress. In order to have carried all before it, the new religion needed to have been conveyed by intense face-to-face communication. Moreover, this was not simply a one-way top-down process; the Church assessed the effectiveness of its communication by reference to how accurately and how well its message was received and understood. Even though conversion to Christianity can sometimes be seen as an intrusion into people's lives, and sometimes a violent intrusion at that, the process is always subject to the imperatives of communicative clarity. It is impossible to reproduce here the mass of documents which demonstrate that every stratum of society was involved in and affected by this process. Whether people approved or disapproved, whether they welcomed or feared its coming, the expansion of the new religion was the result of such intense communicative activity that the limitation we mentioned earlier – namely that 90 percent of speakers are inaccessible to study – ceases to apply.

Christian communication is both energetic and invasive (Brown 1992; Cameron 1991; Piétri *et al.* 1995).

- (3) *Internal considerations having to do with linguistic choices themselves.* The arrival on the scene of Christianity, and its adoption by high-status intellectuals who, despite being converts to the new values, were imbued with the Latin tradition, gives us new insights into the conditions of general communication in Latin. The rhetoricians of classical antiquity did address the question of how communicative needs might constrain the linguistic ingenuity of a stylistically assured orator; but it is always difficult to know for certain what audience they had in mind, and our knowledge of Roman society in the Republic and early Empire makes us sceptical that they were targeting all speakers, regardless of their social, cultural or linguistic level. Moreover, attempts to moderate one's style during this period are always limited by the criterion of *elegantia* and by the canons of elite good taste (Banniard 1988). Both these limitations are called radically into question by pastoral Christianity. In their place emerge two new precepts. First, simplicity: the simplicity of the original Apostles included simplicity of language; they had humble occupations and their speech smacked of the 'language of fishermen [*sermo piscatorius*]'. Second, intelligibility: the overriding imperative of communicating the Good News to every single individual meant that the first task of a proselytizer was to use intelligible language. This is not to say that there could be no overlap between the new principles and ancient rhetoric: they could concur, for instance, when it came to expressing the sublime. The fact that Christ's message was sublime, even though it was expressed in a humble style, was a novel idea that literate speakers who were imbued with learned culture nevertheless found easy to accept. The theory and practice of vertical communication reach their apogee in St Augustine. His linguistic choices can be systematized to yield a sort of portrait of the type of Latin which he deemed to be effective, notably the linguistic register which he famously promoted, the *sermo humilis*. This represents a purely stylistic register of what late spoken Latin had become by the fifth century. The linguistic options are clear: the fundamental consideration is the ordinary spoken language of Late Antiquity, and we have no evidence at all which would justify speaking of a Latin/Romance dualism (Auerbach 1958; Borst 1957; 1958; Lentner 1963; Norden 1898).

St Augustine's comments on problems of language itself, as opposed to problems of style or subject matter, have often been over-interpreted. Time and again, commentators have shown little concern for his work as a

whole and scant interest in its cultural context, yet have been keen to focus on the handful of references which are relevant to the evolution of Latin (Banniard 1995a; 2000a). His comments often concern phonological problems, and it is inappropriate to extrapolate too much from them; in any event, their relevance to the flow of dialogue and communication between speakers is rather limited: it is unlikely, for example, that the reduction or shortening of all occurrences of final unstressed [a] proved a real impediment to communication. Homonymy was increased as vowel quantity ceased to be phonological and became simply a matter of conditioning context, but it was easy to overcome or sidestep any problems which this gave rise to. Take present/preterite distinctions which rely on vowel quantity, such as present *uēnit*, with short [e] vs. preterite *uēnit*, with long [e:]. Ways round the problem emerge: first, and most straightforward, is the restructuring of the phonological opposition of quantity as one of quality; the long vowel becomes close and the short vowel open, so that the opposition between present and preterite comes to be conveyed by a distinction of vowel height. Sometimes the distinction is hypercharacterized, so that the /e/ of the preterite is subject to regular phonological raising to /i/ (*uēni* > **vini*, for instance), which has the effect of creating an opposition with a higher functional yield. And, in any case, every language contains homonyms and ambiguity. The fact that these variations exist does not mean that speakers have said goodbye to Latin. Indeed, St Augustine frequently notes that the ordinary language of Roman Africa has changed or is changing; but he attributes these changes to the remarkable vitality of Latin itself. This position emerges so consistently that one wonders why so many scholars misread this evidence as indicating change for the worse. And why believe St Augustine when he draws attention to discontinuity whilst according him no credence at all when, as most of the time, he stresses continuity?

St Augustine was a Roman from Africa who displayed great learning and prodigious linguistic gifts, but he was unafraid to exploit language in all its complexity, including clashes of register and asperities of style (Mandouze 1968). He created a hitherto unknown literary genre, articulated in completely original Latin prose, but prose in which any reader familiar with the period will sense the cadences of late Latinity (Fontaine 1987). Nor was he reticent about modelling the phrasing of his sermons on the everyday Latin of Africa, going so far as to stud it with expressions and constructions which Romanists would pounce on delightedly, were it not for the fact that they are simply fine examples of late Latin idiom (Banniard 1998b).

3.1.2. Robustness of vertical communication
during the sixth century

It is increasingly common for historians to view the sixth century as an extension of Late Antiquity. Clearly, there are external changes during this period, but they take place slowly enough for the dynamic of the Empire not to be suddenly swept away. Vertical communication continues to function robustly during this period, even though certain subtle alterations are beginning to make themselves felt.

Let us return, *mutatis mutandis*, to the three headings which served as a framework for our earlier discussion.

- (1) *Considerations having to do with the shift from an extension of Late Antiquity to the beginnings of the High Middle Ages.* The unifying effect of the bonds created by imperial institutions is now considerably reduced (although we should not forget the impact of a certain number of reconquests by the Eastern Roman Empire); however, by this stage, the Church's institutional network is dense and effective enough to replace, at least in part, the weakened focus on an imperial centre. As far as changes in social relationships are concerned, we should note a certain decline in the number of *litterati*; however, the immense majority of the Empire's inhabitants had always been illiterate, and even an appreciable reduction in the number of those who could read and write would not have had a significant effect on society as a whole. Moreover, the link that has been made between the 'decadence' of Latin and the demise of the school system (which likewise only ever involved a minority) is not the result of linguistic analysis, but rests on cultural prejudice. The evolution of ordinary language has very little to do with the school system; Latin was spoken before schools (or the Empire) existed and continued to be spoken after they had disappeared. In the very long term, the effect of the education system on spontaneous speech is infinitesimal, if not non-existent. It is true that the decline in educational standards in some Germanic kingdoms led to a decline in people's ability to handle the written language; but that is not the same thing as a change in speech habits (Banniard 1989; Riché 1973; Wood 1990).
- (2) *Considerations having to do with the spread of Christianity.* This period is characterized by three phenomena: the promotion of the old Roman élites in the new *cursus honorum* of the Church; the increasing breadth and depth of Christianity in the countryside (the development of networks of parishes as more land was settled); and the wholesale spread of monasticism, with the founding of thousands of monasteries, a development which was

both reinforced and challenged by exotic forms of Christianity, such as hermitism. During this same period, the cult of saints flourishes, to the extent that any settlement of any size acquires a 'patron saint', with his or her own place of worship (where their relics are kept), a cemetery (where burials *ad sanctos* take place), and moreover with a proliferation of hagiographic material (*Lives* and the like). In these trends we can discern the social influence of the *litterati* – even though their culture was often inferior to that of their imperial counterparts. Such developments have provided modern scholars with ample material for study (Atsma 1989; Heinzelmann 1976; Périn 1992; Piétri 1998; Riché 1993).

- (3) *Considerations having to do with linguistic choices.* The appearance and diffusion of the *Lives of Saints* create a denser communicative network in the Latin West and at the same time indicate that vertical communication was working well. The works in question, which were written and recopied so that they could be read aloud on the saint's annual feast day, are, together with sermons, the most reliable yardstick of what was intelligible to illiterate speakers. As in previous centuries, intelligibility depends not on choice of language, but on choice of style. The vocabulary used by the literate to describe their linguistic preferences simultaneously maintains, diversifies and consolidates the principles which had held sway under the Empire. There is little point in trying to find precise labels in order to translate their terminology into diastatic or diatopic terms. In the sixth century, St Caesarius of Arles distinguishes between a literate élite with refined taste (*eruditi, eruditae aures*) and the community of illiterate believers, who have simpler desires and intellectual capacities (*imperiti, simplices*). In order to be received and understood, the Christian message must come down from the rhetorical heights (*scholasticorum altitudinem*) and descend to the level of the speech that walks the countryside (*pedestri sermone*), which in turn requires preachers to speak like the illiterate (*rustica uerba*), in order to make up for the intellectual limitations (*ignorantia*) of their audience. Clearly, there is an element of provocative affectation in describing simple but clear and elegant language used by a literate speaker as the speech of the illiterate. But, no matter how hard we look, we shall find no reference to anything other than a difference of register; we are merely dealing with a normal situation of mass communication. Moreover, although we have ample evidence that the faithful remained obstinately deaf to the exhortations of their priests, there is nothing to indicate that their resistance to these admonitions has anything to do with their failure to understand the message (Banniard 1996b).

Very similar conclusions can be drawn from the work of Gregory of Tours half a century later; but in this case our information comes not from his sermons, which have not come down to us, but from his hagiographic work. Although he lacks St Augustine's range, Gregory shares his creative boldness (the language of his *Libri historiarum* is at least as innovatory as that of St Augustine's *Confessions*) and his concern for communication. He evinces a sophisticated awareness of linguistic registers, from the most refined (as witness his delight in reading Sidonius Apollinarius) to the most simple. His comments on his lack of grammatical education may be dismissed as affected self-deprecation, but they nonetheless demonstrate that his mind and his observational abilities are sufficiently keen for us to trust his judgement when it comes to matters sociolinguistic. He is constantly aware of diastatic variation in speech, but his conclusions about the effectiveness of vertical communication confirm that a linguistic continuum is being maintained, at one extreme of which we find a learned style of Latin (*philosophantem rhetoricem*), brimming with oratorical effects (structure of periods, archaisms), and at the other an illiterate style of Latin (*loquentem rusticum*), which continues the *sermo humilis* (short sentences, repetition, everyday vocabulary) (Banniard 1995a; Herman 1999; Norberg 1966).

These conclusions are valid for all the Latin areas of the ancient Empire, as shown by studies of specific localities (for instance, in Italy). The results of more significant studies of the following century are even more conclusive, and point to the fact that, in this century, too, the rest of the Latin area behaved in a similar way to Gaul.

3.1.3. The continuity of vertical communication in the seventh century

In the seventh century, vertical communication continues more or less unchanged in both northern and southern Gaul. Although there are no great literary texts dating from this period, the *Lives* (van Uytendaele 2001) contain numerous pointers to its continued functioning (Banniard 1992a:253–303; 1992b; 2006b; Bayer 2007; van Uytendaele 1987; 1989; 1994). The vocabulary used by the (often anonymous) authors when describing their language is rich and varied enough to enable sociolinguistic conclusions to be drawn. In this connection, we should mention particularly the preface to the second version of the *Vita Leudegarii*. Its precision, its originality and its elegance all make this work a worthy successor to the self-analytical reflections of Gregory of Tours,

and it marks the end, around the year 700, of a period of relatively stable continuity, before the problems which were to surface in the eighth century.

A fortiori, vertical communication remained robust in Italy under the Ostrogoths, then under the Empire once again, and finally under the Lombards, as indicated by all *testimonia*. And now we once again have a major author to assist us with our enquiries, in the form of Pope Gregory I, who is just as concerned with pastoral communication as were St Augustine and St Caesarius of Arles, and whose work enables us to give a clear description of the vertical communication of the time and even to make out the types of audience and the different levels of effectiveness, in a way which confirms the credibility of his evidence (Banniard 1992a:105–79; 2005c; van Uytanghe 2005). We may distinguish:

- (a) high-register literary Latin, as represented by the *Moralia in Job*, whose target audience is an élite of clerics and monks. This variety is equivalent in difficulty to the language of the *City of God* (but does not constitute an argument for Latin *per se* being unintelligible);
- (b) Latin which is difficult but accessible to a selected lay audience, as represented by the *Homilies on Ezekiel*;
- (c) popular Latin, which is accessible to all *illiterati*, as represented by the *Homilies on the Gospels* and the *Dialogues*.

In Visigothic Spain, we likewise have a source of the first order for this period, in the person of Isidore of Seville (Banniard 1975; 1992a:181–251; Fontaine 1983: vol. 3; Velázquez 2003; Wright 1982; 1994; 1995a). In addition to his roles as bishop and royal counsellor, Isidore was deeply concerned with pastoral issues, and was prodigal with his advice to preachers. He, too, distinguishes at least two types of speech:

- (a) speech for restricted communication, *scholasticus sermo*;
- (b) speech for general communication, *apertus sermo*.

There are no grounds for interpreting anything in Isidore's work as a reference, explicit or implicit, to a Latin/Romance dichotomy.

He divides his audience into three categories:

- (a) literate;
- (b) semi-literate;
- (c) illiterate.

He is fully aware that late spoken Latin (which he calls *lingua mixta*) is significantly different from the 'ideal' Latin of the past, but still recommends finding a middle way between basic expression which is too humble on the one

hand, and excessive elegance on the other, without shying away from using simplified language and style (*sermone plebeio uel rustico*) where appropriate.

In sociolinguistic terms, we may conclude that the communicative continuum remained in place in the Latin West, although a significantly greater effort now had to be made to overcome the increasing gap between different linguistic registers.

3.1.4. The eighth century: continuity and hesitation

The eighth century provides us with a variety of evidence to suggest that vertical communication is beginning to encounter real difficulties. The deep-rooted nature of changes during this period can be illustrated with reference, *mutatis mutandis*, to the headings used above for the fifth and sixth centuries.

- (1) *Considerations linked to a shift in the centres of gravity of civilization.* The dawn of the Carolingian Empire completes the process of shifting the Mediterranean and Latin centre of Late Antiquity to the north-east of our area and its hordes of new Germanic peoples (Angenendt 1999). A similar shift takes place in the south, where south-western Latin encounters a new civilization, and, most pertinently, a new written culture, that of Islam (Collins 1999; Sénac 2002). It would be wrong to take an apocalyptic view, and all too easy to exaggerate the discontinuity involved; but horizontal communication is nonetheless profoundly disturbed as a result of these changes, and this, in turn, favours a centrifugal dynamic in various cultural, religious and linguistic spheres.
- (2) *Considerations having to do with the spread of Christianity.* The spread of Christianity now takes place under very different conditions. On the one hand, it is making conquests in the north-east, thanks to the powerful military and administrative backing of the Carolingians (Piétri 1993). However, in the south-west it finds itself in the novel situation, unknown since the Peace of Constantine in 313, of having to resist conquest by another religion, supported by another army (Picard 2000; Laliena and Sénac 1991). In other words, a qualitative change takes place during the Carolingian period, and adapting one's message to the general public becomes an absolute imperative. The desire to retain the Roman-Visigothic inheritance in an environment which is no longer propitious suddenly creates tension and urgency in Spain and in southern Gaul.
- (3) *Considerations having to do with linguistic theory.* The context in which linguistic choices are made in this period has three components: (a) the inertia and force of habit which are the inheritance of Late Antiquity, now

in its last throes (McKitterick 1994); (b) reaction to contact with new ‘colloquial’ languages, in very contrasting situations: on the one hand, Germanic languages, elevated to the status of apostolic tongues by the simple fact that the people who spoke them had converted to Christianity (Haug 1997; Hannick 1999; Pohl 1999), and, on the other hand, Arabic, which already had enormous prestige, because it came with a complete range of registers, from colloquial dialect to holy scripture (Millet-Gérard 1984); and (c) the presence or absence of a centralizing authority, there being a huge contrast between St Boniface’s authoritarian efforts to establish central power with the support of the papacy and the vacillating subsidiarity of Mozarabic areas, where we occasionally find compromises stemming from individual initiatives, such as that of Elipand of Toledo.

External conditions are therefore beginning to favour profound sociolinguistic change. And modern scholars, with the benefit of hindsight, can project back from the documentary evidence of the ninth century to draw conclusions about the period that immediately preceded it. It is, after all, logical to suppose that between the seventh century, which was characterized by a continuity of vertical communication, and the ninth century, characterized by a discontinuity of vertical communication, a state of ‘prediscontinuity’ emerged during the eighth century. However, we must be prudent when assessing this phenomenon. First of all, *testimonia* often lend themselves to contradictory interpretations. Two errors above all have to be avoided: seeing the undoubted evidence of gaps in the religious training of the clergy as proof of an acute communicative crisis (the major Christian prayers are straightforward in appearance only – orthodox trinitarian monotheism is in itself quite a stumbling-block to communication, and this has nothing to do with language as such), and confusing in more general terms the alleged decadence – or even corruption (although historians are now able to put these notions into perspective) – of the Merovingian Church with linguistic *laissez-aller*. Inasmuch as vertical communication does not break down completely in the ninth century, and does not break down simply because of internal developments, it is logical to see the eighth century as a frontier zone, with the communicative continuum still dense around 700, but becoming much weaker around 800 (Banniard 1992a; 1994; 1998a; Calboli 1992; Herman 1996b; van Uytenghe 1976; 1994).

Any significant regional variation we come across during this period is more likely to be a geographical reflection of chronological developments. For instance, the historical scholar Paul the Deacon describes the settlement of Benevento by a tribe of Bulgarians between 650 and 700, and the granting of a

‘dukedom’ to their leader. Paul, whose competence as both a germanist and a grammarian is well known (he accounts for a number of Germanic anthroponyms and is one of the specialists of the language reform), notes with surprise that these Bulgarians have become bilingual: they have learnt to speak Latin (*quanquam et latine loquantur*) without abandoning their own language (*linguae tamen propriae usum minime amiserunt*) (Paulus Diaconus, *Historia Langobardorum*, 5, 29). His surprise has nothing to do with the Bulgarians’ having learnt Latin; it concerns their continuing knowledge of their original language, more than a century later. Other *testimonia* point to similar conclusions (Banniard 2000c; Wright 2000).

To sum up: the example of the *Life of Saint Sigiramnus*, the appearance for the first time of attempts to parody Latin in somewhat de-Latinized *scripta*, as well as our knowledge in hindsight of subsequent developments, all point to this century being ‘the end of an era’.

3.1.5. Discontinuity in the ninth and tenth centuries

(1) From the ninth century onwards, the history of communication varies according to country. The ninth century has been a particular focus of enquiry, on account of the Carolingian intellectuals (McKitterick 1994; 1999) – amongst other (good) reasons, because they accelerated the reforms of the previous century, thereby creating an exceptional sociolinguistic framework (Banniard 1992a:306–422; Wright 1982; 1991; 1998; van Uytvanghe 1976; 1984). Other areas follow these developments at a certain distance. *Testimonia* now give positive and direct evidence that the continuum of vertical communication, which had hitherto survived more or less intact, is now being torn apart. However, this evidence has not straightforwardly led to contradictory hypotheses being replaced by facts on the ground, as there is room for several different interpretations, even within historical sociolinguistics. In what follows, I shall start with what is most certain and end with what is most controversial.

(a) The meaning of *rustica romana lingua* (Canon 17 of the Council of Tours of 813; see Banniard 1992a:410f; 2008b).² This term is not as new as it might appear. The phrase *romana lingua* is still during this period a

2 We give here the the original Latin text (MGH, *Concilia*, t. 2, *Concilia aevi kar.*, 1 (ed. A. Werminghoff, Hannover, 1908)) of Canon 17 of the Council of Tours of 813, at which it was ruled that the clergy should adapt their sermons to the ‘rusticam Romanam linguam’ for the benefit of their illiterate parishioners: *Visum est unanimitati nostrae, ut quilibet episcopus habeat omilias continentes necessarias ammonitiones, quibus subiecti erudiantur, id est de fide catholica, pro ut capere possint, de perpetua retributione bonorum et aeterna damnatione*

synonym of *latina lingua*, and there are no significant differences between them; *rustica* simply means ‘illiterate’, and has no diatopic connotation. So the phrase as a whole means ‘the Latin of the illiterate’, and this is the only interpretation which is compatible with the perceptions and the vocabulary of the time.

- (b) The appearance of a break in the linguistic continuum (Banniard, 1992a:413–14, 500–5). Those responsible for Canon 17 did not retain the former terminology, *sermo rusticus*, because their mindset is a new one: defining the language of everyday communication not as the **language** of the illiterate but as the **style** of the illiterate would amount to recognition of Merovingian Latin as an acceptable Latin style (albeit at the most humble level). So otherness of style is elevated to otherness of language.
- (c) The meaning of *transfere* (Banniard 1992a:411–13; contra, Wright, 1982:121). This is a topic of great debate. Are we talking about a genuine translation (as, say, from English into French) or a simple transposition (whose nature remains to be determined)? The construction’s strict parallelism with the *thiotisca lingua* (‘the language of the people’; Thomas 1988; 1990) is a strong argument in favour of the meaning ‘to translate’. This meaning would reflect the fact that the Merovingian *sermo rusticus* was now no longer seen as Latin (see above). One cannot help thinking that the Carolingian intellectuals are being devious again: by equating the strangeness of ‘the Latin of the illiterate’ and ‘the language of the [Germanic] people’, they are once again rejecting the Merovingian period, or even sneering at it (Geary 1988), and banishing Merovingian Latin and Latinity.
- (d) The meaning of the words *facilius* (lit. ‘more easily’) and *aperte* (lit. ‘openly’). These adverbs reveal that that the situation had become critical without degenerating totally: the reformed Latin of Carolingian sermons and the *Lives* ruptured the communicative continuum without destroying it entirely.
- (e) The actual nature of the target language for this ‘translation’ or ‘transposition’. Here, too, I shall adopt a sociolinguistic perspective. In the light of the thinking and vocabulary of the Carolingian intellectuals, the canon allows spoken Latin of the Late Merovingian period, as it existed

malorum, de resurrectione quoque futura et ultimo iudicio et quibus operibus possit promereri beata uita, quibusque excludi.

Et ut easdem omelias quisque aperte transferre studeat in rusticam romanam linguam aut thiotiscam, quo facilius cuncti possint intelligere quae dicuntur.

prior to the Carolingians' abortive attempt to reform everyday communication, to be readopted (by those who had followed the directive) or retained without implying disobedience or secrecy (by those who had resisted the directive, out of principle or incompetence) (Banniard 1998). Canon 17 of the year 813 cannot really therefore be described as 'the birth certificate of the Romance languages'. The fact that, from our modern perspective, spoken Merovingian Latin of the eighth century merits the name of 'proto-French' should not obscure the fact that, even during the crisis of the Carolingian era, the rift between the written and the spoken language is not seen in terms of immediate and total incompatibility (Banniard 1992a:497–505; 2008b; McKitterick 1989; 1991; van Uytendaele 1977; Wright 1982; 1991; 1997).

- (f) The nature of the Carolingian reform of Latinity. This is an important element of the debate, which I shall leave open in this discussion. In overall terms, the *reformatio in melius* of written and spoken Latin in the second half of the eighth century consisted of identifying and reimposing the norms of the patristic Latin of Late Antiquity with reference to a battery of specially rewritten grammars relying on sources which were deemed sound (Law 1982). Was this a reform of pronunciation (Lüdtke 1964; 1993b; Wright 1982; 1991; 1996)? This reform clearly took place in the monastic *scriptoria*. However, its impact on public communication is less easy to assess (Banniard 1985), even though it is beyond doubt that oral communication with the mass of the faithful underwent significant changes. Whatever the precise phonetic consequences of the reform and its extent, the re-establishment of a normative grammar brought substantial changes in its wake, particularly as regards morphology. All the inflections which the Merovingian spoken language had treated logographically (Wright 1982) were now treated as phonographic. Thus, for instance, inflections which had been written but not pronounced suddenly reappeared in speech, causing massive interference in utterances to which illiterate speakers had become accustomed; and where previously the ending *-ibus* (used freely in the plural for the accusative case as well as for the standard dative and ablative) had been realized unproblematically as [es]/[os], it was now given a spelling pronunciation, to spare the ears of immigrant grammarians who were appalled to hear texts where 'solecisms resound [*perstreperere soloecismos*]'. But, by the same token, the illiterate speech community lost its bearings in the face of such audacity (some might say foolhardiness) on the part of the religious authorities (Banniard 1989: ch. 6).

- (2) Does Spain follow a different communicative trajectory? In a context which is clearly very different, and which is conducive to contrasting developments in different regions, all research points to Spain having experienced a certain time lag with respect to Gaul. There seems to have been something of a crisis in Córdoba in the mid ninth century, when the religious extremism of some Christians found expression in a corresponding linguistic extremism, leading to a sort of revanchist reaffirmation of complex Latinity. This move must have placed vertical communication under enormous strain, although it is possible that the Cordoban *energumens* were restricted to a small group of initiates (Banniard 1992a; Wright 1994). But in general the situation meant that Merovingian-style compromises were kept for longer in Spain – although no doubt to a lesser extent in *Al Andalus*, which was rapidly being Arabized, than in the Christian kingdoms of the north (Menéndez-Pidal 1964; Pérez González 1993; Wright 1995b; 1996). Here, motivated by the desire to maintain the continuity of Christianity in the face of their burgeoning monolingual rival to the south (although, as noted above, Arabic was not monolithic, and had a range of registers), the literate perhaps avoided emphasizing the gap between traditional writing and the innovations of speech (Collins 1990). Nonetheless, the glosses and the handful of translations/transpositions which emerge from the tenth century onwards incline one to see the ninth century as a watershed for vertical communication.
- (3) Does Italy lag far behind these other two regions as far as vertical communication is concerned? The kind of detailed sociolinguistic investigation that has been carried out on the ninth and tenth centuries in Carolingian France and Mozarabic Spain has not so far been pursued to the same degree for the comparable period in Italy. But it does not seem to be the case that there was a crisis in vertical communication comparable to that which affected *Francia* globally and Spain sporadically. We also know that the first glosses in Italian *scripta* and the first translations go no further back than the mid tenth century. On the other hand, there are numerous texts in LSLat2³ dating from this period, but the linguistic awareness of their authors reveals no radical divergence between writing and speech, nor a fortiori between Latin and proto-Italian. It transpires that Italy, as so often, is an area of compromise, with the accommodations put in place in the High Middle Ages continuing to survive, a survival which is facilitated by political fragmentation. The names given to the language do not point to any real change until the tenth century. The everyday spoken language

3 Late spoken Latin Phase 2 (sixth and seventh centuries).

(*vulgaris sermo*) is distinguished from the careful spoken language (*latina uox/latinitas*) in two independent *testimonia*, in contexts which make it clear that we are dealing with spoken communication, and, in one case, with vertical communication. The conclusion to be drawn is that, for literate speakers in the tenth century, the common spoken language is incorrect Latin and that vertical communication is effected through translation/transposition from one linguistic register into another. External knowledge leads us to believe that there was an awareness of bilingualism, even though the gap between the languages will not have seemed enormous to people at the time (Banniard 1992a:543–50; Norberg 1999:159).

3.2. *Stages in the relationship between writing and speech*

3.2.1. Evaluative criteria

In this history of the transition, the problems are layered. There is the linguistic question of the status of the language people speak at a given time and in a given place; the sociolinguistic question of the name people give to the language they speak and the ways in which they use it (Janson 1991; van Uytanghe 1991); and the cultural question of how this spoken language relates to the corresponding written language. Let us try to shed some light on this last question by turning it round: How good a fit is there between the traditional written code and the spontaneous oral code?

This question should be hedged with some methodological caution. First, we must avoid two extreme positions: a Manichaean approach, which we rejected at the beginning of this chapter, claiming that the difference between the two codes is absolute; and another, according to which a period of perfect fit (spoken Classical Latin; Lüdtke 1993a) is contrasted with a period of complete lack of fit (proto-Romance). This second position, too, is easily dealt with. The fit between written and spoken codes is always relative, as a simple comparison of modern-day Romance and Germanic languages will show. As is well known, the gap between (standardized) writing and (spontaneous) speech is much greater in the case of French than in that of Spanish, and likewise much greater in the case of English than in that of German. However, this perspective is of limited value if we restrict ourselves to the relationship between graphemes and phonemes. True, French and English have writing systems that are fairly remote from the articulatory realization of their phonology. But if we broaden our discussion to include other descriptive categories, the picture changes, because differences between individual languages in this respect are now effectively ironed out. The goodness of fit between writing and speech can no longer be measured so

precisely once we begin to look at morphology, syntax, the lexicon, idiomatic expression and phrasing. Variation and difference at these levels depend much less on strictly linguistic factors, and much more on cultural considerations, such as literary genre and communicative context.

Overall, if we turn the argument on its head, it is clear that the only possibility that all languages systematically reject is direct phonographic representation. If we take every other level of grammar into account, then all registers of a language have equally effective access to writing. Yet there is one element which writing has always had problems representing – accentuation, and in particular intonation. With the development within acoustic phonetics of the new and complex sub-discipline of intonology (Rossi 1999), we can now measure – literally – the information we lose if we have access to a language only in its written form. In this sense, we shall never have full access to any language prior to the twentieth century, regardless of the age of our written sources.

3.2.2. Outline

These observations are not designed to reintroduce the radical approach which we criticized earlier, but simply to put the emergence of the Romance *scripta* in its proper context: it was not a sudden and dramatic development, but grew, organically, as it were, out of a process of adaptation (Sabatini 1968; Herman 1992; Kramer 1998; Wright 1991). In what follows, I shall sketch the development of the relationship between writing and speech from late spoken Latin to proto-French, linking the various stages to the stages proposed for the development of speech and communication (Banniard 2001a).

Development of the relationship between speech and writing between CSLat. and ProtoFr.

Language	Linguistic situation	Fit	Tension
1] CSLat.	Monolingualism	Good	Low
2] LSLat ₁	Complex monolingualism	{ Reasonable Poor	High Very high
3] LSLat ₂			
4] ProtoFr.	Diglossia	Minimal	Maximal

NOTES:

LSLat. = late spoken Latin

LSLat₁ = late spoken Latin Phase 1: third to fifth centuries

LSLat₂ = late spoken Latin Phase 2: sixth and seventh centuries

The term ‘fit’ here refers to the overall goodness of fit between the written language (Latin, or Latinized) and the natural spoken language – speakers’ spontaneous mother tongue. The term ‘tension’ refers to the effort required to convert spontaneous speech into written text or vice versa. These terms are, of course, impressionistic, in that neither phenomenon has been computed mathematically (it is interesting to speculate on whether such calculations will ever be possible). Nor does the proposed classification take account of the different levels of language that we have previously alluded to.

The extrapolation of this model to the other major Romance languages may require the later periods to be extended somewhat. In particular, it seems that the final stage (stage 4) needs to be at least a century longer than in the case of the *oïl/oc* area. The table should therefore be modified as follows:

Christian Spain

Language	Linguistic situation	Fit	Tension
4] ProtoCst.	Diglossia	Poor	Very high
5] ArchCst.	Diglossia	Minimal	Maximal

The abbreviation ‘ProtoCst.’ is here used to denote proto-Castilian (eighth century) arising from contact with Visigothic LSLatz and ‘ArchCst.’ as a convenient label for archaic Castilian (ninth and tenth centuries).

Carolingian and Ottonian Italy

Language	Linguistic situation	Fit	Tension
4] ProtoIt.	(Partial) Diglossia	Poor	Very high
5] ArchIt.	Diglossia	Minimal	Maximal

The abbreviation ‘ProtoIt.’ is here used to denote proto-Italian (eighth century) arising from contact with Lombard LSLatz, and ‘ArchIt.’ as a convenient label for archaic Italian (ninth and tenth centuries).

If goodness of fit between graphemes and phonemes were taken as the sole criterion, then there would be significantly greater variation between regions, because changes in this relationship were more differential.

The end of stage [4] (diglossia, minimal fit, maximal tension) was essentially determined by external (that is, in the final analysis, cultural) factors. There was in principle nothing to stop the guardians of written Latin from retaining an archaizing orthography, within which innovatory forms could have been included; external factors needed to come into play before the change to Romance *scripta* could take place (Wright 1997). In the north-west of our area, this qualitative change came about through contact with Germanic language and culture in Austrasia (Pitz 2000). The Carolingian language reform was at the heart of two complementary and apparently contradictory developments. It put an end to the two-way traffic between speech and writing which had continued to mark the Merovingian era, whilst also introducing a new written entity into Germanic-speaking society with the creation of a *scripta* for 'the language of the people [*theotisca lingua*]' (Banniard 1991c; 2003a; Edwards 1994; Fourquet 1980b; Haug 1997; Jolivet and Mossé 1941; Mossé 1945). As a result of the second innovation, the paralysis resulting from the first was rapidly overcome. Germanic varieties now had access to a *littera* derived from an age-old Latin tradition (Haubrichs 1995), and thus the notion arose that these varieties had their own grammar; and the Romance varieties were henceforth free to follow suit. 'Illiterate Latin [*romana lingua rustica*]' began to impinge on intellectual consciousness, in the same way as 'the language of the people [*theotisca lingua*]', and, in the same way, benefited from the grammatical reflection which was necessary for it to become a written language in its own right (Haubrichs and Pfister 1989).

4. Creating a linguistic model

4.1. Chronological foundations

On the basis of the foregoing, we have established a historically informed approach to the transitional phase between Latin and Romance with reference to at least three criteria:

- (a) Vertical communication in Latin: this was still flourishing in the fifth century, but had become unstable by the eighth century, and then in the ninth century underwent a crisis, the gravity of which depended on the area involved;
- (b) Nomenclature: the traditional terminology used to designate the different registers of a single language began to break down in the ninth century, indicating that new entities were emerging;
- (c) Access to the written register: the Latin written code is the only model for writing until the end of the eighth century, at which point we find the

beginnings of a new written code which distances itself from traditional orthography.

Taken together, these factors identify a critical period which, for the sake of argument, we can define as the 100 years from 750 to 850 (Banniard 1992a:485–505; Herman 1967:114f.; 1996b; 1998b; van Uytvanghe 1976; 1994; 2000; Wright 1982; 1997). In sociolinguistic terms, this period represents the end-point of Latin as a living language, or, in more modern terminology, the final transitional phase during which the speech community relinquishes the use of Latin in vertical communication. If we now interpret the tensions which emerge before this period with the benefit of hindsight, we can define a longer period of transition, consisting of the beginning of (Christian) vertical communication in Latin (third and fourth centuries), its peak (fifth and sixth centuries), its continuation (seventh century), its disintegration (eighth century), and its abandonment (ninth and tenth centuries).

4.2. *Methodological bases*

Work in historical linguistics which is informed by this sociolinguistic chronology will clearly have to confront a series of problems, and, as a result, will call into question some of the descriptions and datings of traditional Romance philology. In general, the chronology outlined above has been well received by European and American scholars of Late Antiquity and the High Middle Ages, since it corresponds to new thinking in this area and hence to their own view of the periods in question (Banniard 2000a). The real debates are taking place amongst linguists, Romanists, Latinists and ‘generalists’, each of whom brings the methodology of his or her own discipline to bear on these problems, often attempting to maintain the traditional dating at all costs (Sanga and Baggio 1994; 1995; Zamboni 1998; 1999). Rather than discuss all these views in detail, which would in any case be out of place here, I shall propose a linguistic model informed by the new chronology of communication.

4.2.1. Rejected models⁴

- (a) The dualist model. The Manichaean opposition between ‘Literary Latin’ and ‘Vulgar Latin’ which was elevated to the status of scientific principle

4 Banniard (1993a; 1995b; 1998a; 2001b; 2002a; 2008a).

in the nineteenth century is merely the pseudo-scientific projection on to language of cultural and aesthetic differences. It invents a linguistic dualism which would make Latin a unique exception amongst languages (Flobert 1999; Koch 1995; Oesterreicher 1995).

- (b) The fixist model. By adopting this dualism and additionally claiming that only so-called 'Vulgar' Latin was subject to change, this model rejects the possibility that the language spoken by the élite could ever change. For a language to remain fixed in aspic in this way is also exceptional. We must distinguish between the conservative statements of grammarians and how literate people actually spoke.
- (c) The decadentist model. A consequence of the two previous models is that the development of Latin into Romance is systematically presented as a negative process. A quick trawl through work published by philologists, and even historical linguists, produces a plethora of examples: there is an 'upheaval' in the vowel system; the case system and the passive voice are 'lost'; Latin 'degenerates', passing through a period of 'chaos', etc. The evolution of Latin is generally presented as an inevitable attrition of good language.
- (d) The diglossic model. This model, which is a development of both the dualist model and the fixist model, had its hour of glory (Lüdtke 1964). I shall not go here into the arguments which have essentially led to its being invalidated, but will merely note that we have to reject any notion of diglossia when describing the linguistic situation prior to the eighth century, and perhaps even the ninth. In a situation of vertical communication, any profound difference between two languages (Latin and Romance) would have represented a real barrier to communication, and would have been quite untenable over a period of several centuries, contrary to what has been claimed. (Even Lüdtke, one of the chief proponents of this theory, came to reject it, despite having spent thirty years arguing in its favour; see Lüdtke 2005.)
- (e) The authoritarian model. Linguistic theories are, alas, often arbitrary. We are still some way from a complete scientific description of language. Recent work has tried to apply the methods of current theories of generative grammar to language change (Salvi 2004; Vincent 2001). Although we have much to learn from this type of research, we must constantly bear in mind that the formalisms it uses are mostly self-referential heuristics rather than keys which will unlock the mysteries of historical linguistics. And the speed with which modern research proposes models of language, accepts them, and then rejects these same

models out of hand in favour of some counter-proposal (Butler 2003; Fuchs 2004; Jackendoff 2002) suggests that historical linguists should be cautious in their use of such models. It is crucial that we respect actual data (the global view from sociolinguistics) and avoid imposing arbitrary models (Rastier 2001a; 2001b).

4.2.2. Proposed models⁵

(a) *The continuum*

- (i) *Diastratic*. Two considerations guide this model. First, Latin was a living language and, as such, had no reason not to follow the same rules as any modern living language. Second, Latin belonged to the class of languages which also have a written form and are therefore appropriate for the interactions required by an advanced literate civilization. Its instantiations therefore fan out or diffract into a continuum which can be divided, albeit arbitrarily, into a series of strata, ranging from the most 'literary' to the most 'colloquial'. There can be a very large number of these strata, without any implication that the top and the bottom of the cultural scale are linguistically alien to each other (Oesterreicher 1995; Kramer 1999). Within this continuum, a plethora of communicative acts may flow freely throughout the diasystem (Stefenelli 1995).
- (ii) *Diatopic*. Although there was undeniable horizontal fluctuation across the geographical extent of the Empire, this fact should not be seen as implying any fragmentation of the Latin speech community. On the contrary, everything points to communication between different regions having been straightforward and untrammelled during the third and fourth centuries, perhaps precisely at the moment that the wholesale spread of Christianity was acting as a powerful unifying force on language (Brown 1992; 2001). When we have evidence for communication breaking down, this is nearly always due to specific circumstances arising on the fringes of the Empire; it does not indicate that there were 'Latin' speakers who spoke something that was not Latin (that is, Romance), but rather that there were communities of foreign speakers who had not yet adopted Latin (Herman 1995).

5 Banniard (1992a: ch. 9; 1993a; 1995b; 1996a; 1998a; 2001a; 2004; 2005d), Beckmann (1963), Wright (1982; 1991; 1993; 1997), van Uytendaele (1977; 1984; 1994; 2000).

(b) *Positive internal causation*

In our search for the causes and cadences of the development of Latin, we shall give priority to dynamic internal factors, that is, to positive factors which point in the direction of observed change from within Latin speech itself. This amounts to saying that the whole system is involved in the change, regardless of diatopic or diastratic variation. We shall disregard random fluctuation linked to a particular place or social group, to focus on the overall trend followed, or rather generated, by the speech community, involving both conscious and unconscious change (Coseriu 1973; 1977d; 1992; 2000). In this scheme of things, the fundamental building blocks of future developments can be found at the heart of literary production (Banniard 1996; 2002a; Oesterreicher 2001b), just as much as in so-called 'secondary' texts (recall that we are not regarding external social factors as fundamental causes of change), and in the heart of the cities – much more so than in the countryside, which is by nature conservative. Once again, external geographical and ethnic factors are irrelevant here.

(c) *The stability of complexity*

The level of linguistic complexity remains unchanged, despite all the changes to the system – the Romance languages are neither more nor less complex than Latin. The replacement of the case system by a system of prepositions scarcely represents an effortless shift (Vandeloise 1993). The passive voice did not vanish; it was recreated in a different guise. Speakers are creative regardless of the century they live in or the social group they belong to (Labov 1972b; Trudgill 1991; Vincent 1996; Gadet 2003). Essentially, one must accept that a living language is complex regardless of the level of literacy of its speakers (Manessy 1981; Pinker 1994). The creation of the large number of subsystems which lead spoken Latin to become spoken Romance is as complex a grammatical development as the no doubt similar process which gave rise to spoken Latin itself, a millennium or so earlier.

(d) *Modularity and different levels*

A change in pronunciation does not necessarily correspond to or result in a morphological change. The modularity of the different levels used to describe and analyse a language synchronically is even more pronounced over a sufficiently long diachronic time-span (Andler 2004; Biber 1998; Butler 2003). Changes at these different levels may to some extent be correlated, but they are not simultaneous (Banniard 1980b; Herman 1995).⁶

6 In some quarters, there is an unthinking tendency to view any change as starting with pronunciation, rather as schoolteachers tend to assume that pupils with a bad accent

(e) *A historically informed approach*

However arbitrary diachronic classifications may be, we shall seek to ensure that our model is historically informed. This involves postulating the existence of diachronic dialects which form an unbroken chain between Latin and Romance. These intermediate entities will enable us to respect the sociolinguistic evidence and to take account of both change and continuity. In particular, we must be aware that the emergence of a new structure is not the same thing as the incorporation of that structure into the system, nor, a fortiori, the same thing as the elimination of the earlier structure which it may eventually come to replace (Beckmann 1963; Herman 1989; Klausenburger 2000; Löfstedt 1933; 1941; Pei 1932; Sas 1937; Stotz 2004; Wright 1995b).

(f) *Fluctuation*

Developments are neither inevitable nor linear. They can take place at a differential rate according to the phonema involved. The distinction between nominative and accusative case was reinforced in LSLat. in the third declension, with the emergence of nominative plurals in *-i*, such as **flori*, which created an opposition with accusative *FLORES*; meanwhile, the genitive plural ending *-ORUM* was extended from the second declension to the third, replacing classical *-UM*. Both these developments reinforced the case system in advance of its ultimate demise (Uddholm 1954:60). A new pronunciation or morpheme or rule of syntax may remain latent for a century or two, before expanding exponentially through the language.

4.2.3. Modelling

Rethinking the chronology of linguistic change in the light of historical sociolinguistics therefore leads us to look for models which can handle data from very different levels in a uniform way. At the same time, we should be under no illusions about our ability to construct a unitary and multifactorial theory, given our present state of knowledge. We have established that the transition entered a significant phase in the third century and was completed in the eighth century; we shall try to construct schemas which offer insights into this

have no knowledge of a language. There are good examples of this in French, where the loss of certain final consonants (notably [s] and [t]) is regularly cited as the reason for the loss of the two-case system and the increasing use of subject pronouns, and itself attributed to the fact that speakers of old French were lazy. It is easy to argue against this claim and suggest that, on the contrary, speakers of old French were simply continuing a deep-seated and centuries-old tendency to replace suffixes by preposed elements (Banniard 2005b).

process, not from a 'microstructural' point of view (rule by rule or form by form) but from a 'macrostructural' perspective akin to that of a satellite photograph.

Since it is impossible to deal with every type of phenomenon in a chapter of this nature, I shall single out morphology, or rather morphosyntax. Putting together sociolinguistics and the models we have outlined suggests the following scenario:

- (1) Stage 0 (preliminary phase). A number of variables (morphemes or syntactic structures) begin to surface in the language, although they remain very much minority variants compared to usual forms and expressions. These variables constitute a meagre quarry of building blocks, prototypes which are available for use diffusely in linguistic exchange. Their emergence and use result from the operation of complex principles, including considerations of style (poetry). They are on the fringes of the diasystem, and their very existence is stochastic.
- (2) Stage 1 (initial phase). The frequency of these building blocks progressively increases, to the point at which their numbers become significant. They now contribute to an expansion of the diasystem, by moving from its fringes towards its centre. On each occasion, the process involved is the same: alongside the normal morpheme or piece of syntactic structure which recurs in the language of the speech community as a whole, an alternative form or structure is introduced – created or reconfigured – by one or more speakers. The creativity of these speakers is driven by their desire to express themselves more accurately, more clearly, or even more individually: the new structures are marked with respect to the ones with which they alternate, but can be seen as being in free variation with them (which is not to say that the choice between the two is completely unmotivated); they spread through the spoken language by a process of fractal diffusion (to borrow a term from the mathematical modelling of 'natural' phenomena; see Peitgen and Richter 1986). There is a shift in the diasystem of Latin, which nonetheless remains coherent. The forms in question are no longer stochastic, but now have a determinable probability.
- (3) Stage 2 (intermediate phase). The innovation which is being adopted polycentrically ceases to be a random or arbitrary variable and starts to generalize across the Empire as a whole and to become part of the grammar. The old and the new forms begin to compete in the diasystem (unlike reconstructions which take Romance as their starting-point, this

model does not posit the immediate elimination of the earlier forms). As it advances, the marked form tends to become less motivated, and so *ipso facto* less marked. At this point, an intense polymorphism begins to appear, as speakers freely and constantly adopt the innovations or revert to earlier usage. Now, however, it is the earlier forms which are starting to move towards the fringes of the diasystem, which has begun to suffer from hypertrophy and instability.

- (4) Stage 3 (final phase). The original building blocks are now omnipresent in speech. Forms which were marked in Stage 1 are now unmarked, and are the forms in normal use by speakers. As the frequency of these forms in utterances increases, so they become less motivated, and hence less expressive. Conversely, and in direct proportion, the earlier unmarked form, which was the usual form in Stage 1, becomes less frequent and gradually acquires the marked status originally held by the newer form; by now it is a rare form, or even an archaism, used for stylistic effect, and is likely to disappear. At this point, the earlier forms are being eliminated from the diasystem, and their occurrence is stochastic. It is at this point, with the inversion of the diasystem, that a new language has come into being.

It is obviously difficult to date these four stages with any precision, and actually impossible to do so in purely linguistic terms. The only evidence we have linking these stages to particular periods is provided by the phases in the history of vertical communication. It seems likely that Stage 3 of the linguistic model corresponds to Phases 4 and 5 of vertical communication. One might therefore propose that Stage 2 of the linguistic model is simultaneous with Phases 3 and 4 of vertical communication and Stage 1 to Phases 1 and 2. We therefore end up with the following periodization (Gleßgen 2007):

STAGE 0:	CSLat.	Second century BC to second century AD
STAGE 1:	LSLat ₁	Third to fifth centuries
STAGE 2:	LSLat ₂	Sixth and seventh centuries
STAGE 3:	ProtoRom.	Eighth and ninth centuries

Each change leading to the inversion of an element of the system forms part of the change in the overall diasystem of Latin. Each one represents a diachronic morphological isogloss. Once a sufficiently large number of these

inversions has taken place, we have, in fact, not a generalized inversion of the Latin diasystem, but rather its complete dislocation, as a clump of new diachronic isoglosses forms (Banniard 1980b), so that, by the time we reach Stage 3, the diasystem is no longer Latin, but proto-Romance. Earlier minority structures have been integrated with structures which have been preserved, so creating a new language, whilst earlier common structures now survive as randomly distributed residue. The processes of selection, retention and discarding speed up towards the end of this development, as the new elements become increasingly immanent in the system; at this stage, maintaining a large number of competing variants would have been uneconomical.

In order to give rise to this rapid evolution, the inversion of the morphological diasystem must have begun, during Stage 2, in a tentative and fragmentary way: we may use the metaphors of 'islands' or 'bubbles' to indicate how microstructures of the new type emerged into the system. Alongside them there remained microstructures of the earlier system. The complexity and intricacy of the system increased as its mechanism changed on an ever-larger scale between the third and eighth centuries. These microrestructurings emerged and spread randomly; change must have propagated by a process of fractal diffusion (see above). The reorganization took place according to a model of 'deterministic chaos', also known in mathematics as a 'non-linear dynamic system' (Bergé and Pomeau 1995; Gleick 1989; Lurçat 1999; Stewart 1989). The nature of this process explains why it is so difficult to link a particular change to a specific region, locality, author or text.

In this way, the transformation of the system follows two curves which are inverted with respect to one another. The marked form of late spoken Latin eventually becomes a Romance form when it becomes so overwhelmingly prevalent that it is now a central element in the new system. The unmarked form of late spoken Latin eventually becomes a Latin form when it becomes so vanishingly rare that it is now on the margins of the system. Whether it is eliminated or retained depends on factors which we shall not go into here. It may remain available as a stylistic archaism, having effectively changed places with the form that was marked in Stage 1 (compare the survival of non-prepositional indirect objects in the oblique case in 'classical' old French), or it may survive in a new or complementary function (compare the preterite, which becomes the 'simple past' once the creation of a 'resultative past' has led to the 'compound past' being established).

This type of curve describes formal developments at the level of active competence. Formal developments at the level of passive competence follow

a curve which is comparable, but which crucially lags behind the 'active' curve in respect of the area becoming marked. It is this part of the language which is threatened with extinction by simply being forgotten. Whether or not it survives in collective memory depends on three factors:

- (a) Accommodation across two or more generations;
- (b) Inertia, arising from the fact that there may be limits to the rate at which the language can incorporate new material;
- (c) Diastratic accommodation between literate and illiterate speakers.

What can be described for one piece of morphology can be reproduced for every other piece, one unit at a time. Seen from a diachronic perspective, we therefore have a series of S-curves which evolve in loose parallel with one another. If we put these curves side by side, this adds a third dimension to the two dimensions of each individual curve, and this three-dimensional representation will give us a structure that now resembles a mountainside. This gives us a picture of positive changes in frequency (from rare and marked to frequent and unmarked). Alongside it will be a complementary picture of negative changes in frequency (from frequent and unmarked to rare and marked) which will have the same sort of appearance. Synchronically, at the summit of the 'mountain', the sum of the positive curves will yield a network of isoglosses which will represent the typology of the new language. The sum of the negative curves will yield a less coherent picture of default isoglosses, where persisting features describe broken or erratic lines.

A detailed schematization of this model would result in an exponential increase in complexity, and would soon become impossible to describe in a chapter of this length. This fact alone should alert us to the complexities involved. There is only room here for a handful of illustrative examples. If we begin by postulating one curve for every morpheme – genitive case, dative case, synthetic passive infinitive, infectum of deponent verbs, etc. – then each of these can be further divided into as many strands (or 'microcurves') as there are declensions or conjugations, or even lexemes, whose individual history may deviate slightly from the overall curve, whilst conforming to the general trend; conversely, curves and microcurves may clump together in 'macrocurves', involving, say, a number of lexemes or a whole declension.

These intricacies remind us that modelling language change in real time is an enormously problematic task, and historical linguistics would be well advised to draw on every possible source, from large-scale diachronic grammars to microdescriptions of texts and varieties, and never to lose sight of the fact that those who are responsible for language change are speakers, not linguists.

As this model also makes abundantly clear, the typological change from Latin to Romance is firmly anchored in the internal history of Latin. Essentially, the linguistic transformation obeys the same logic everywhere, from the centre of Rome to the far-flung marches of the Empire: the logic of the internal dynamic of Latin speech, which has little to do with changes in external circumstances, and which probably also changes little over time. Inscriptions on the most modest Christian tombs in the catacombs testify to this fact (Guyon 1987; Herman 1991). By this reasoning, the responsibility for the phonetic and phonological fragmentation of Latin speech is essentially shifted on to the ebb and flow of substrates, superstrates and adstrates. These fluctuations will be attributed to minor differences in the 'initial states' of phonetic microsystems which lead to long-term developments independent of the internal typological dynamic outlined above. These developments are cyclical and exponential, and result in the emergence of the individualized contours of different Romance 'accents'.

5. Transitions

5.1. Chronological linguistic cross-sections

The problem of providing a precise historical account of language change according to this model has still to be tackled. The first step is to isolate differences between strata over the long term, and, if possible, to date these. Once again, I shall restrict myself here to a discussion of morphology. We may distinguish four types of structure.

(1) *Persisting structures*

These are structures which are carried over from Latin into proto-Romance, and often beyond, for instance:

	LATIN	PROTO-ROMANCE
1 <i>Present</i>	+	+
2 <i>Preterite</i>	+	+
3 <i>Past of infectum</i>	+	+
4 <i>Past of perfectum</i>	+	+
5 <i>Present subjunctive of infectum</i>	+	+
6 <i>Past subjunctive of perfectum</i>	+	+
7 <i>Suffixal marking of person</i>	+	+

It would be easy to extend this list, especially if one takes account of forms such as the imperfect subjunctive, which continue to be attested into the Romance period in circumscribed areas and periods, and which were therefore in all probability still in use in the LSLat2 period.

(2) **Innovating structures**

These structures represent new developments whose generalization makes a major contribution to changes in the diasystem, for instance:

- (a) *Analytic passive in the infectum*
- (b) *Analytic past (so-called 'compound past')*
- (c) *New future in -r-*
- (d) *New future-in-the-past in -r- (new conditional)*
- (e) *Reflexives using se/sibi*

The emergence of these new forms is complete by the eighth century. They are directly attested in Romance texts from the ninth century onwards and indirectly attested in Latin texts of the LSLat2 period. If we assume that they emerged according to the exponential model proposed above, their presence in the language will rapidly have become preponderant. However, it is not axiomatic that literate speakers of the period will have been able to identify them: indeed, the first two structures have a long history in the language (they have, so to speak, an 'alibi'), whilst the third and fourth, although representing more of an innovation, may sometimes have been confused with the second future (the future of the *perfectum*), which was also formed in -r- (Banniard 2002b).

(3) **Metastable structures**

These structures involve forms which do not have a fixed status in collective speech. Sometimes they are becoming minority variants as a prelude to disappearing completely; sometimes they are being allocated a new function, this latter development often being associated with regionalization.

- (a) *Imperfect of the perfectum (pluperfect)*. This form, which survives especially in proto-French and then in archaic 'high' old French, is also robust in medieval Occitan. Additionally, it survives into Portuguese. It is also found in Castilian, albeit with a new function (second form of the past subjunctive).
- (b) *Future of the perfectum (future perfect)*, conflated with the present subjunctive of the perfectum (perfect subjunctive). This future form in -r- competes with the first form of the future (the future of the *infectum*) before being ousted by the new future formations. It survives into old Spanish and modern Portuguese after being allocated the function of future subjunctive.

- (c) *Synthetic genitives in -oro-*. These were extended from the second declension to other declensions and were attested as persisting forms; they are late to disappear from speech (Pei 1932; Sas 1937).
- (4) *Evanescent structures*
- Without disappearing completely from collective speech, these forms are restricted for the most part to literate speech. Even here, they are rare before the Carolingian reform. Isolated instances can be found in illiterate speech, where they represent imitation of the prestige of literate speech (Díaz y Díaz 1992; 1998; Falkowski 1971; Gaeng 1977; 1984; 1992; Green 1991; Herman 1998b; Löfstedt 1961).
- (a) *Genitive singulars in -i and -is, genitive plurals in -um*
 - (b) *Dative/ablative plurals in -ibus*
 - (c) *Neuter plurals in -a*
 - (d) *Synthetic passives in -ur*

All this implies that taking cross-sections – say, once a century – of these developments should assist us in mapping the transition, taking account of historical reality and periodization in a sufficiently nuanced way, so as not to do violence to linguistic realities ‘on the ground’, so to speak.

A schematic account of selected strata might look something like this:

- (1) LSLat₁, third century, Carthage. All forms of types 1, 3, and 4 are part of the diasystem. Forms of type 2 are in their early stages.
- (2) LSLat₂, sixth century, Rome. Forms of types 1 and 3 are part of the diasystem; forms of type 2 are beginning to be incorporated into it.⁷
- (3) ProtoFr., eighth century, Paris. Forms of type 2 have become generalized and combine with forms of type 1 to form a new diasystem. Forms of type 3 have been marginalized. Forms of type 4 have disappeared.

A systematic account of the relationship between speech and writing could be based on this type of model. In the long term, we might use these insights to produce a history of the formation of the Romance languages which would follow asymptotically the fluctuations of speech in Western Latinity.

5.2. *Reception and understanding*

The thorniest question of all is how to model the conditions and manner in which texts through which vertical communication took place were received

⁷ The Spain of Isidore of Seville provides a good example of this step-by-step evolution of morphology (Díaz y Díaz 1992; 1998), in which we can see the *lingua mixta* (in fact late spoken Latin) of the *Etymologies*.

and understood (Banniard 2005b; Herman 1996b; Koch 2008; Lüdtke 2005). This answer to this question is linked to the conditions in which utterances are produced; but production and reception are not simple mirror-images of one another. How a communicative signal is deciphered depends on many factors, both objective and subjective.

- (1) *Subjective factors*. These play a major role in vertical communication, in which a literate minority dominates an illiterate majority. The effects of such domination may range from influence and orientation all the way to constraint and oppression.
 - (a) *Ambivalence* (Graus 1965; Le Goff 1977). The complex web of relations linking this illiterate majority on the one hand and, on the other, the guardians of religion and language involves a great deal of ambivalence, ranging from forward-looking conviction to backward-looking irredentism, and does not exclude deliberate mimicry. This description is to some extent a sociolinguistic reading of the words of the Gospel: 'He that hath ears to hear, let him hear [*Qui habet aures audiendi, audiat!*']'. The effectiveness of vertical communication is particularly dependent on fluctuations in this factor: some apostles and saints were able to reach out to their flocks, whereas the Carolingian reforms were an assault on language and culture, to which the faithful to some extent remained deaf.
 - (b) *Ritual* (Jungmann 1951–54). A significant part of the Christian liturgy involves time spent on ritual: the same prayers, the same hymns, the same readings from the Gospel, and especially the same *Lives of the Saints*, read aloud down the centuries, until they form part of the common cultural heritage. The audience comes to recognize these stories that are repeated year in year out, and they start to expect and anticipate them. The many interferences between popular legends and tales of miracles performed by saints that are found in the folk tales collected by nineteenth-century anthropologists show just how strong these influences are and how far they go back. Since the same miraculous story is regurgitated every year, or even on every saint's day, a tripartite synergy arises in the hearer's mind between the topic of the text, its narrative devices and its language. It is easier to understand what one already knows. At the same time, those elements of the language that are becoming archaic are kept alive in a prestigious showcase. But the corollary is that any change to the narrative structure (independently of any change of linguistic register) will inevitably undo these timeless traditions.

- (c) *Localism* (Brown 1981; Duval 1988). The fact that the worship of a given saint is firmly rooted in a particular locality also aids the functioning of vertical communication. Over time, every *pagus* and every *ciuitas* acquires its own saint, who becomes the locality's patron, successfully marrying the universalism of the message of the Gospels and the need to create a specific link with the place where an individual believer lives and dies. The development of the *inhumatio ad sanctos* (burial near saints' relics) in the sixth and seventh centuries gives concrete effect to this geographical particularism, which rural priests were eager to promote (Godding 2001).
- (2) *Objective factors*. This is where language itself comes in. We have dealt above with linguistic complexity and with choice of style and pronunciation; it remains to discuss the reception of a message *in situ*. First of all, however, we should add some further guiding principles to the rather partial model already proposed.
- (a) *Linguistic 'distance' and its limits*. If a complex message (and even a sermon in a simple style falls into this category) is to be received and understood by its audience, the distance between the language of the source and the language of the target will have to fall within certain limits, which can be no greater than those between two dialects (say, the difference that separates Languedocian from Limousin in our own time), and which will rule out a clear difference between two languages (such as the difference between modern French and modern Spanish).
- (b) *Hierarchy of elements determining distance*. We shall not deal with here with pronunciation: there is fluctuation, but there are no significant differences between the literate and the illiterate until the Carolingian reform. As far as morphology is concerned, provided one accepts that competing variants continued to exist side by side within a diasystem characterized by intense polymorphism, then the communicative gap cannot have been that great before the eighth century, for the following reasons:
- (i) Speakers will have retained a passive awareness of forms which were being eliminated or had already been eliminated from the language, but only recently, and then not everywhere.
 - (ii) Many 'idiom chunks' or 'syntagmemes' existed, which had been repeated down the centuries and were ingrained in memory.

- (iii) Texts designed for general communication will have contained a significant amount of redundancy, with expressions regularly repeated with only minor differences.

As far as the lexicon is concerned, the rules of rhetoric applied to pastoral discourse provided a sufficiently sound basis for information to pass easily, as witness the large number of Christian terms which entered ordinary speech. And diachronic studies of the geographical distribution of lexical items have shown that the Latin West formed a relatively coherent unit until the eighth century (Stefenelli 1998; Wright 2003).

The main criteria according to which linguistic distance can be established therefore involve syntax and semantics (albeit not the rigorous formal syntax and semantics of recent linguistic theory). Syntactic and (structural) semantic considerations enable us to abandon out-of-date models based on simplistic dichotomies and what amounts to an insultingly negative view of the communicative capacities of illiterate speakers. The degree of reception and understanding of an utterance in vertical communication depends first and foremost on the ordering of blocks of discourse.

Let us consider these blocks as a whole:

1[NP₁ {acc/ DOC/ DO}] // 2[NP₂ {dat/ IOC/ ATT}] // 3[NP₃/ PP {Prep. + abl./ Prep. + IOC/ CCPlace}] // 4 [V].

We regularly find the following orders in CSLat.:

[2/ 1/ 3/ 4] or [1/ 2/ 3/ 4] or [1/ 2/ 4/ 3] (Charpin 1977; 1989; 1990; Hofmann 1959).

They are impossible or very rare in CSFr.

On the other hand, in LSLat₁ and LSLat₂ we frequently find:

[1/ 2/ 4/ 3] or [3/ 1/ 2/ 4] (Löfstedt 1933; 1942; Pinkster 1991; 1995).

In 'classical' old French, the most frequent order seems to have been:

[4/ 1/ 2/ 3]

although other orders are widespread:

[3/ 1/ 2/ 4] or [2/ 1/ 3/ 4] (Herman 1954; Marchello-Nizia 1995; Richter 1903).

We can attempt to establish correlations between these element orders and different registers or literary styles, and so on; after all, all linguistic norms are part of linguistic knowledge. But let us nonetheless assume that there is a

significant discrepancy between element order in ordinary language and the element orders found in literary texts. It remains the case that these literary word orders can be perfectly well received and understood by the medieval French speakers who are their target audience (unless we assume that, in this case too, there has been a breakdown in communication!).

This comparative approach can shed light on how vertical communication in Latin worked ‘on the ground’, so to speak. Utterances with element orders which are well attested, albeit as minority variants, in ‘classical’ old French correspond precisely to those orders which represent the majority of utterances in LSLat1 and LSLat2. As we have been at pains to adopt a careful historical approach, we should also note that the natural element order in the eighth and ninth centuries is close to [2/ 1/ 3/ 4], up until the point at which the new order takes over. In other words, we must postulate a transitional phase in the case of element order, as well. An obvious consequence is that the members of the speech community retained considerable capacity to understand utterances based on ‘Latin’ element order until at least the eighth century.

The reasons for the alleged insensitivity of literate speakers to the changes taking place in popular speech thus become clearer; even at its most inelegant level, popular speech retained a rhythm and a word order which were a fairly faithful reflection of Latin. So the signals which hearers rely upon arrive in an order which does not take them unawares. Since the pastoral lexicon is broadly speaking the same as the ordinary spoken lexicon, hearers have a formal structure to latch on to which is not too disorientating, at least up until 700–750. Misunderstandings which might arise from morphological differences can be resolved by relying on semantics and pragmatics, linguistic context, and pointers in other parts of the grammar (van Acker 2007). This type of macro-analytical description complements rather than contradicts micro-analyses based on modern methods.

5.3. *Text, communication and speech ‘on the ground’* (linguistic archaeology)

In practical terms, this period can be seen as presenting us with more transitional texts than is usually claimed. In the same way that during the classical period we find the building blocks of future developments at the heart of literary production, especially poetry, so, too, literary texts of the sixth to eighth centuries provide all manner of ‘missing links’ between Latin and Romance (Avalle 1965; Stotz 2004; van Uytenghe 1995). Let us take the following short extract from a *Life of Saint Radegund* (*Vita Radegundis* MGH, SRM, t. 2) written in Poitiers in about 600 by an elderly nun who had a

smattering of Christian Latin. It is a less sophisticated version than the one produced a short while earlier by Venantius Fortunatus (in this sense, the history of communication is moving in exactly the opposite direction from Alcuin's rewriting of the *Vita Richarii*).

- (1) *Illud quoque quis explicet, quanto feruore excitata ad coquinam concursitabat suam faciens septimanam? . . .*

'How can anyone describe her excited fervour as she ran into the kitchen, doing her week of chores?'

- (2) *Aquam de puteo trahebat et dispensabat per uascula.*

'She drew water from the well and poured it out into basins.'

- (3) *Holus purgans, legumen lavans, flatu focum uiuificans, et ut decoqueret escas, satagebat exaestuans, uasa de foco ipsa levans, discos lauans et inferens . . .*

'She scrubbed vegetables and washed legumes and revived the hearth by blowing so that she might cook the food. While it was busy boiling, she took the vessels from the hearth, washing and laying out the dishes.'

Everything necessary for this language to be received and understood when read aloud is present. Sentence 2 has the element order [NP₁/ P+NP₃/ V + V/ P+NP₃], which would be immediately comprehensible to any illiterate speaker. Each NP₃ is the object of a preposition which will survive into Romance. The neuter gender has not yet been eliminated from speech, and it would in any case have been a robust part of speakers' passive knowledge, owing to their awareness of its use at earlier stages of the language. Finally, the semantics of the two VPs creates a crucial link in meaning which enables the sentence to be understood.

In the third sentence, we need only note the phrase *flatu focum uiuificans*. The audience would certainly not have been troubled by the element order [NP₃/ NP₁/ V]. But what are we to make of the bare instrumental ablative? We might postulate a stylistic affectation on the part of the author (the modest alliteration f-f-v-v), which could in itself explain the shift to a slightly higher register. But it would be wrong to interpret this as the author resorting to a type of expression which was no longer current in the language. Structures like these still form part of the diasystem; although they are increasingly marked and marginalized, these characteristics merely give them the character of archaisms or features of elevated style. Merovingian polymorphism is shot through in this way with considerations of discourse and stylistics.

This short text can therefore be seen as an example of totally effective vertical communication (stage 2 of our chronology). The relationship between graphemes and phonemes is clearly looser than before (for instance, *focum* will

have been pronounced something like [fuɔwo]). Globally, however, the relationship between writing and speech remains quite close, and there is therefore only moderate tension between written and oral communication.

However, the true relationship between the two types of communication is extremely difficult to gauge, given that, in our attempts at reconstitution, we are denied access to a factor whose importance is gradually emerging as a result of work by modern instrumental linguistics – intonation (Rossi 1999). As noted above, intonation plays a crucial role in immediate oral communication, yet its importance is only now being realized (in addition to or as part of its description). The sentences which we attempt to analyse and understand using our two-dimensional grammatical descriptions also involved a third dimension of transmission, in the shape of intonational conventions which enabled an orator or an individual declaiming a text to provide clear pointers to its meaning.

Conversely, from the eighth century onwards, we begin to find utterances which treat proto-French as an acrolect, for all that it may be clothed in Latin garb (Banniard 1994; 2003b; 2008b). The Romance languages were therefore not engendered in an obscure area to which they were relegated by *grammatica* (Avalle 2002; Frank *et al.* 1993; Koch and Oesterreicher 1985): provided they are analysed using the methods of differential dialectology, original documents of the transitional period (Atsma and Vezin 1981–82) give us access to this area of diachronic interference. The notion of ‘preliterate French’ should therefore be abandoned in favour of ‘Romance written as if it were Latin’ from the eighth century onwards. Every serious study of the newly Romance-speaking Europe of the eighth century, be it of Spain (Wright 1982; 1995b) or of Italy (Lazard 2007; Sornicola [to appear]), begins by examining all the data we have at our disposal and then proceeds to subject these data to an analysis founded on complex models of historical sociolinguistics and non-linear dynamic systems, thereby opening up solutions to what had hitherto been intractable problems. Naturally, these discoveries come at a price – that of increased intellectual complexity (Peitgen and Richter 1986; Stewart 1989) – with the result that notions such as ‘medieval Latin’ now also require substantial refinement (Ravier 2005; Andrieux-Reix 2005). But that is a discussion for elsewhere.

Periodization

ROGER WRIGHT

1. Periodization in history

Historians like to divide history into periods; hence they have invented such labels as 'Contemporary France', 'Pre-Columbian America', 'Medieval Italy' and, more generally, 'The Middle Ages', 'The Renaissance', 'The Enlightenment', etc. (see, for example, Dumoulin and Valéry 1991, or Jordanova 2000: ch. 5). Romanists, looking back, also like to periodize. Banniard, for example, who has taught us not to overdistinguish diastatically with reference to the Earliest Middle Ages, distinguishes periods diachronically to a surprising extent; he hypothesizes three successive separate and precisely dated stages of late spoken Latin, 'latin parlé tardif 1' in the second and third centuries AD, 'latin parlé tardif 2' from the fourth to the seventh centuries, and 'protoroman' from the seventh to the ninth centuries (Banniard 1997:30–36; see also Banniard, this volume, chapter 2). Some historians of language have invented linguistic periods based on great historical events or periods, such as 'Renaissance French' and 'Golden-Age Spanish', or on relative time, such as 'Old French', 'Modern Portuguese'. This latter step is not self-evidently justified, since most historical events do not obviously have direct effects on the development of a language. Even periods based on the names of centuries, such as 'Fourteenth-Century French' or 'La Lingua del Settecento' are misleading, for not even the Millennium Bug could have caused speakers to change from (e.g.) 'Twentieth-Century French' to 'Twenty-First-Century French' overnight. Increasingly, the important events in language development are thought to be language-internal, with the result that such externally based names now seem to be inadequate. Periodization has thus long been a topic of interest to Romance linguists, although not to historical linguists in general, and the term is, surprisingly, absent from Trask's *Dictionary of Historical and Comparative Linguistics* (Trask 2000). Even so, recently, Penny (2000:5), has declared baldly that 'all notions of periodization are misplaced in

language history [. . .]; linguistic development is as seamless as all other cases of linguistic variation'. He is right, but it is still worth considering further the implications of this realization, since Romanists have wanted for practical purposes to call separate chronological stages of Romance by separate labels. In particular, we need to investigate why Latin had its name changed to Romance, and why Romance then came to be thought of as being several separate languages with different names, before considering the validity of such labels as 'Old French', 'Middle French', etc.

2. Changes of language name

The question of periodization is thus not so much that of how and when a language changes, for change is continuous ('seamless', in Penny's description, who is there comparing diachronic stages with the seamlessness of the synchronic dialect continuum), as of why a language should change its name. For there is no necessary or even sufficient reason why a language should ever change its name at all. However much a language changes, all speakers live in a sense of continuity within their own lifetime, and we would not like to think that we are speaking a completely different language at the end of our life from the one which we were speaking at the age of ten; nor that we are speaking an early version of something that is only due to achieve full validity later, or a decadent version of something that flourished much earlier. Although all language states are indeed transitional between earlier and later ones, the language of each synchronic period has validity in its own right. Describing 'Middle French' as a period in which some of the characteristics of 'Old French' are disappearing, and some of the characteristics of 'Modern French' are being introduced, hardly helps us understand the language of the time; similarly, calling the language spoken in tenth-century Spain either 'barbarous Latin' or 'incipient Spanish' seems equally inappropriate, but both labels have been applied.

Some linguists of a structural persuasion (including generativists) have assumed that every time there occurs a change in a language, then the whole language changes. Lightfoot (1979:14–16, discussed further in Wright 1999), for example, says this explicitly, despite the obvious problems of pinpointing the time at which any change actually happens. This perspective is unhelpful if we wish to understand what occurs in real life, and for that purpose we need to seek help from sociolinguistics. For what there is in the real world are people, not idealized structures, nor the overschematized handbooks of historical grammar of the nineteenth-century tradition, nor

the generative grammars of the twentieth; it is well known, for example, that the nominal case systems presented in university handbooks of 'Old French' are considerably neater than the wide variation which most of the surviving textual evidence attests. What we need to try to appreciate is the language as used by real speakers, and, when we consider periodization, what speakers think about their language should not be ignored.

According to the normal definition of humanity, human language has existed as long as the human species. Whether or not all languages derive from a single original language, all the languages that are spoken on this Earth today have descended directly from languages spoken long ago (although some specialists would wish to exempt genuinely creolized pidgins from this generalization). Since language changes constantly in all communities, the languages spoken today are not identical to their ancestors of ten thousand years ago, nor even of two hundred years ago, but in a perfectly intelligent sense it is meaningful to say that modern versions are the same language as their ancestor, only later. Contemporary French is a temporal continuation from old French, for example. Romance is the same language as Latin, only later; what has happened in these cases is just that the name has been changed. And the change of name often occurs long after the event; nobody at the time thought they spoke 'Old Spanish', for example, but a modern historical linguist wants each separately identifiable linguistic period to have a different name. This applies to languages of the past as much as to those of the present, so we often give them a name *a posteriori*. Specialists, that is, apparently have few qualms about referring to a language of the past with a name different from that which the actual speakers gave it. But this can confuse the issue.

A decision to change the name of a living language is rare. When it occurs, it is a deliberate decision taken by a powerful group of speakers, even perhaps by one man. The decision to call the Ibero-Romance spoken in the kingdom of Castile 'castellano', for example, may have been taken by King Alfonso X himself (ruled 1252–82). The reason for creating one or more new language names is usually based on geography, in an attempt to distinguish the speakers of one geographical area from those of another area within the same dialect continuum. Such a change of name is thus most liable to happen at times of nationalistic or regional self-assertion. That was certainly the motive of Alfonso X; it is also the motivation of those in the modern officially monolingual Spanish 'autonomías' who like to claim that they have their own language, just as the officially bilingual Catalans, Basques and Galicians do. They establish the concept and then pin it down with a geographical label that identifies the referent as existing in a particular area ('aragonés', 'andaluz',

'murciano', etc.). For such reasons many language names are the nominalizations of what were previously toponymic adjectives, invented for political purposes rather than being based on any sociolinguistic analysis; and linguistic periods can reasonably be based on such a name's first use.

One complication in the modern world is that these geographically based language splits can occur even though the speakers are still in regular contact with each other and often mutually intelligible in practice (as with, for example, Galician and Portuguese). In such a case a later historical linguist who did not know the politics would not be able to see any reason for the metalinguistic split. In earlier times, the situation was often more straightforward than it is now. For example, we can accept that proto-Indo-European was a single language once, even if we cannot be sure where or when it was spoken, and even if that language contained considerable variation within itself (which we now know is only to be expected anyway); and that the reason for the split of this single language, into many languages with separate identities and names, is merely that several groups of speakers separated physically and lost touch with each other, as they spread over the huge land mass of Europe and western Asia. Having its speakers separated physically is probably the most straightforward circumstance in which one language can come to be thought of as being two, thereby marking the start of a separate linguistic period. But there is no list of necessary and sufficient conditions which have to be fulfilled in order for us to be sure that we are talking about a different language rather than another variant of the same language: is Sicilian a language? The answer is not self-evident. Since new languages can thus in effect be created by politicians, we now have the apparently paradoxical situation that the number of languages thought to be spoken in the Iberian Peninsula is increasing at the same time as the general range of linguistic diversity over the peninsula as a whole is in fact decreasing. The same may be due to happen in Italy over the next century, if political regionalism grows even as the dialects slowly converge.

3. Possible internal criteria for periodization: phonetic

Linguists would like to be able to identify linguistic turning-points, such as what Herman (1996b) calls the 'End of the History of Latin', or the start and end of 'Middle French', etc., on the basis of reconstructable internal chronologies of changes in phonetics, morphology and syntax. But in practice this is difficult, for a single language is not a single system. Any language is a

collection of phenomena of differing types, often varying from place to place, only sometimes having a substructure of some kind available to hold the units together in paradigmatic opposition. What changes during an individual 'linguistic change' is not the identity of the language itself but one or more of its component linguistic features, even though phonetic and semantic changes, in particular, can indeed have direct repercussions on neighbouring units within a substructure.

But that perspective too is over-idealized, because even individual changes are hard to date. For example, a historical grammar of the old school will tell us that in the Iberian Peninsula $[-t-] > [-d-]$, ostensibly implying that this was a single event. The modern Romanist who knows more about sociolinguistics (such as Penny 2000) will be aware that in real life a voiced pronunciation with $[d]$ can only have come in hesitantly at first – what Milroy and Milroy (1985) call the innovation, rather than the change – and then gradually spread, leading to a time when both the old and new coexisted, with some words at the head of the lexical diffusion queue preferring the new form with the $[d]$, others occurring often in both forms (equally intelligibly), and yet others resisting the form with the $[d]$ for reasons that may or may not be clear to linguists (see Cravens 1991). The chronology of the 'change' is thus not sharp. A state of variation can last for centuries, and the variation need not even simplify out in the same way everywhere. This innovation arose when Romance was still one speech community, but words with originally intervocalic unvoiced $[-t-]$ still have $[-t-]$ in much of Italy, for example. Nor need this change have coincided with any other; it is certainly tempting to see the voicing of the intervocalic plosives as a single phenomenon, but the dentals seem to have changed before the labials, even so.

If we are going to propose linguistic periodizations on the basis of internal developments such as this, we need to decide a date for when they actually happened. So when should we date a change? To the date of the first use of the new form? But at that point the older form is still the one used in the great majority of cases. Since one thing that happens during a change is that gradually the new form becomes less marked and the old form becomes more marked, could we perhaps date the change to the time of the shift in markedness, i.e., the move from roughly 51–49 percent to 49–51 percent variation (if this is possible to gauge)? Perhaps; that would, however, offend both the old-fashioned prescriptivist and the modern generativist, who hate to allow that variation might be a normal fact of life and would prefer to see only one of the variants as an acceptable part of the language. Should we date the change to the time of the loss of all the old forms? But few phonetic changes

are unconditioned, and not many changes affect 100 percent of the lexical items that might be thought to qualify, so the continuing presence of any form which had escaped unchanged would then mean by definition that the change had not yet actually happened. There is, in short, no obvious precise time we can pinpoint and say 'that is when this change occurred'. We have to adopt a vague dating based on statistical evidence (which is for many purposes good enough): e.g., 'The Spanish sibilants devoiced around 1500?' But even then we are forced to admit that phonetic developments from 'Latin' to 'Romance' pronunciation were continuous, starting before the millennium in some cases (such as the loss of initial [h-]), and continuing today. Some specific changes occur over several centuries, at different times in different places; the loss of word-final [-s] happened in northern Venezuela over a millennium later than it did in Italy, for example, and in Castile it has hardly begun and may never happen at all. Overall, then, it seems that phonetics cannot easily be used to define when any 'period' began.

4. Possible internal criteria: morphosyntactic

Other levels of language develop separately from phonetics, naturally. But the morphosyntactic chronology is not internally consistent either. The loss of the neuter gender from Latin seems to have occurred after the loss of the ablative case, for example, insofar as we can date either development (Herman 2000: ch. 5, §1). Nominal morphology changed earlier and probably more quickly than verbal morphology. Much of the verb morphology of Latin, indeed, has not changed at all even now, in many areas, neither as regards the system nor its exponents. The use of auxiliary verbs is still increasing slowly even now, particularly in Spanish, being as long-drawn-out a process as the loss of [-s]. It is feasible as a statistical abstraction to point to an epoch and say that 'this is when Latin morphology became Romance morphology', and indeed Romanists sometimes do that, usually pointing to the seventh century. For this reason, since morphosyntax is thought by several linguists to be the most important branch of linguistics, we are sometimes given AD 600 as a date for the change of Latin to Romance, although such periodization based on morphosyntactic statistics is no more convincing than one based on phonetics.

Specialists in syntax do not agree with each other, either. For example, López García (2000) has argued in detail that, from a functional perspective, Romance syntax began with the Bible translations of the fourth century AD; yet his analysis depends on the controversial view that previous Latin syntax was not based centrally on features of the verb in the way that Romance

syntax and that of most other languages are, and there are many Romanists and Latinists who do not accept that idea. Syntactic criteria seem to most specialists to be particularly unhelpful in the present chronological dilemma, in fact. The main difficulty is that many of the syntactic features of Romance which are sometimes mentioned as being diagnostics for the identification of Romance rather than Latin syntax did in fact exist in Latin all along, even if only as occasional and unfavoured variants rarely occurring in texts. The plays of Plautus attest features which Romanists want to categorize as 'late', despite the ostensibly embarrassing fact (except that Romanists do not seem to be embarrassed by it) that Plautus wrote in pre-Classical times. The use of grammatically reflexive *se* with an inanimate subject and passive meaning, for example, is found in Plautus, and also turns up in the discursive text of some of the Roman Grammarians themselves. Even the use of *se* as a non-agentive passive with an animate human subject, often said to be a distinctive feature of Romance syntax (such as French *se brûler* and Spanish *quemarse* for 'to get accidentally sunburnt'), sometimes even said to be a recent development, is attested in the *Cena Trimalchionis* of the first century AD. The use of *quod* and a finite verb after a verb of saying or perceiving, rather than an accusative and infinitive, is also often said to be a sign of Romance, and is also found in Plautus as an available variant, possibly marked but certainly not peculiar. Conversely, accusatives and infinitives are often used now, particularly after verbs of perception (Italian *la vedo ridere*, Spanish *la veo reír*, 'I can see her laugh'). Often, then, in syntax, the supposedly new is not very new and the supposedly old is still with us, and even then usually remains intelligible in older texts long after it has become unusual in active speech.

Thus there is rarely, perhaps never, a syntactic turning-point. Syntactic change is always slow. Word-order change is a prime example of this. It is wrong, as Pinkster has pointed out (e.g., Pinkster 1991), to state baldly – as people often do – that Latin had SOV word order and Romance has SVO; other word orders often turn up in Latin, and in Spanish, at least, all six possible word orders are now acceptable. For periodization purposes, there seems to be nothing tangible to date, even if we could. The usual word order of these three constituents in the *Peregrinatio Aetherae* is similar to that of the Ibero-Romance texts of the thirteenth century. In short, despite the impression given by manuals, in syntax and morphology the new does not neatly displace the old. The state of coexistence of both old and new during the last two thousand years is such that once again statistics are the only means available to decide when the once unmarked becomes marked and the once marked becomes unmarked. But statistics have no more ontological validity

than the data they summarize, and we surely cannot say when Latin ‘becomes’ Romance on the basis of the syntactic phenomena that have been used as diagnostics to identify one rather than the other; and still less to identify any subsequent linguistic ‘period’ of merely two or three hundred years, such as ‘Old French’ or ‘Modern Portuguese’.

5. Potential internal criteria: lexical

Vocabulary is hardly a diagnostic tool for periodization at all. Even avalanches of loanwords are not sufficient to create a new language. It is generally agreed, for example, that the huge presence of French loanwords in fifteenth-century English has not prevented that language from being identifiable as English. Even after the influx of Latinate vocabulary into western Romance languages in the central Middle Ages, it would be exaggerating to say that the actual languages had changed their identity enough to deserve a new name; although this has in the event been used at times as one of the criteria in the temporal delimitation of ‘Middle French’. Despite Anglicisms, Contemporary French is still French. It is at least possible to propose borrowing as a criterion for periodization, since it can indeed fluctuate between one time and another, but there is probably no such fluctuation at all in other areas of lexical and semantic development. For example, existing affixes and free morphemes are combined intelligibly every day without it being clear when (or if) their combination is lexicalized, and this seems to have been as true in the past as it is now. Lexical innovation achieved via derivational morphology never stops; neither does semantic change, for words alter their meaning all the time, and often both the old meaning and the new are around simultaneously, forestalling the existence of any clear periodizable boundary; semantic change is particularly likely, perhaps, at lively times such as the Early Middle Ages and the present day, but it is nonetheless an ongoing phenomenon of all times.

6. Orthographic criteria for periodization

There is thus no generally convincing way, based on internal linguistic data, that Romanists can date the start and end of any coherent ‘period’ in any way other than the merely statistical: ‘vaguely quantitative’, in Smith’s (2002) admirably lapidary phrase. But if we change our perspective, and consider the question from the viewpoint of the speakers of the time, rather than of internal analyses carried out with the help of modern theory, the change from *Latin* to *Romance* happened after, and probably because of, the medieval

invention of an alternative method of writing (cf. Lüdtke 2000; but also Varvaro, this volume, chapter 1). This change in written texts was essentially only orthographical; evolved syntax, post-Roman vocabulary, and even at times non-Latin nominal and verbal morphology, had all been represented with the old spelling for many years already. This was the great discovery of the Italian scholar Sabatini (e.g., Sabatini 1968) with reference to the Italian peninsula, since developed by other scholars, such as Emiliano (1999) for Portugal, and Blake (1991) for Castile and León. There are many documents from before the age of ‘written Romance’ which – despite the spelling – attest the syntax, morphology and vocabulary of the time in their non-formulaic sections, what Sabatini called their ‘parti libere’. These internal phenomena were not what constituted the novelty of ‘written Romance’; the novelty lay in the deliberately non-Latinate spelling.

This relative chronology of new *scriptae* (see also Kabatek, this volume, chapter 5) and new language names is normal (Janson 1991; Wright 2001). Although some scholars, including Banniard (e.g., Banniard 1991c), have tended to assume that a new language name can only be a label for a distinct metalinguistic concept that has already been developed, what happens in literate communities tends to be that the advent of a new writing system becomes the main defining catalyst for the idea that there exists a new language. It seems natural for literate human beings to identify their language with its written form. There may be no case in which two forms of speech which are generally thought to be separate languages are written identically. Those who claim that Valencian is not a kind of Catalan try hard, for precisely this reason, to write in a different way from that of standard Barcelona Catalan. Speakers of Judaeo-Spanish in Israel are encouraged to accentuate their metalinguistic and cultural independence from Christian Spain by writing words, even those pronounced identically as in Spain, in a distinctively non-Castilian orthography.

Using this criterion means that we are identifying a language with the orthographic system used to write it. Which is what happens in real life. Linguists complain, because we have been trained to think of writing as a relatively unimportant and even parasitic mode, but the normal assumption made in most literate societies is precisely that a different orthography implies a different language. That is why texts written in twelfth-century France, Spain and Italy are said to be in Romance if their spelling is reformed, and to be in Latin if their intended orthography is still the traditional one, despite the obvious presence of definitively Romance morphosyntactic features and words in many of the ostensibly Latin texts, and the noticeable presence of Latin features in many of the ostensibly Romance texts. That is, the spelling

decides the language which we think the text to be in. This is still the case. Those who oppose modern spelling reforms on the grounds that they would change the identity of the language concerned are entirely right; in general, the only time a whole language is thought by its speakers to change, and thus to deserve to change name, is when the way it is written is changed. Thus histories of the French language locate the first written texts in 'French' to the ninth century, rather than to any earlier time, on orthographical criteria, even though there is no obvious internal linguistic development exactly coterminous with that periodization.

Changes of name are not compulsory even in such cases, however; in 1881 the reform of Romanian spelling on the French model, which was undoubtedly of great significance, did not lead to a change in the language's name, and there was no reason why it should. And if orthographic change is indeed the only usable criterion, there seems to be little to be said for periodizing at all within western Romance after c. 1300, other than for administrative convenience.

7. Latin and Romance

The need to distinguish between the language name given to a language by the speakers of the time and that given to it by modern specialists is particularly acute when investigating the early Romance period. As Herman (1991) has shown, the difficulty is as much the fault of the subject matter as of the Romanists. Speakers of the time dated the start of the 'Romance' period (as opposed to 'Latin') later than most modern historical linguists do. As van Uytfanghe (1991) and others have demonstrated, for a long while the words *Latinus* and *Romanus* (and various suffixed cognates) could be used apparently as synonyms, and both could be applied to texts in what looks to us like Latin form. But the advent in ninth-century France of a choice of writing systems, which we would now identify without hesitation as Latin and Romance, led to a gradual conceptual separation of the two names to apply to one each of the two written forms. Latin writing continued for centuries after the introduction of Romance writing as an alternative, of course; but the presence of the new written Romance as an intentionally closer counterpart to speech meant that in due course – during the twelfth-century Renaissance – Latin became a foreign language for all, even for the Romance speakers. This development could even perhaps be seen as marking the end of the historical period of the 'Early' Middle Ages (as in Wright 1998).

We can then say that the Romance period began in France in the ninth century, on the grounds that the earliest texts we want to call 'Romance' are the *Strasbourg Oaths* of 842. Even so, we need to be aware that the novel form of these *Oaths* arose for merely contingent reasons. It was a consequence of the introduction of the reformed Latin of the Carolingian scholars, which was based on spelling pronunciations; and the subsequent need to write texts for oral performance, such as oaths, songs or sermon notes, in a mode which led educated readers or singers to approximate vernacular rather than reformed pronunciation. The distinction (between Romance and Latin) survived the end of the Carolingian period in France, particularly in enterprising cultural centres such as Fleury-sur-Loire, and by the eleventh century Romance had become an established concept both north and south of the *oc/oil* divide (for the earliest written Romance seems to have flourished in the Occitan area to a greater extent than further north). This was a distinction made by the educated, of course; all the earliest Romance texts were written down by expert Latinists, and the advent of written Romance is in no way connected to any decline in the use of Latin. The number of both Latin and Romance texts expands markedly in the twelfth century.

The name 'Romance' (variously spelt) was given to the language of some of those earliest texts written in reformed spelling. Some writers were quite explicit about this (Wright 1999). For a time the word 'Romance' lacked its subsequent geographical subdivisions, and could be used indeterminately to refer to what seem to us as being different 'Old' Romance languages. Subsequently the distinguishing words, if used, were geographical adjectives before they became nouns. This is the stage we have reached in the modern English-speaking and Spanish-speaking world; it is, for example, quite common to hear the noun phrases 'Australian English' or 'español venezolano', but it is still thought of as a bit exaggerated to use the noun phrases 'Australian' or 'venezolano', as if the speech of those areas no longer qualified to fit within English or Spanish. But if the Australians or Venezuelans were to carry out a full-scale orthographic reform confined to themselves, that could indeed initiate a new period.

8. Early Romance, proto-Romance

We might wish to give a name other than 'Latin' to the speech of the time between the Roman Empire and the Carolingians, in order to establish a 'period' in which people spoke neither the Latin of the Empire nor the post-Carolingian reformed ('Medieval') Latin of the ninth century onwards. It

seems fair enough to call it ‘early Romance’, even though obviously they did not use that phrase themselves. But we need to be aware that this phrase ‘early Romance’ has been forced on to the evidence of the time in order to aid our own modern thought processes, rather than in order to clarify the meta-linguistic world view of those who spoke it.

‘Proto-Romance’, on the other hand, is probably a label worth avoiding. The name derives from the reconstruction techniques that were used to hypothesize its existence. Hall (1976) and others reconstructed the pronunciation of Latin, on the basis of the later Romance languages; but they thought they had discovered the pronunciation of a different language altogether, and since their techniques were modelled on those that had been used in the reconstruction of proto-Indo-European, they called it ‘proto-Romance’. But reconstruction in itself cannot give us a date for when phonetic changes happened, so ‘proto-Romance’ cannot be a serious candidate for the name of the speech of a historical period. We need in addition careful analysis of written evidence, such as that which has been painstakingly provided over the years by Herman (e.g., Herman 2000), from epigraphic and other written sources, to tell us that.

9. Gallo-Romance, Ibero-Romance, Italo-Romance

The arrival of the new ‘Romance’ writing, identity and linguistic period outside the Carolingian sphere was related directly to the level of Carolingian influence in an area, and this is naturally less easy to trace outside France (cf. Kabatek, this volume, chapter 5). It also means that we may need to periodize the start of Romance at different times in different places. In the ninth century, northern Italy and the Catalan region were at least to some extent under Carolingian control and influence, but it is not clear how quickly or seriously the new reformed Latin pronunciation, and the new reformed Romance writing that this eventually necessitated, were adopted in different cultural centres. There is little such evidence before the tenth century.

There was no pre-Carolingian linguistic reform movement at the Papal court (Wright 2000). Yet the earliest explicit reference to a conscious difference between Latin and Romance in the Italian peninsula comes from Rome, in the *Gesta Berengarii* (c. 915–23; Norberg 1968:34). The famous *Placiti cassinesi* (960–63), also from the south, are brief passages in deliberately reformed orthography incorporated neat into a longer Latin text, much as the *Strasbourg Oaths* had been included in Nithard’s ‘History’ (Lauer 1926 gives the text and a translation of the latter). The concept of ‘Romance’ as a separate

language from Latin would have been reinforced by these contrasts, as the Church was becoming bilingual (Schiaffini 1959:707). Shortly afterwards a further distinction was made in Rome between different geographical modes of Romance; when the elegiac epitaph of Pope Gregory V (999) mentions his ability to preach in three manners of speaking (*usus francisca, vulgari et voce latina / instituit populos eloquio triplici*), this probably refers to what we might call Gallo-Romance, Italo-Romance and medieval Latin.

The period of the concept of 'Romance' was thus brief in Rome, and it soon split into those of separate Romance languages. Even though most of the tenth-century evidence we have from Italy seems to come from the south, both the distinction between Latin and Romance and that between Gallo-Romance and Italo-Romance were also made in the second half of the tenth century in northern Italy. The first explicit evidence there is said (since Novati 1926:32) to be the reference by Gonzone di Novara in 965 to *usu nostre vulgaris lingue, que Latinitati vicina est*, as compared to that of the French. This date closely follows that of the *Placiti*. Gonzone could well have heard of them; accordingly we can conclude that the concept (and thus the existence) of 'Romance' was present in Italy by the 960s, although perhaps this would have varied from place to place. Some scholars with international contacts seem also to have distinguished further, between Italian and French Romance.

In the Iberian Peninsula the situation was different. The Catalans were part of the Carolingian cultural world, and knew of the conceptual distinction between Latin and Romance, but before the thirteenth century the Catalans used the written Romance modes already elaborated in Provence rather than anything specifically Catalan; even the twelfth-century *Homilies d'Organyà* seem to have been written in an adapted Occitan, rather than, as is sometimes said, Occitanized Catalan. This is understandable, since Provence was ruled by the Kings of Aragon from 1167 to 1213. So we can date the arrival of the 'Romance' period there long before the arrival of 'Catalan'.

The rest of the peninsula was not directly affected by the more imperialistic aspects of Carolingian culture until the Europeanizing reforms of the late eleventh century (see Wright 1995a). There was probably no general presence of the concept of 'Romance' before then, although Hispanists feel justified in using the labels 'Romance' and 'Ibero-Romance' to refer to the speech both north and south of the religious frontier between Muslim and Christian states. Much of the Visigothic intellectual tradition survived. For example, both Isidore of Seville in the seventh century and several writers of the eleventh century used the words *vulgo* or *vulgariter* to refer to the normal speech of all (e.g., with reference to straightforward toponyms), rather than as yet

contrasting the usages so described with reformed Latin equivalents. The emergence of Romance writing in the non-Catalan Iberian Peninsula is usually located in the famous *Glosses* of San Millán; but these are also of famously uncertain date, as their proposed time of elaboration seems to advance at a rate faster than real time: that is, in the last eighty years their dating has advanced by 100 years, from Menéndez Pidal's (1926) proposal that they were written in the 970s, to the latest suggestion that they may come from the Europeanizing milieu of the 1070s. The deliberately non-Latinate form of many of these glosses probably only attests the idea of a new spelling, rather than of a new language, but since orthography is the usual criterion for periodization they are normally seen as the start of an Ibero-Romance period – or, anachronistically, even of Castilian, despite the fact that they do not come from Castile (see the wide-ranging Round Table on these Glosses in García Turza *et al.* 1999). 1070–80 could thus be a usable periodizing date, being the decade that also saw the introduction into most of the Christian half of the peninsula of the Roman rite and what we now call 'medieval Latin', from France.

The twelfth-century Renaissance marks a period, for that is the intellectual atmosphere which established the definitive conceptual distinction of (Medieval) Latin and Romance. Romanian speakers were not in this cultural orbit, so missed out on this development; it is also quite possible that this distinction never became established in those relatively remote areas whose Romance speech is still referred to with a word descended from *latinus*, such as 'ladino' or 'Ladin'; nor in non-Christian communities who had no need for Latin, most notably the Jewish Ibero-Romance speakers who referred to their Romance, written and spoken, as 'ladino' (as some still do). But the idea that there were three different Romance languages seems also to be commonly held by mid century; a few comments in Latin histories from twelfth-century Spain, for example, seem to be contrasting French and Ibero-Romance vernacular usage. This is also the conclusion to be drawn from Burnett's evidence (2001) deduced from comments by scholars who came from different areas of Europe to meet in recently reconquered Toledo. Dante was to make another tripartite distinction a century later. Unlike the distinction between Latin and Romance, which had been made (rather unusually) on a speech-community-wide orthographical basis, these later conceptual splits were geographical in the ordinary way. The famous 'descort' of Raimbaut de Vaqueyras was written in 1199 in five different writing systems, which seem now to be identifiable as French, Italo-Romance, Occitan, Gascon and Ibero-Romance; perhaps we have here the basis for periodizing the start of Gascon as a separate

language, but we might not yet be conceptually justified in subdividing further.

Phrases such as ‘Old French’ are still commonly used, but reference to ‘Old’ Romance languages has become controversial. Maiden (1995), for example, would not wish us to use the phrase ‘Old Italian’ to refer to the Italo-Romance of a thousand years ago. One reason for this reluctance is that the first Romance written forms were in context new and even revolutionary, rather than ‘old’ (perhaps ‘Young Italian’, ‘Young Portuguese’, etc., would be less inappropriate); but the main reason concerns the anachronistic nature of the implied perspective, which is mere hindsight. A phrase such as ‘Old Italian’ seems to imply that the speakers could see into the future, and were all on a road that led to modern ‘Italian’. It also indicates a greater homogeneity than is likely to have been the case, given the huge variability of the language over the Italian peninsula (which may explain why geographically more restricted labels such as ‘Old Neapolitan’, etc., seem less unacceptable). Diatopic variability was wider in medieval Italy than in other Romance areas, but Lodge (1993) is also for similar motives reluctant to endorse the phrase ‘Old French’, despite the time-honoured use of this phrase to refer to Gallo-Romance between the *Strasbourg Oaths* and 1300. A vague concept of ‘Old Spanish’ is also often used (sometimes intended to mean merely ‘Old Castilian’), without even having a generally accepted end-point. A compound alternative such as ‘Italo-Romance’ looks back into the past, certainly, but if we want to recapture the attitudes of the speakers this is more justifiable than phrases such as ‘Old Italian’ which look forward to the future, since speakers knew to some extent what had happened in the past but had no idea what was going to happen in the future. The later development of Italo-Romance into a state that privileged the dialectal habits of Tuscany cannot have been foreseen in the tenth century, any more than the similar later development of Ibero-Romance into a state that privileged the dialectal habits of Castilian, or the development of Gallo-Romance into a state that privileged the dialectal habits of the Île de France.

10. From 1200 to 2000

The thirteenth century brought several different Romance writing systems, deliberately standardized to some extent, and thus several different Romance languages. These are political events; for example, the conceptual break of Catalan from Occitan follows the political break of Aragon from Provence in 1213, and the conceptual split of Portuguese from Galician followed (by some distance) the political break between the two areas in the mid twelfth century.

Leonese seemed to be becoming another Romance language – that is, with its own writing system – but the definitive incorporation of León into Castile in 1230 stifled the chance of Leonese becoming conceptually independent. The important fourteenth-century texts in ‘Aragonese’ were hardly standardized at all, and largely written in Schismatic Avignon, and as a result the status of medieval Aragonese as a separate language has been ambiguous. Further north, the modern term ‘Anglo-Norman’ created more problems than it solved, since the language so referred to was French. The choice of Tuscan as the model for written Italo-Romance is also a product of these times; Renzi (1998:22) refers to ‘l’italiano antico, cioè il *fiorentino antico*’ as if the identification were self-evident.

Since there have been no further language splits, there is no need for further periodization in Romance, and we would be probably less confused without it. But many further names have become common. ‘Contemporary French’ is almost self-defining, although its end boundary advances a year every year. In between the ‘Old’ and the present day some languages have acquired a ‘Middle’ period, most notably ‘Middle French’, roughly comprising the fourteenth to sixteenth centuries (in most Romanists’ view, if they manage to date it at all), even though internally nothing dramatic happened to French around 1300 (see Smith 2002). Indeed, 1400 might seem a more sensible boundary than 1300 on socio-philological grounds, in that Froissart, probably the last important French writer not to aim to represent the standard Parisian dialect, died shortly after 1400. The label ‘Modern’ is misleading; ‘Modern French’ is supposed to have taken over from 1600 to the storming of the Bastille, whereas ‘Modern’ Spanish and other Romance languages are, in contrast, still with us today.

External criteria based on writing continue to be available. Sometimes Romanists see periods as being introduced by attempts at standardization; in Italy this would mean identifying the thirteenth, sixteenth and nineteenth centuries as turning-points, for France the seventeenth-century establishment of the *Académie française*, for Castile the reign of Alfonso X (1252–84) and the founding of the *Real Academia de la Lengua* (1713), and for Catalan, the work of Pompeu Fabra in the early twentieth century. Catalan has more claim to sensible recent periodizations on external grounds than most, in fact, given the sociolinguistic eclipse of Catalan between the Union of Castile and Aragon in 1479 and the nineteenth-century *Renaixença*.

Even for those who prefer to concentrate on periodizing the spoken language on internal criteria, choices of label can still depend largely on how changes are envisaged as occurring. French Romanists, in particular,

like to think of language change as proceeding largely via brief revolutionary spasms of rapid change occurring between long comparatively tranquil and change-free periods. This is why it seemed natural from this perspective to date a change (from 'Modern' to 'Contemporary' French) at the French Revolution, although the internal evidence is hardly so clear; and the Romanians follow them, dating 'Modern Romanian' from 1780 (although Romanian before this date is said to be 'Old Romanian', since written texts in Romanian started in 1521); see Rosetti (1973) and Todoran (1989). Similarly, Bruni (1990) dated a change of linguistic period at the mid-nineteenth-century unification of Italy (from *Il Primo Ottocento* to *Il Secondo Ottocento*).

Most scholars, though, see change as less sudden than that. Internal evidence of related changes is still sometimes exploited to propose new and relatively sharp periodizations, but historical linguistic theory does not encourage this procedure; related changes are now more often thought to be pull-chains, which are almost necessarily drawn out over a lengthy period of time (and thus hardly envisageable as being a hinge between periods), rather than push-chains, which would need to be instantaneous to exist at all, and could thus form a turning-point. Eberenz (1991) produced an interesting variant on this theme by proposing to establish a 'Middle Spanish' period during which several interrelated developments occurred. This is the opposite of the more usual procedure, as Eberenz is offering us an identifiable period characterized by related changes, rather than one characterized by stability between briefer times of change. This was an intelligent suggestion, but unfortunately the possible existence of a 'Middle Spanish' period has become a fashionable topic for discussion in Ibero-Romance studies just as the usefulness of the term 'Middle' is being questioned by those who study the linguistic period often known as 'Middle French' (e.g., Ayres-Bennett 1996:98–99). These periods are in any event conceptual inventions of a much later age (as 'Middle English' was invented in the nineteenth century; see Matthews 2000); fortunately, nobody as yet seems to have invented 'Middle Portuguese' or 'Middle Italian'.

It is also noticeable that in the modern world geographical separation is not often seen as a reason for changing a language name. Given the likelihood of future slow but universal convergence within dialect continua, such splits need never happen again. Mere physical separation has not meant that people in French-speaking areas outside Europe think that they do not speak French, for example, or that the countries of Spanish-speaking America no longer think that they speak Spanish; and when some Brazilians in the 1930s wanted to call their language 'Brazilian', for nationalistic reasons, that suggestion was

laughed away. Portuguese is still thought to be Portuguese in Brazil, as well as in Angola, Guinea-Bissau, East Timor, etc. The sea has as often psychologically united as separated emigrants from their original home (see Nielsen and Schøsler 1996), and it may even be that in future creoles will come in general to be more like their related lexifier languages, perhaps deliberately.

II. Traditionally suggested periods

For the periodizations that have been suggested for French, see Eckert (1990) and Ayres-Bennett (1996); for Spanish, see Marcos Marín (1992) and Quilis Merín (1999); for Portuguese, see Messner (1994), although Emiliano (1999) supersedes Messner for the earliest period; for Italian, see Stussi (1980), Krefeld (1988) and Maiden (1995); for Romanian, see Rosetti (1973) and Todoran (1989).

Evidence and sources

ROGER WRIGHT

The question of the nature of linguistic evidence, and of what textual sources actually attest, is a slippery one for synchronic linguists investigating their own native language, so it is unsurprising that it should also be a hard one for the diachronic analyst. The question has become harder because of the development over the last fifty years of sociolinguistics, which has led to the general understanding that variation is not just a normal feature of a living language but a necessary one; thus no text is likely to attest more than a small subset of available variants. There is always not merely diatopic variation between different ways of speaking in different geographical areas, which has traditionally been privileged by historical linguists as the apparently most salient kind of variation, but also variation between different people in the same area – often analysed in the late twentieth century through a quasi-Marxist analysis of speech communities in unobvious terms of social classes – and, as is increasingly being appreciated, between the way the same individual speaks and writes in different circumstances. The discovery that features of speech and writing are always to some extent different, even within the same speaker/writer, has seemed of no great import to the synchronic investigator interested in the speaker/hearer (particularly if he or she is the ‘ideal’ speaker/hearer who provides malleable data through the convenient conduit of the investigator’s own intuition), but has provided considerable food for thought for those who hope to recreate linguistic details of the distant past, from which the only direct evidence is written evidence. (These problems have been well discussed in the Romance field by Wulf Oesterreicher: e.g., Oesterreicher 2005.)

Given that individuals vary with each other, even when incontrovertibly speaking the same language, it has been natural to allow for such variation to exist within synchronic descriptions of a single language viewed as a whole; the only alternative would be for each idiolect to be studied separately. Ontologically, such a procedure might be justified, and many conscientious

studies have been made of the language of Romance-speaking individuals writing in the past (Boccaccio, Molière, Camoens, Hugo, Unamuno, etc.) without any discussion of how far the individuals studied should be seen as representative of the language of their age. Indeed, since most of the writers chosen for this kind of analysis stand out for their overall individuality and untypical linguistic skill, it can be argued that literary sources such as theirs are the wrong data to turn to for evidence and enlightenment about the contemporary language in general. For example, Spaniards sometimes like to refer to their language as 'la lengua de Cervantes', but not because they speak and write now, nor indeed because they think or have ever thought that they ought to speak and write, like Cervantes did. The style of *Don Quijote* has a cheerful long-winded pedantry that is most attractive, but it could never be claimed in detail to represent seventeenth-century Castilian speech as a whole. So it has seemed sensible to many historical Romanists to investigate data of other kinds, such as administrative documents that preserve verbatim accounts, wills designed to be read aloud to interested parties, diaries, journals, inscriptions, letters (all of which usually have the great advantage of being precisely datable and locatable, unlike many more literary manuscripts), or oral genres (even though they don't have that advantage) such as ballads (rather than drama, which can be as stylized as any other literary genre despite ostensibly representing speech). Data need to be securely dated and located, and since so many literary and historiographical works only survive in copies that were, or might well have been, prepared at a different time and place from their original, these can be awkward to use. A notable example concerns the twelfth-century *fueros* of New Castile (law-codes granted to towns recaptured from the Muslims), often originally drafted in Latin but translated into Romance in the following century, where the dates explicitly included are naturally all those of the original grant rather than of the translation. Most of the source texts of whatever kind, though, even if anonymous now, were composed by an individual just as much as *Don Quijote* was, and similarly represent an idiolect rather than a generality; even oral ballads, which have indeed often been shaped by several different performers, have stylistic tendencies belonging to the genre which it would be risky to extrapolate on to the normal speech of even their performers, let alone that of everybody else. The evidence of poetry, in fact, is often disregarded by historical analysts, on the grounds that poetry is not a natural register, tending to include archaisms, unnatural word orders and diction, either for ostentation, or more prosaically *metri gratia* (although syllabic verse, such as early medieval Latin 'rhythmic' verse, can be a help when studying syllabification).

The dichotomy between speech and writing has been useful to explore, but is less stark than it is sometimes made to seem; for there are many kinds of spoken register and many kinds of written register. Written texts can be, and in the Middle Ages usually were, read aloud. Spoken styles can be, and often have been, written down, with greater or lesser degrees of faithfulness. In addition, there are now media which seem to be both at once, such as text messages. Thus the clear diglossic relationship which has sometimes been envisaged between speech and writing is as much an idealization, if postulated as existing within the competence of the ideal speaker/writer, as any ideal speaker/hearer is. Diglossia has probably been a useful concept in the analysis of bilingual societies, as opposed to individuals, even though it is now apparent that under the laxest of the available definitions most literate societies would be classified as diglossic (given the presence of several registers in monolingual literate societies anyway); and even under a stricter definition, all societies which have been described as diglossic are in fact rather different from each other. The widespread use of the term 'diglossia' may thus have done little more than create confusion in the linguistic analysis of individuals in complex monolingual speech communities and the textual evidence which they provide, which is why it will not be mentioned again in this chapter.

There have, naturally, been valiant attempts to overcome these potential hurdles. The existence in the modern world of huge computerized linguistic corpora has meant that it is possible to circumvent or ignore the problem of the necessarily idiolectal nature of all evidence by amassing vast cumulations of detail. If we have available millions of words provided by millions of speakers, we can at the least draw out a highest common factor. There is, however, a human tendency to want to decide between competing variants, and depressingly often to want to call one 'right' and the other 'wrong' rather than just accepting that both exist. If there are two ways of expressing the same concept, and if we choose our questions with care, Google, for example, will tell us which of the variants is the commoner, and we can, if we are of a mind to be impressed by statistics, choose that one as 'correct', or at least as belonging to the supposed norm; even though it seems fairer not to feel compelled to finger either of them for such an honour. Such search engines may soon even supplant the largest dedicated corpora for those investigators who are only interested in synchronic states of the present, being so easy to use. The available evidence from the past can also be amassed into such searchable stores, as with Mark Davies's ever-increasing databank of written Spanish (www.corpusdelespanol.org), or the *Real Academia Española's* similar corpus (www.rae.es/corde), provided that genuinely unemended texts rather than editions are used as the

input; the fact that these two storehouses do not always agree in detail is an interesting revelation, and the fact that inevitably they are based only on written texts is an elephant so far merely lurking in a cupboard, still waiting to emerge into the room.

In practice, then, our direct sources and our direct evidence for the past have to depend on written manifestations of particular idiolects, even when the collective data are vast enough to allow us to offer statistically valid generalizations which might not in total apply to any one individual. And despite the caveats just mentioned, there are intrinsic gaps between speech and writing which need to be taken on board. Not just because the two registers always and inevitably manifest specific differences; as well as the obvious contrasts between phonetic media, which involve intonation, pitch, rhythm, sandhi and other phenomena unrepresented in writing systems, and graphic media, which involve punctuation, spaces between words, capitals, fonts, etc., there are large statistical morphosyntactic distinctions such as the greater presence of, for instance, deictics and diminutive suffixes in speech, and of, for instance, subordinate clauses and third-person verbs in writing. A greater presence: obviously not exclusive. But the greatest difference between the two is one which in practice has often not been taken into account by the diachronic Romanist; speech comes to us all naturally, but, in any literate society, writing needs to be taught. Writing at all, even if badly and clumsily and unintelligibly and inaccurately, needs the writer to have been at some point taught how to do it. Incompetent written evidence still comes from a literate source. And to the eternal disappointment of all sociolinguists and philologists, and the sociophilologists who try to combine both kinds of expertise and insight, elementary students are not usually taught to write phonetic script. Teachers always aim to have their pupils writing their words in their traditionally correct graphic shape, even in societies where the standard orthographical systems were established in a long-distant past time which predated any number of phonetic changes. In the pedagogical traditions of most societies, any graphical variant which reflects such changes, or indeed which is incorrect for any other reason, is *ipso facto* by definition wrong and to be chastised. In this respect, writing disguises speech rather than providing a photograph.

This means that as a general rule we can take spelling mistakes as potentially being evidence of evolved phonetic features, but correct spelling not to be in itself evidence of unchanged phonetics. There is, for example, a form *lueco* in the eleventh-century Riojan glosses, representing a word which comes from Latin *LOCO* and eventually becomes Spanish *luego* (['lweyo]). The glossed word in the text is *repente*, which suggests that the gloss's meaning there was that of

the later *luego* rather than the earlier *LOCO*; the novel spelling <ue> definitely attests [we], since otherwise no scribe would ever have thought of writing it that way, but the letter <c> is in itself evidence neither for [g] nor [k], since writing the correct letter <c> in intervocalic position in words that have come to be pronounced with [g] would by then have been normal. Such orthographic conservatism has to be taken into account by the analyst. A letter <h>, for example, in Romance and in Latin, has often failed to attest [h], both when found as itself (e.g., French *huit*, Spanish *hallar*) and as part of a digraph; for instance, the modern Portuguese and French digraph <ch> represents [ʃ], the Italian *ch* represents [k] and the Spanish <ch> represents [tʃ], but nowhere does *ch* represent [kh], and in none of these cases does <h> represent [h]. Such digraphs were popular when Latin borrowed Greek words containing <χ>, <θ>, <φ> or <ρ>, and <ch>, <th>, <ph> and <rh> came to be used when writing them; whether the Latin pronunciations of such words ever included the same aspirations as the Greek is rarely clear, but even if they did, the aspirations certainly did not continue into Romance.

Anglophone Romanists should not complain about this use of the letter <h>, since English and the Celtic languages adopted similar expedients. The parallel helps us understand the problems faced by both the scribes of the time and the modern historical Romanists. For example, the fact that my own surname, and the homophonous bound morpheme that gave rise to it, is always spelt in the English-speaking world as <Wright>, with the digraph <gh>, does not constitute evidence that it is pronounced as [wright], nor even as the [wɹɪxt] which originally inspired this graphical form; the letter <w>, and the digraph <gh>, when found nowadays in this word in a written source, are instead evidence of a good education and successful teaching. My name and the morpheme are never written as **<rajt>, which happens to be the closest alphabetical approximation to the normal [ɹaɪt]. Education can also lead to mistakes which would not otherwise have been made at all. Spanish authors and printers often get the name wrong, but not for phonetic reasons; in Spain it tends to be <Wrigt> or <Wrigth>. The motivation for the latter form is the same as that behind the common misprint in Spain of the *Foreing Office*: the writers know from their education that <th> is a common English digraph and -ing is a common English suffix, and here their education has created, and accounts for, errors which could not be explained if we did not take it into account.

The intrinsically confusing influence of education on the written evidence is not only to be found in spelling. Late fifteenth-century writers and printers were often led by the prevailing pedagogical orthodoxy to believe that

syntactic and stylistic Latinism was a virtue. When seventeenth-century Romance poets imitate the most artificial kinds of Latin phrase construction, such as manifesting multiple simultaneous cases of hyperbaton, the modern analyst smiles appreciatively and ignores these poems in his or her data. But the evidential value of writers whose antiquarian enthusiasms are less obvious can be difficult to assess; do apparently Latinate turns of phrase indicate merely an effective classical education in the writer or a more genuine feature of the language? (See Pountain, volume 1, chapter 13.) And if it is the former, is it still not the case that that writer speaks and provides evidence for the same language as his compatriots who would never let a hyperbaton pass their lips? There is no clear answer to this. It would be misleading automatically to ascribe all cases in which Romance grammar is similar to that of Latin to the influence of Latin; the grammar could just have failed to change. Lexically, on the other hand, words were borrowed from Latin into Romance writing all the time. Some such words became general in the language as a whole; even some affixes could do that. Others remain there as beached whales, oddities of an idiolect, curiosities that never escaped those texts. It may not have been determinable at the time which words belonged to which category.

Similar problems can also be present in otherwise potentially helpful metalinguistic comments. That is, grammarians and others can make illuminating remarks on contemporary practice, but they tend to tell us what ought to happen rather than what actually does. The literate learnt to write in the early Middle Ages from teachers who knew the tradition, and perhaps occasionally even the text, of the *Ars Minor* of the fourth-century scholar Aelius Donatus (who taught Jerome); but Donatus had not known that he was founding a long pedagogical tradition, or that he was educating the apprentice scribes of the future, since he was writing for contemporaries wishing to read the classics of the past. Speech was hardly on his radar at all, and, insofar as it was, his data are likely to have been out of date even at the time he wrote. Other grammarians are more helpful; Velius Longus in the second century, for example, and notably the fifth-century grammarian Consentius, probably working in Africa, sometimes tell us about speech. Consentius is regarded by József Herman (2000), for this very reason, as his favourite grammarian. The metalinguistic source need not be a specialist linguist at all; Augustine, for one, gives us a few revealing explicit insights into his speech (the Latin of Africa around AD 400). The Carolingian educational reformers (working c. AD 800), including most notably Alcuin of York, a non-Romance speaker with an outsider's eye, and Theodulf of Orléans, originally from Zaragoza, who apparently wrote much of the work ostensibly attributed to Charlemagne

himself, can tell us something if we let them; Alcuin's *De Orthographia* is, for example, as much about the reformed pronunciation as about spelling, but even that tells us more about what Alcuin wants people to say in reformed Latin rather than what they actually do say, either in Latin or in their vernacular (Wright 2003). After the Carolingian reforms, metalinguistic comments that help the Romanist become fewer and implications need to be deduced rather than being explicit in themselves. One kind of source that has so far hardly been exploited at all by modern Romanists, being so difficult to analyse, is that of the monolingual Latin Glossaries, particularly those which derive from the seventh-century Hispanic tradition (Wright 2008). Grammarians of the central Middle Ages are often capable of manifesting astute intelligence and providing valuable information for us as well as for their students, such as those who produced the *Artes Lectoriae* in eleventh-century Aquitaine, but inevitably after the Carolingian reforms they tell us more about medieval Latin than about Romance. The elaboration of a grammar of a Romance language rather than of Latin needed to wait for the Renaissance, but after Antonio de Nebrija's pioneering study of Spanish grammar in 1492 (Nebrija 1492a) such grammars became fashionable. Over thirty were prepared in the Spanish Golden Age (up to 1660), and there were many scholars on the case in Italy, although in Italy the intellectual field was further confused by the *questione della lingua*, usually interpretable as being the question of what kind of Italian Romance to write. Anipa (2002) has made a comparison of the features mentioned by grammarians and those attested in realist literature of the Spanish Golden Age, concluding that both may be better witnesses for Golden-Age spoken Spanish than they tend to be given credit for, in particular insofar as they attest the survival of obsolescent features in speech for longer than we might otherwise expect. Nebrija also prepared both a Latin-Castilian dictionary (Nebrija 1492b) and a Castilian-Latin dictionary (Nebrija 1495), but monolingual Romance dictionaries only appeared in the seventeenth century, as it became scientifically respectable to study vernacular languages with as much attention as Latin. Since then, metalinguistic comments on language have been abundant in writing, if sometimes contradictory. This is still the case. Unfortunately, not everybody who comments on language has sensible things to say, and the analyst always needs to remain wary.

Thus the gap between speech and writing is largely created, or at least exaggerated, by education. The way in which scribes were trained needs therefore to be studied explicitly when it can be. If we know enough from contemporary comments on education, or from any metalinguistic remarks that survive, to make deductions concerning linguistically relevant features of

the educational methods of the age in question, we should do so; and we should make relevant deductions from the written data when there is no direct evidence of scribal training practice. If a text exhibits features which we really can't envisage as being normal features of contemporary speech, not even of the speech of the writer, we are justified in wondering if the data look the way they do because the writer had been told in his training to write in that fashion. For example, when some Leonese writers adopt Castilian forms in writing in 1350 which their forerunners would not have used in the 1250s, this may be because they were being instructed to write in such a way rather than because their speech had developed in that direction (see Morala 2005). If a medieval centre starts producing documents with many proper names including a letter *k* which had been absent there before, this may well attest the arrival of a new teacher who liked to encourage the use of the *k* rather than anything phonetic. It is also possible, naturally, that such pedagogical predilections might have had an initial phonetic rationale.

Thus there is no necessary direct connection between textual change and language change, and the relationship is not the same in every linguistic aspect. Even languages which are written with alphabets tend to have teachers and speakers with logographic rather than phonographic attitudes to writing; that is, English speakers are trained to write 'rite' right as <rite> and write 'write' right as <write> and write 'right' right as <right> and write 'wright' right as <wright> even though the forms all represent [ɹaɪt] and nobody ever says [rite], [write], [right] or [wright]. French speakers are taught to write [ɛm] as <aime>, <aimes> or <aiment>, but never (nor <aim>); they are also taught to distinguish in writing between the forms <chanter>, <chantai>, <chantais>, <chantait>, <chantaient>, <chanté>, <chantée>, <chantés> and <chantées>, which for many of them all have the same sound. And yet spelling reform, however desirable it may seem to the outsider, always meets obstacles in practice, because in any language, for those who have learnt to read, it is naturally easier to read, as well as to write, the traditionally correct form, rather than a novelty inspired by phonetic transcriptions; while those who cannot read at all cannot read any form, so a form based on a phonetic transcription would be of no more use to them than the traditional counterpart.

The consequence is that phonetic change can progress a long way without evidence for it being directly attested in a written source. The realization of this dissociation was the main motivation for the twentieth-century reconstruction of 'proto-Romance' on the basis of extrapolating backwards from medieval Romance written evidence (taken to be largely phonographic) rather than by examining the written texts provided for us by those who supposedly

spoke it. This enterprise was modelled on the previous reconstruction of the 'Proto-Indo-European' spoken five thousand years ago or more, wherein there is perforce no contemporary written evidence to take into account (Hall 1976; 1983; Dardel 1996b). The changes implied in the proto-Romance reconstructions cannot be dated by this method, but Hall and his colleagues effectively established most of the phonetic features of spoken Latin. Unfortunately, this was at the expense of inventing alongside Latin a rival and supposedly contemporary coexisting separate language of 'proto-Romance', now generally thought to be chimeric, even though most of the reconstructed phonetic details are generally accepted. That is, what was reconstructed was simply the pronunciation of Latin. The main surviving protagonist of proto-Romance reconstruction is Robert de Dardel (e.g., Dardel 1996b; 2007), who has been attempting to extend the methodology into syntax. Reconstruction of an ancient state of a language on the basis of later data has helped analysts work out details of many languages all over the world, and is enthusiastically defended by historical linguists (such as Lass 1997), but it is also always recognized to be a *faute-de-mieux* kind of procedure; the specialists in Proto-Indo-European themselves are delighted when new ancient written evidence is discovered, for example, a delight which presents a marked contrast with Dardel's insistence on never using written sources, on the grounds that preferring written testimony of the late Latin period to reconstructions based on later medieval Romance is 'unscientific'.

It is clear that there are details of Latin, as attested in writing, which could never be worked out by reconstructing backwards from Romance. This limitation is particularly noticeable in the vocabulary. It is a moot point whether the converse is also true, and reconstructed features that are not attested at all at the time concerned did indeed once exist, particularly in the case of syntax. As a result of his decontextualized calculations, Dardel (1989) even presents, and apparently believes in, reconstructions of several successive changes in the unmarked proto-Romance word order, including an order which is otherwise textually unattested in Europe, seeing this conclusion not as a problem but as the crowning glory of his theoretical approach. That is, Dardel does not accept that syntactic developments can be studied from texts written by those who spoke the language he is reconstructing, which seems to Romanists of other perspectives to be a drawback of the method. Another problem with reconstruction has been that until recently it has seemed unable to accept variation; in particular, any variant reconstructed (or attested) in the past had in the theory to be allotted to a separate dialect or language in which that was the canonical form, rather than being allowed to be a simple variant

in a single multivariable vernacular; thus the many subbranches of proto-Romance, in particular as envisaged in geometric tree diagrams by Hall (1974), including various kinds of 'Gallo-Romance', etc., are of necessity dated so early (during or even before the Roman Empire) as to seem unconvincing to the classicists and historians who know that even if Latin during the Roman Empire varied diastatically and geographically, it was still one language. Outside the field of 'Proto-'Romance, this problem has begun to be confronted; the problems of combining reconstruction with an understanding of socio-linguistic variation are thoughtfully considered by several studies in the volume edited by Cravens (2005), including Cravens (2005) himself, concentrating on Italian intervocalic consonants, and by Harris-Northall (2005), on late fifteenth-century Castilian (see also Dossena and Lass 2004).

Even so, despite all the caveats, phonetic evidence can be attested in writing. It is particularly worth considering such possibilities at length in the late Latin context, where the nature of sources and evidence is most thorny. It is at times possible to deduce differences in the vernacular pronunciation of different places from the different errors made by scribes, as Herman (1990) could do from epigraphic inscriptions, despite the reservations expressed by Adams (2007: ch. 10), and as Janson (2008) was able to do concerning vowels in the Merovingian and Lombard areas. Sometimes the attestation of evolved phonetics is the consequence of a deliberate act. Occasionally, though rarely, this is the result of a wholesale decision to write everything in a new more phonographically based manner; this is what happened with the *Strasbourg Oaths* and the other early complete texts in Romance (see Ayres-Bennett 1996). More commonly, such intentional novelties resulted from an attempt to represent in writing a single word that had no standardized written form, rather than a whole text. In late and medieval Latin this could apply particularly to words borrowed from other languages, such as from Germanic in Gaul or, later, from Arabic in the Iberian Peninsula. Greek borrowings had been adapted into Latin writing earlier than these, and their Latin form (for instance, in the Vulgate Bible) can reveal something about the pronunciation of the word in one or both of the languages. Conversely, Latin words written in the Greek alphabet, a practice we find in seventh-century Ravenna, can also tell us a great deal, although even there we need to beware the temptation to treat the evidence as if it were a phonetic transcription. Personal names and placenames often offer a way into the problem for the modern investigator, although such extraneous considerations as popular etymology can make these processes less than transparent.

Usually the problem addressed by the scribe was how to write a non-Latinate name using the symbols available to him in the Roman alphabet, but occasionally new letters could be adopted from elsewhere. Thus in late Latin texts we sometimes see the originally Greek letter <k> rather than a correct <c>, particularly before the letter <a>, and regularly in the dating formula based on *kalendas*, however abbreviated; not because the sound [k] was new, but usually, it seems possible to deduce, because words written with <ca> in northern Gaul could sometimes have those letters representing a palatalized sibilant (as *CANEM* was becoming *chien*, and *CATTUM* *chat*), and <ka> could seem preferable because it unambiguously always represented [ka]. The originally Greek letter zeta, variously drawn but usually not unlike <z>, which was probably adopted to represent the affricate [dz], also came into general use in Late Antiquity; it is particularly common in the patronymic suffix of witnesses to tenth-century legal documents in the Iberian Peninsula, usually written as <-iz> or <-izi>; this led to modern <-ez>, as in *Sánchez*. Otherwise new letters were not adopted, not even from the Arabic alphabet to represent Arabisms (unlike what happened in the Anglo-Saxon context). The name just mentioned was occasionally written as *Sanggiz*, but not because it was ever pronounced with [gg]; the scribes were casting around for a way to represent [tʃ] with the symbols that they had, given that Latin had had no equivalent, and <gg> was one possibility bruited before the general recourse to the digraph <ch>.

Unintentional novelties can be enlightening too, naturally. Some are directly motivated by phonetic reality, but we usually need to reconstruct backwards from later evidence to understand what was happening in such cases. For example, when the non-classical late Latin suffixed verb *AUCTORICARE* turns up in tenth-century evidence from Galicia, formed by joining the existing *AUCTOR* and *-ICARE* into a whole which had not previously existed, as the written form <*obtorigare*>, it seems from later attestations that the letter <g> represented [g] and the digraph <ob> represented [ow] (it became Galician *outorgar*, ‘grant’). But textual change can happen even when phonetic change hasn’t. A remarkable example, also from tenth-century Galicia, is that of Latin *SOBRINUM* (‘cousin’), which became eventually Galician *sobrinho* (‘nephew’); this phonetic trajectory seems to leave no scope for confusion in the first three letters, since the sounds did not change, but the word turns up in the documentation more often as *super-* or *supr-* than with any other initial combination of letters. We can hardly postulate [sup-], and this form can only have been chosen because the Ibero-Romance [‘sobre] was written every day, correctly, as *super* (both as a free word and as a prefix), and the forms with *super-* or *supr-* must be the result

of misapplied intelligence from writers who wrongly, but intelligently, identified [sobr-] as a prefix. Such misdirected intelligence is at times a more plausible explanation of ostensibly absurd written forms than is the capricious stupidity and barbarous ignorance often imputed to the writers by generations of smugly unsympathetic modern Romanists.

What are usually seen now as the earliest sources usable as a whole as evidence for Romance are those in which the spelling system has been deliberately adapted in a novel way to represent the colloquial sound of existing words via new combinations of existing letters (as in the *Strasbourg Oaths* of 842). Thus in such textual evidence the novelty lies mainly, or in some cases only, in new written forms of words. Direct evidence for other aspects of Romance had been attested in writing already for a long time. Romance syntax (including word order), Romance morphology (including the loss of obsolete inflections, unless scribes had been explicitly told to use them), Romance semantics (in which old words acquired new meanings, as in many of the technical terms used by Christians) and Romance vocabulary (whether borrowed from elsewhere or internally created by normal derivational mechanisms) could all be, and indeed regularly were, represented in the normal writing systems before the time of the *Strasbourg Oaths*. There had been no problem there. For example, when the various forms of *ILLE* came to be used in speech as what we would think of as definite articles as well as the original pronouns, there was no need to change any of the written representations of the word; whether article or pronoun, clitic or not, the four letters of <ille>, and the other graphical forms, were taught and learnt and written and still thought of as correct. When *HABERE* came increasingly to be used as a future tense auxiliary, there was no corresponding need to change the written form; whether auxiliary or full verb, and however pronounced, *HABEO* could continue to be written as <habeo>. When grammatically reflexive *SE* increasingly found itself being used with non-reflexive non-agentive semantics, there was no need to change its written form. When *DE* changed meaning and function to rival and then overtake those of genitive inflections, it could continue to be written as <de>. And so on: as in all modern languages, so in late Latin, new grammatical and semantic features could easily be accommodated within existing spelling systems.

Thus features of morphology, syntax, semantics and vocabulary which we would now prefer to call 'Romance' rather than 'Latin' (if we have to choose one or the other) were attested in writing before the deliberate decisions were made to change the spelling systems as a whole. This has become obvious now in the evidence from seventh- and eighth-century Gaul, most notably as the result of the illuminating and groundbreaking research of Michel Banniard

and his colleagues at Toulouse, and of Marieke van Acker at Ghent (e.g., Banniard 1992a; 1993a; Verdo 2010; van Acker 2007). Francesco Sabatini's work in the 1960s on Romance elements in sixth-century Italian documentation was illuminating in itself (e.g., Sabatini 1968; see also 1983), and provided a light-bulb moment for António Emiliano's perspicacious and acute studies of similar, though rather later, evidence in texts from Portugal, both those that are ostensibly Latinate in their spelling system and those that can be described as being in old Portuguese – fittingly, given that the texts concerned are short and replete with proper names, this demarcation line is much disputed (Emiliano 1999; 2003b; Emiliano and Pedro 2004). I have prepared similar studies for Ibero-Romance, usually with reference to particular documents and individual idiolects, although also based on the concordances of the texts kept in the Cathedral Archive at León (Wright 1995a; 2003; 2004). The only conclusion that can be drawn is that, all over the Romance-speaking world, most of the texts of the centuries preceding the Carolingian reforms, apart from those of the most deliberately antiquarian writers such as Avitus of Vienne, can be exploited as valuable sources of evidence for Romanists, if used with care, and above all if not 'emended' into relative uselessness by myopic modern editors. The same caveat applies to later medieval textual evidence as well; see Fleischman (2000) on medieval French, and in general the justifiably strong views expressed by Lass (1997; 2004: 'an amended text is a falsehood', p. 31). Even relatively well-known late Latin and medieval Romance texts are likely still to contain linguistic data as yet imperfectly understood, which are in danger of being lost to scholarship if a modern editor decides not to let his readers see them; several late Latin works have been studied mainly, or only, for their knowledge of the ancients, or for their use of earlier sources, which is of little relevance to us, for it is precisely those aspects that have no source which are likely to be of interest to linguists.

The texts of those centuries unsurprisingly also attest relatively antique or obsolescent lexical and morphosyntactic features which Romanists would not reconstruct, nor like to think of, as belonging to early Romance, and there is scholarly discussion still as to how Romanists should react to these. Lexically, we can often believe what we see. Words attested in early medieval documentation, including those which were not going to survive into later Romance, can often be taken at face value as belonging to the lexicon of their writer, even if not of everybody else, given that individuals can in any event vary greatly with each other in the words they use and that the literate often have wider and older lexical resources than other speakers. Since most of us in any language community have a more extensive passive vocabulary than

active, words which are actively used by only a minority can still often be understood by a majority of those who hear a text read aloud (e.g., in Church offices). In view of the fact that most texts of that age were indeed read aloud, the non-literate were in no way cut off from the texts and written documentation which continued to form the basis of society in the early medieval centuries, and words which were readily intelligible, even though not widely used, need to be included in the lists of words which still existed in the language. The same probably applies to obsolescent morphology (such as -ARUM, -IBUS and -AMINI in late Latin); French speakers do not now use their preterite forms in speech, but they understand them when they hear them, so we can hardly say that the preterites do not exist at all any more; similarly, Spanish speakers never (or hardly ever) use their future subjunctives, unless they are lawyers working professionally, but will understand such forms as *fuere* if they hear them; so maybe those late Latin inflections should still be included in the list of the inflections of the age. Obsolescent vocabulary and inflections of this kind are still part of the language for as long as scribes are encouraged to use them, but will fade away from even passive competence if the technicalities of the written mode get updated and people no longer hear them regularly (as happened after the invention of the new Romance writing, which shows no sign of the continued existence any of these three inflections).

Written, and to some extent spoken, Latin were integrally involved in what have come to be labelled by modern historians as the 'Carolingian reforms', which the Carolingian scholars themselves referred to as *renovatio*, of around AD 800. Their educational component aimed, among many other things, to standardize linguistic usage along respectable antiquarian lines. This process can be seen as the invention of what we now tend to call 'medieval Latin' (Wright 1982: ch. 3). As a consequence of these reforms, the relation between written evidence and natural spoken usage became much more opaque, and the idea that this newly re-established educated Latin is a separate language from the normal vernacular (which we would call 'Romance') began to spread after that time; such a conceptual distinction is unlikely to have been made earlier. These reforms spread to most of the Romance-speaking world outside the Carolingian realms with the intellectual movements sometimes summarized with the phrase 'twelfth-century Renaissance'.

Thus the pre-reform textual sources are evidence of speech to a greater extent than those written in Latin later. Syntactically, there is little in the pre-reform textual evidence which cannot be accommodated in a plausible account of the grammar of the time (unless the author was deliberately imitating a model from the distant past). For example, we tend to think of one of the

features of Latin syntax as being the ‘fact’ that verbs came at the end of their sentence. This was perhaps statistically true during the Roman Empire, but in no sense obligatory even then, as several studies of Latin word order have shown (e.g., Pinkster 1991; most recently Spevak 2010; see also Ledgeway, volume 1, chapter 8). Roman grammarians didn’t tell their readers to do this, and it doesn’t seem to have formed a part of most scribal training. It is true that verbs in modern Romance, particularly in Spanish, often appear earlier in their sentence than they might have done in the first century AD, and earlier than they would in English, but that is also true of much of the documentary evidence from the pre-reform period, including in both eighth-century France and tenth-century Iberia. József Herman’s succinct account of *Vulgar Latin* demonstrates how such non-classical grammatical features in particular were attested in writing (Herman 2000; see also 1990). And our conclusion is the same as before: new features could appear in texts without the written mode needing to be updated at all. QUOD was written as *quod*, whether being used ‘classically’ as a relative or in the ‘Vulgar Latin’ fashion as the default complementizer.

These texts tend to manifest old and new together. For example, the same document can show both inflected genitives and prepositional phrases with *de*, with similar meaning; both accusative plus infinitive constructions and synonymous constructions with *quod*; or both active and apparently novel deponent morphology with active meaning. There are two possible reasons for this. In the case of the apparent deponents, given that there is no Romance evidence at all to allow us to reconstruct deponentization of originally active verbs as a feature of normal speech, the textual manifestation of unexpected inflections in *-ur* is likely to be an unusually obtrusive symptom of scribal training rather than of speech; that is, it seems that scribes were told that endings in *-ur* looked proper on the page, possibly without any indication from teacher to scribe that this added *-ur* might affect the meaning. Some texts thus acquire more, and increasingly unnecessary, morphological passives in successive copies and redactions, a phenomenon which supports this interpretation of that evidence in the sources as being the result of pedagogically inspired overenthusiasm.

This might also have applied to the two other examples adduced here of textual cohabitation, but more likely not, because later Romance evidence allows us to reconstruct the continuing presence of the older feature in the speech of the time. Cases other than the nominative and accusative had dropped from nouns and adjectives before the time of the earliest Romance texts (and even this distinction survived only in Gallo-Romance and Ræto-Romance), but oblique cases (other than vocatives and ablatives) had not disappeared from pronouns, so the genitive itself was not alien to the speakers: Italian *loro* and

French *leur*, from *ILLORUM*, for example, still survive, and although *leur* is now best analysed as an adjective or determiner meaning 'their', with its own plural form in *leurs* (but no feminine ***leure*), *loro* still seems to be a pronoun meaning 'of them', with no feminine ***lora* nor plural forms in ***lori* and ***lore*. Proper names, in fact, could still have genitive forms in old French as well as in the late Latin texts (e.g., in the phrase *ecclesia Sanctae Mariae*). And the accusative plus infinitive construction has not entirely gone from Romance either even now, particularly with verbs of perception (e.g., Italian *ci vede venire*, 'she sees us arrive'). So even though genitive inflections and accusative plus infinitive constructions may indeed have been recommended in class, and were thus on the list of those features which were commoner in writing than in speech, they were not as alien to speech as we might suppose if we felt impelled to categorize every feature as either 'Romance' or 'Latin' and thus to analyse such intratextual variation as evidence of a mixture of two languages. It is altogether misleading to see such data as evidence of a mixture of different languages ('Latin' and 'Romance'); theirs was one single language still at the time, which had competing variants as all literate languages do, some of which were newer in origin than others. The two were only going to be conceptually separated out at a later time when the general idea arose that there were two languages involved here rather than one, a conceptual recategorization which it seems can only have followed the development of new written Romance modes rather than preceded them. It is no coincidence that the earliest uses of the word *romanz/romance* (variously spelt) were all used to refer to the new modes of writing rather than to any kind of speech; the concept of 'Romance', semantically contrasted with 'Latin', had not existed before the new ways of writing had been deliberately invented (even though the modern historical linguist cheerfully applies the name to language states of earlier times).

Semantic change can be attested directly in texts of any kind. A surprising number of words have changed their meanings over the centuries, as any glance at an etymological dictionary will show. But speakers and writers, whatever their level of education, usually do not know if a word used to have a different meaning in the past, and would find it difficult to use the word with an obsolete meaning which it no longer has even if they knew about it. So a word which is still found in new texts with an old meaning probably attests the continuing currency of that meaning, at least in the register involved, particularly if the newer meaning would be incongruous. For example, Latin *CONLOCARE* ('place') eventually became Spanish *colgar* ('hang') and French *coucher* ('put to bed'); so the use of *conlocemur* in the seventh-century Visigothic prayer book, in a request that people be placed by God's right hand, seems to be

evidence that those semantic developments cannot yet have begun. Conversely, the regular use in the pre-1200 León Cathedral Archive documentation of the word written as *sedeat* as a passive auxiliary, with both the grammar and the semantics of Ibero-Romance (as would be later attested abundantly in the written form *sea*), meaning 'be' rather than the original 'be seated', suggests that what is being attested in this evidence is simply the Romance grammar and semantics of the writer's time and place, regardless of the word's being dressed in the normal traditional spelling (Wright 2004).

Such a combination as we often see attested in the evidential sources from the late Latin pre-reform documentation, of contemporary semantics, vocabulary, morphology and grammar with traditionally correct orthography (however unlike a putative phonetic transcription this is in eighth-century Gaul, twenty-first-century Paris, etc.) is exactly what literate speaker/writers of English, French, Spanish, Portuguese, Italian, Catalan, Romanian, Galician, etc., do every day of the week. It is what one would naturally expect in any language community in which the spoken language changes unavoidably and the graphical forms taught to apprentice scribes remain the same; which may well be all of them. So it is hard to see why this perspective is sometimes thought to be controversial when applied to early Romance. And in general, the moral is clear: reconstruction is valuable, but written evidence can and should be used as sources for our analysis of speech if we continually bear the context of production in mind. There is no text without context, as the sociolinguists say. Diachronically, this kind of analysis is the role of sociophilology, which aims to combine sociolinguistic insight with philological concentration.

If no direct phonetic evidence survives the coming centuries, as it may not, the historical linguist of a thousand years hence who looks back at the Romance languages of the early twenty-first century will have similar problems concerning textual evidence to those which Romanists have now. In particular, he or she will feel as uncertain as the contemporary Romanists do about the nature of the evidence attested by the sources, because written forms have no direct connection with spoken forms even in languages thought to be phonographic, such as modern Spanish or Italian. In the year 3000, it will, for example, take careful philological analysis to realize that there is no [g] in twenty-first-century Italian *sbaglio* any more than there is in *right*, no [v] and no [ll] in Spanish *valle*, no [x] nor [ks] nor even [s] at the end of French *heureux*, and that all the many Portuguese words ending in *-os* are, at least in Europe and Rio de Janeiro, pronounced with [-uʃ] rather than [-os]. Conversely, he or she may not realize that the French preterite tenses are both now already absent from speech and still intelligible even so when heard in a written text

read aloud. He or she may even be tempted to deduce that forms of words found in twenty-first-century text messages represent a different phonetic entity, or even a different language, from the more conventional written forms of the same word; for they don't, and in this respect text messages are roughly similar to the Riojan glosses, both being new ways of writing the same language as before, rather than as yet attesting the birth of some new language. Metalinguistic comments are still as much based on what writers think people ought to say as they always have been, although given such works as this *Cambridge History*, such data from 2013 might perhaps be easier to follow with certainty in 3000 than Velius Longus is now. It remains probable that if we miraculously overheard a Romance conversation of the year 1199, or a performance of the multi-Romance *descort* written in that year by the Provençal *trobador* Raimbaut de Vaqueyras, we would hear a number of details which we didn't expect; but even Romance historical linguistics is a progressive science, despite the dead ends which have been put in our way, and we probably have a better idea of those details now than our forebears did a hundred years ago.

Koinés and scriptae

JOHANNES KABATEK

I. Introduction

This chapter discusses processes of linguistic convergence in medieval Romance: the development of supra-regional spoken varieties (*koinés*) on the one hand and the emergence of supra-individual orthographic conventions (*scriptae*) on the other. The term *koiné* is sometimes applied to both written and spoken varieties, whereas the term *scriptae* frequently refers only to non-literary orthographic conventions. For reasons of clarity, I shall distinguish the two terms by reference to their medium, using *koiné* for all supra-regional *spoken* varieties and *scripta* for all supra-regional *written* varieties, whilst recognizing their mutual relationship. The chapter consists of two main parts, which are in turn subdivided. In this first part (§1), I shall address some fundamental theoretical and methodological issues concerning the general problem of koineization and the evolution of *scriptae*. The second part (§§2–3) will begin by distinguishing different periods of development and go on to describe convergence processes in the Romània of the Middle Ages (up to the Renaissance). Each section will include observations on the current state of research and suggestions for further investigation.

In order to provide an adequate description of the processes of linguistic convergence in the Middle Ages, it is essential to consider both internal and external historical factors. This requires extensive consideration of the role of institutions and the existence of centres of gravity, and analysis of the prestige values of the time within a comprehensive model of linguistic variation, as has become increasingly widespread in medieval research in the last few decades. Increasing use of corpus-linguistic methods and the possibility of working with extensive data-sources have enabled historical hypotheses to be refined; however, particularly in recent years, they have also led to an unfortunate neglect of theoretical understanding of linguistic variation, since sheer mass of data is sometimes seen as a substitute for philological differentiation. It is true

that large quantities of data often enable us to construct rich and subtle representations of long-term developments, but this should normally be the starting-point rather than the end-point of an analysis. For instance, the identification of a specific process of linguistic convergence based on large quantities of diachronic data should not lure us into seeing a metaphorical ‘invisible hand’ (Keller 1990) as responsible for processes of language change. Rather, the identification of this process should lead to a detailed analysis of the sub-processes (*innovation, adoption, diffusion, selection, mutation*; cf. Coseriu 1983) which lie behind the change, in order to achieve a comprehensive historical description or ‘explanation’.

1.1 Terminological preliminaries

Even if one of the basic assumptions in historical linguistics is that we can ‘use the present to explain the past’ (Labov 1975), taking general linguistic knowledge obtained in our current linguistic situation and applying it to the Middle Ages does not necessarily produce valid results. For this reason, Medieval Studies has tended to be wary of embracing developments in general and synchronic linguistics, and has seen itself as a quite separate discipline. In the Middle Ages, ‘everything is different’, because of the lack of spoken evidence and the great divergences between manuscript culture and the culture of the printed book. Nevertheless, within several frameworks – such as New Philology (Wenzel 1990; Gleßgen and Lebsanft 1997), historical sociolinguistics (Romaine 1982; Gimeno Menéndez 1995), historical pragmatics (Schlieben-Lange 1983), historical dialectometry (Goebl 2000; 2007), discourse tradition research (Kabatek 2005b; 2008), to mention just a few – attempts are increasingly being made to apply recent linguistic methodology to the Middle Ages. Manuscripts constitute our sole record for this period and we have only indirect access to its languages, so, if we are to establish a method enabling us to reconstruct what is no longer directly accessible, our starting-point must be our general linguistic knowledge.

A first distinction that must be drawn here is that between *text*, *text tradition* – or, as we prefer, *discourse tradition* – and *language*. When we study the text of any medieval manuscript, we can attempt to deduce the grammar of that text and describe its lexicon. Subsequently, we might compare the language of that text with that of other texts (from other areas or periods) and assess its representativity as regards a specific *état de langue*. Nonetheless, a differentiated study of language should reject this method, however tempting its simplicity, and opt instead for a more complex approach.

A text is a concrete individual utterance, whereas a language is a supra-individual system of signs. These are two distinct levels, and some models of grammar consider the one level to be directly derived from the other. The study of linguistic variation, however, shows us that in general an individual masters not one single linguistic variety, but several, albeit to varying extents, and that in the creation of individual texts, multiple linguistic varieties may be merged. A text *can* in fact be based on a single linguistic variety, but this does not necessarily *have* to be the case. In everyday life, it is completely normal in many speech communities for the spoken language of individuals to be characterized by elements of varying geographical origin or by a blend of dialect elements mixed to a greater or lesser degree with the standard language (see Auer *et al.* 2005). When dealing with the written language, however, we rather assume uniformity, a notion that plays a striking role in the conception of language in many branches of linguistics.¹ The study of language since the invention of the printing press, the purism of the Academies and Jacobin uniformism, which has greatly influenced Western conceptions of language since the French Revolution (cf. Schlieben-Lange 1996), have all served to obscure the true heterogeneity of language, which, when recognized at all, is often considered to be an exception to the rule of a uniform, standardized language.

In the Middle Ages, the monolingualization characteristic of modern Western linguistic culture had not yet taken place or was only just beginning. In consequence, we must adopt a dynamic conception of language and accept the possibility that several languages and/or varieties are present within a single text. In principle, for medieval languages, as for languages in general, we assume three dimensions of possible variation (Coseriu 1980): varieties according to geographical areas (*diatopic* varieties), varieties depending on social groups (*diastratic* varieties) and varieties pertaining to the style of language (*diaphasic* varieties; for *diamesic* varieties, see below), which can occur even within the speech of a single speaker in a constant group and in a constant place. In diachronic description, one must actually describe the

¹ Owing to the variation found in manuscripts, in medieval philology a variation-orientated perspective has long been common in certain linguistic traditions (cf. for example Menéndez Pidal 1926; Brunot 1905). It would be valuable for the field of Medieval Studies if a critical synthesis could be established between the older tradition of the study of variation on the one hand, and more recent medieval variationist findings based on the adoption of contemporary linguistic thought on the other hand, which would serve as a corrective to the linguistic monolithism derived from the focus on written language (cf. for example Cerquiglini 1989 and, for several proposals along these lines, Hafner and Oesterreicher 2007).

diachrony of a three-dimensional construction, the ‘architecture’ (Flydal 1951) or ‘diasystem’ (Weinreich 1954), of the ‘historical language’ (Coseriu 1980). The notion of a specific, more or less homogeneous ‘idiolect’ expressed in the texts of an individual is misleading; rather, one must assume a multiple competence that encompasses knowledge of different varieties, each one of which may be voiced in a single text. This remains the case even in a specific, stable constellation of communication. The crucial tension which appears to characterize the texts of an individual is the antagonism between the mother tongue(s) on the one hand and varieties and languages learned later in life on the other (cf. Miestamo *et al.* 2009), which poses the question of how to deploy the multiple competence in a specific situation, depending on such factors as the assessment of the interlocutor, the content, and the prestige of the linguistic forms involved. The greater the number of varieties involved, the more complex this tension becomes. A further problem arises from the fact that texts themselves are not only individual utterances but also part of traditional settings, which means that writing (and speaking) involves knowledge of specific textual traditions, so-called *discourse traditions* (cf. Koch 1997; Oesterreicher 1997; Kabatek 2005a; 2005b; 2005c); the latter may influence the choice of elements used in a text – that is, not only the ‘textual’ characteristics selected, such as the specific text form, but also lexical or grammatical elements or even a particular combination of languages (cf. Kabatek 2008).

When analysing an individual’s texts or utterances, one must identify two main characteristics of the varieties, apart from the possible influence of a certain discourse tradition: first, the variety the individual is seeking to employ, and second the interference of other varieties forming part of the individual’s competence, which also influence the utterance/text. Except for completely hybrid texts resulting from aleatory language mixture, a text is generally at any given moment orientated towards a particular language or variety; a speaker or writer is always seeking to realize a specific linguistic system. However, this orientation can change within the text as a result of code-switching. The identification of code-switching in the case of very closely related medieval Romance systems can be an extremely difficult task, so that in the analysis of medieval texts it might sometimes be just as important to discover which variety is intended to be realized in a certain passage as to examine the specific linguistic properties of the passage under consideration. Sometimes code-switching is clearly identifiable and might even appear together with metalinguistic comments – for instance, when Romance passages are incorporated into medieval Latin texts (such as charters and chronicles) as literal quotations, as in the case of the *Strasbourg Oaths* or the

Placiti cassinesi (see §§3.1 and 3.3.1 below). But even within a text segment orientated towards a single specific language, elements of various languages may appear, when varieties copresent in the competence of the speaker/writer interfere with the target variety. Here we must distinguish four types of interference (Coseriu 1977c; Kabatek 1996). The first type is the one most frequently discussed, namely, the overt appearance of elements of a different language/variety from that towards which the text is orientated – for example, Occitan elements in a northern French text, or Latin elements in a Romance text, or Romance elements in a medieval Latin text. This type of interference can be called *transposition interference*, and it results in the presence of *positive*, namely, effectively identifiable ‘foreign’ elements in the text. The opposite is the case with the second and third type of interference, which are frequently neglected in the study of linguistic variation. Their results cannot be observed directly in the form of foreign elements in a text. The second type of interference, which we may subsume under the term ‘negative’ interference (following Coseriu 1977c), consists in a preference for what has been found to be concordant between two varieties. This type of interference can be referred to as *convergence interference*, or simply *convergence*. The third type is complementary to the second and is based on a preference for diverging elements (*divergence interference*, *divergence*). Neither of these negative types of interference leads to ‘mistakes’ or to overtly identifiable foreign elements on the surface of the text, but they do alter the frequency with which certain forms are used. Both types can be found more frequently the more closely the varieties copresent in the speaker’s competence are related – that is, the more convergent elements exist between them – and they are based on a mostly implicit analysis of elements of different varieties between which the speaker/writer observes analogies as well as differences (for an application of these concepts in a medieval context, see, amongst others, Bello Rivas 1998). Finally, a fourth type of interference can be observed, which is likewise based on the contrastive analysis of two languages or varieties. The result, however, lies beyond the traditions of both languages, as the analysis leads to the application of transfer-rules in cases where both varieties actually coincide. This type is traditionally called *hypercorrection* and serves as an important indicator for the reconstruction of the language of earlier stages.

1.2 Koinés and koineization

The term *koiné* (for general discussion, see Siegel 1985; 1993; Cardona 1990; Grübl 2010; 2011) obviously derives from a very concrete old Greek language situation, and its application to other linguistic settings has led to a certain

terminological ambiguity, as Siegel (1985) states when he says that of ‘all the imprecise terms used in sociolinguistics, “koine” may win the prize for the widest variety of interpretations’. There are mainly two dimensions that are discussed in this context: on the one hand, the relationship between oral and written supra-regionality and, on the other hand, the difference between conscious (or even legally determined) language planning and ‘implicit’ convergence processes without overtly visible metalinguistic intervention. As mentioned above, we will reserve the terms *koiné* for (only indirectly accessible) *spoken* supra-regional varieties and *koineization* for their emergence, whilst recognizing their mutual relationship with convergence processes in the written *scriptae*. In sociolinguistics, other terms such as *standardization* (Haugen 1966; see also Lodge 2011) or *Ausbau* (Kloss 1987) are being used for modern convergence processes, and attempts have been made to apply them to medieval situations (see, for instance, Kabatek 2005a; Koch and Oesterreicher 2008; Selig 2008). There are obviously no instances of conscious, institutional language planning, such as are found in modern times, in the Middle Ages; however, if we look at Haugen’s criteria for standardization, distinguishing the processes of *selection* of a linguistic variety, *elaboration* of functions (roughly corresponding to Kloss’s *Ausbau*), *codification* of a certain norm and *acceptance* of the standard, we can see that they are to some degree at least implicitly present in medieval convergence processes, as well.

If we assume that all the aforementioned types of interference and code-switching may be present in a medieval text, then reconstructing medieval language areas or convergence or koineization processes may seem a virtually impossible enterprise. In addition, there remains the problem of the written language, which will be addressed below. On the other hand, variation is precisely what enables us to situate a text more accurately. This is a well-known fact in traditional philology, and the notion that an author’s origin may be determined on the basis of foreign elements or instances of hypercorrection is not a novel one (cf., for example, Baldinger 1958). It does seem desirable, however, for the achievements of modern-day variationist linguistics to be applied to medieval studies to a greater extent than at present. Amongst other things, it would be interesting to investigate in greater depth the general circumstances under which certain types and combinations of interference generally appear, so that we can search for comparable constellations in medieval texts. Factors such as the degree of elaboration of a text, the closeness of the interfering varieties and the various structuring levels of the language would have to be taken into account. ‘Phonetic’ and ‘graphic’ hypercorrections presumably have a different status from syntactic or lexical

ones, generally appearing alongside negative interferences leading to shifts in frequency, which means that they are part of the convergence and divergence processes with which koineization and *scriptae* research is concerned (cf. Holtus, Körner and Völker 2001).

A further methodological problem arises from these considerations: how is the difference between variety and interference to be dealt with methodologically, if all we have at our disposal are texts (namely, utterances), whilst the languages and varieties are undergoing dynamic processes of change? One could cite the familiar criticism of the unobservable nature of linguistic change from a synchronic perspective. However, it is by no means the case that all that has come down to us from the Middle Ages is an impenetrable linguistic chaos of variation: we are aware – at least to a large extent – of the Latin basis of the Romance languages. We have access to the present-day varieties which have emerged from the medieval ones and which in part remain surprisingly similar to them (Goebel 2008), and, alongside variation, we can also recognize substantial areas of stability in the languages of the Middle Ages, as well as clearly identifiable phenomena which indicate their dynamic.

The reconstruction of the medieval Romance *scriptae* and *koinés* thus implies their *recontextualization* within the overall architecture of the medieval languages and varieties (Oesterreicher 2001a; Koch 2006). In recent decades, various attempts have been made to add a ‘diamesic’ component to this architecture (cf. Mioni 1983), differentiating spoken and written varieties as well as establishing a universal continuum between ‘immediacy’ and ‘distance’ (Koch and Oesterreicher 1985; 2011). Other authors have pointed out that this expansion of the model involves a level which is in fact distinct from that of linguistic variation, since the relationship between spoken and written language is purely one of medium, characterized by the possibility that basically any variety may be expressed in written or spoken form. On the other hand, it is also a qualitative relationship, stemming from the fact that only certain varieties are actually written, whereas others are not. Moreover, specific written discourse traditions may emerge that are shaped by the particular possibilities which the written medium offers, and these may lead to the creation of specific styles linked to the written language (Kabatek 2000b).

It is precisely the interplay between language and speech, between the system and the creative behaviour of the speaker, that results in the diasystem of the historical language being not rigid, but rather dynamic. This interplay also means that individual processes of convergence and divergence which we observe in texts can lead to the convergence or divergence of languages at an

abstract level. Individual convergence is a correlate of dialogue and of power and prestige relationships, whereas linguistic convergence is a correlate of communicative networks with the corresponding centres of gravity and their power and prestige, which determine the individual dialogue culture. This means that in order to examine linguistic convergence, one must consider the operation of individual processes on the one hand, and the social communication structures within which dialogical convergence processes may occur on the other (cf. Auer *et al.* 2005). These structures are outward correlates of communities organized in various cultural institutions, from the family, through the monastery or castle, to the village or town or other political or religious units, where the complexity of the respective institution may go hand in hand with that of its internal organization.

Our observations give rise to the following challenges for the analysis and characterization of koineization processes in the Middle Ages. First, the texts must be analysed with reference to the underlying languages and varieties. Second – and this is particularly important for the question of linguistic dynamics – the dynamic inherent in the texts must be related to the architecture of the language. A particular text cannot simply be located somewhere at a fixed place in the diasystem of a language; rather – and especially in dynamic situations – a text often originates from a specific language (or variety) and simultaneously heads towards a specific language (or variety). We could call this phenomenon the ‘vectoriality’ of a text or utterance. In this context, it is of vital importance to know which portions of a text can be attributed to the language acquired earlier on and which to the one learned later. For example, a strikingly large number of Castilian words appear in the old Portuguese *Foros de Castelo Rodrigo* (cf. Cintra 1959). One might therefore assume that the author of the text was a Castilian speaker. But it could equally well be the case that the Castilianisms in question are introduced as a result of the prestige of Castilian, or even that they are Castilian loanwords which were already perceived as normal in the Portuguese legal terminology of the time, thus indicating a general process of convergence. It is sometimes very difficult to judge whether a particular element in a text is an individual interference, an expression of individual tendency towards a prestige variety, or evidence of a new linguistic tradition, in which it figures as a loan element. The decision as to which of these we are dealing with must be the result of an interpretative reconstruction process, in which language-external and language-internal information is considered, in order to yield the most probable language-historical interpretation. This in turn leads to a third challenge: that of obtaining extensive knowledge of the available language-internal data of the

corresponding language area; and a fourth: obtaining extensive historical knowledge of cultural institutions and social developments. Metalinguistic comments and the names of languages take on an intermediate position between external and internal data. Language names (cf. Kabatek and Schlieben-Lange 2000, and Wright, this volume, chapter 3) in particular are important indicators of convergence or divergence processes, as they can delineate an established linguistic area as well as create or consolidate linguistic boundaries; they can be derived from, or motivated by, linguistic realities (such as particular linguistic characteristics) or non-linguistic ones (such as political boundaries).

1.3 Scripta

Since all of the steps necessary in koineization research may lead to ever-increasing interpretative uncertainty, and as the information at our disposal must ultimately be derived in its entirety from the mass of manuscripts and from comparative historical reconstruction, so-called *scripta* research has become established as a sub-discipline of Medieval Studies, above all in Gallo-Romance studies. This field is for the most part limited to the study of written phenomena, for which an increasingly refined method was developed in the course of the twentieth century. The term *scripta* goes back to Remacle (1948); it generally designates a particular (and in his narrow approach a non-literary) medieval writing tradition. Between the 1940s and 1960s, *scripta* research was developed mainly by Carl Theodor Gossen (cf. Gossen 1967), who also coined the term *scriptology*. The field of *scripta* research is critical of an approach which naively equates regional written language with regional dialect, and it thus consistently rejects the notion, attributed to Gaston Raynaud, that regional, dated legal documents provide direct insight into medieval dialects.² In contrast, descriptions of medieval ‘writing landscapes’ (‘Schreiblandschaften’, Gossen 1968) have been called for,³ established on the basis of data found in the medieval *Chartes*, which are classified and evaluated

2 ‘So *chartes* present us with the absolute truth of the vulgar language, and are by far the most valuable sources for the study of the dialects’ (‘Les chartes . . . offrent donc la langue vulgaire dans toute sa vérité, et sont de beaucoup les sources les plus précieuses pour l’étude des dialectes’) (Raynaud 1876:54).

3 It is repeatedly stressed that there is a certain, but by no means direct, relationship between regional dialect and *scripta*: ‘The regional written languages of northern France enable us to glimpse, to very varying extents, what medieval dialects were, but they are by no means identical with these dialects’ (‘Les langues écrites régionales de la France du Nord laissent entrevoir, à des degrés très différents, les dialectes du moyen âge, mais elles ne sont nullement identiques avec ces dialectes.’) (Gossen 1968:4).

according to diatopic and diachronic criteria. These reveal the heterogeneity of the *scripta*, which is a 'hybrid and composite continuum' ('continuum hybride et composite') (Goebel 1975:147) in which a single scribe may use various forms to spell the same word⁴ – contrary to the nineteenth-century assumption that this is the case only for the literary language, chiefly as a result of changes made by copyists. Thus, parallel documents written by different scribes at the same time, as well as different documents written by the same scribe, are particularly informative for *scripta* research. There are tendencies both towards a predominantly quantitative analysis and towards a detailed, qualitative philological analysis. In quantitative analysis, a historical linguistic geography has taken root in research on French, akin to Ramón Menéndez Pidal's earlier work on Spanish: linguistic data are copied on to maps, from which the medieval writing landscape can be derived. A significant spur to this line of research has been computer-aided data analysis, which was conducted at length for the first time by Hans Goebel (1970) for medieval Normandy, and allowed immense quantities of data to be analysed. In the 1980s, scholars, most notably Anthonij Dees (1980; 1987), extended this quantitative, statistical approach to other areas, examining the entire northern French area and also including the writing landscape of literary texts (Dees 1987). Despite polemical exchanges between individual researchers (Gossen 1982; Dees 1987, XIV), the field of *scripta* research has become a firmly established discipline with a clearly defined method, particularly in the case of the northern French area. Its method involves the following steps:

- selection of an area to study;
- selection of a corpus to analyse as well as a particular period;
- selection of several 'scriptorial features' ('traits scripturaires') considered to be relevant;
- statistical analysis of the corresponding relevant characteristics and cartographic representation;
- diachronic and historical interpretation of the statistical evaluation.

4 This was introduced by Gossen as an argument against the possibility of a direct relationship between the written and the spoken language: 'If so-called graphemes really correspond to phonological reality, then we ought to ask why one and the same scribe should use several graphies for the same sound in the same word in the same document. It is certain that he will have had only one pronunciation of the word in question.' ('Entsprächen die genannten Grapheme wirklich alle lautlichen Realitäten, so müsste man sich fragen, wieso ein und derselbe Schreiber in derselben Urkunde für denselben Laut desselben Wortes mehrere Graphien verwendet. Er besaß doch sicher für das betreffende Wort nur eine Aussprache!') (Gossen 1967:15).

The selection of a particular area is made according to historical and political criteria or with regard to certain historical linguistic areas. It is wise to outline diachronic areas of convergence or divergence, as certain processes taking place in these areas are to be the focus of attention.

When selecting a corpus, the beginning of the Romance period is generally set as the *terminus a quo*, and the attainment of a more or less unified, supra-regional orthography is set as the *terminus ad quem*. In the case of French, the starting-point is generally the beginning of extensive Romance document production in the thirteenth century, and the end-point the decree of Villers-Cotterêts (1539), in which the *languaige maternel francoys* is stipulated as the exclusive written language for legal documents (see also Sanson, this volume, chapter 7). The latter boundary is not disputed and is connected to the unifying tendencies of humanism and to the development of the printing press, which took place in other language areas besides France. Studies similar to those of *scripta* research for later periods would only make sense if different types of texts (such as private correspondence) were selected. At the other end of the period, the restriction to Romance texts should to some extent be waived, inasmuch as writing landscapes are already beginning to emerge within the Latin tradition, especially in the case of vernacular names for places and people which already appear in their vernacular form in the Latin texts. These may be considered a testing ground ('Versuchsfeld', Goebl 1970:119) for the emergence of written Romance, even if the statistical analysis of elements occurring only sporadically is not possible to the same extent as in texts clearly characterized as Romance. When selecting the corpus, attention must be paid above all to the reliability of the transcriptions or editions, since much of the material transcribed for historical documentation rather than for philological purposes proves to be unreliable.⁵ On the one hand, large collections of documents are important, yet on the other, the significance of smaller collections and individual 'exceptions' has also been stressed (Gossen 1979:265).

The determination of the 'scriptorial features' must be carried out on the basis of an intensive comparison of documents which are as diverse as possible; historical linguistic information beyond what is in the corpus must also be taken into account in order to select characteristics which are likely to display variation in the area in question and in order to exclude homogeneous characteristics. A basic principle is that, generally speaking, there is no single

⁵ An overview of Romance documents up to the end of the thirteenth century can be found in Frank and Hartmann (1997).

trait that is characteristic of a *scripta*, and a quantitative variationist analysis (in the sense of Labov) cannot be based on a particular isolated element. Rather, a *scripta* is characterized by a ‘particular combination’ (‘particolare combinazione’, an expression originally coined by Ascoli 1876) of written forms.

The statistical analysis initially takes place on the basis of the observation of a ‘habitual frequency’ of certain forms within the whole corpus. Next, deviations are measured within individual sections of the corpus, corresponding to divisions made on geographical and chronological grounds. These deviations are transferred to maps which take the aforementioned diatopic and diachronic differentiation into account. It has also proved fruitful to take account of the distinction between original documents and copies, as advocated by Goebel 1995 (cf. also Morala 2002), in contrast to Dees, as the differences between these types of document reveal a certain *direction* in the evolution of a *scripta* – similar to what has been claimed above for hypercorrect forms. It is often the case that a high degree of regionality may be observed in the original texts and a higher degree of supra-regionality in copies intended to be archived in the scriptorium.

The question of considering further parameters of variation in *scripta* analysis has been the subject of more recent discussions (cf. Goebel 1995; Völker 2001; 2003; Videsott 2009). However, the inclusion of diastatic and diaphasic criteria does not appear to be unproblematic, given that the variation due to these factors in the documents might be marginal, with the consequent danger that any interpretation proposed will need to be hedged about with reservations.

1.4. *Scripta and koiné*

The last of the above-mentioned steps, namely, the diachronic and historical interpretation of the data, in fact goes beyond mere *scripta* research, and there is no fully developed method for it, so that in many studies intuitive judgments are applied. This step, however, is the truly relevant one for historical linguistics. In order to arrive at an adequate interpretation, historical data and information on later developments must be combined. As far as the emergent language areas in the Middle Ages are concerned, we may observe that the computationally synthesized data of twentieth-century linguistic atlases present surprising parallels with the synthesized data of medieval *scripta* research. This is why Goebel considers data from the *ALF* alongside the scriptological data in his studies (Goebel 2000; 2002; 2003; 2006; 2008). There is no doubt that medieval documents and medieval dialects are to some extent related, and indeed this relationship has provided the basis for a

document-based medieval ‘dialectology’, although it is increasingly obscured as dedialectalization unfolds throughout the course of history. The ‘scriptorial features’ are in addition always symptomatic of specific social and cultural constellations which, in turn, may correspond to linguistic features with varying degrees of probability. However, we encounter various methodological obstacles in reconstructing the relationship to the spoken language: in manuscripts, the spelling, particularly in the case of questionable elements, is frequently not uniform. The different written forms may have corresponded either to a single phonetic form or to different coexisting ones, owing to the arbitrary relationship between sounds and graphemes, and in the final analysis, a stable orthography does not necessarily have to correspond to a spoken reality, but could instead result from a purely written convention. The rejection out of hand of any relationship between written and spoken language on the one hand, and on the other the claim that dialects are directly mirrored in medieval documents would be equally exaggerated and extreme positions (cf. Dees 1985; Remacle 1992). A relationship is probable, but not necessary, since the writers may originate from other areas (Monfrin 1968) or be orientated towards other varieties. Restricting the basis on which manuscripts are localized to purely extra-linguistic factors, as has been called for since Carolus-Barré (1964), can only offer hints, given that the place in which a document was produced by no means necessarily determines its language. For this reason, *all* available factors must be considered (see also Wright 2001): those that can be derived from the language of the manuscript, from its content, from its outer form (Frank 1994), from palaeographic analysis and, finally, all the external circumstances that might contribute to an adequate interpretation. However, whilst the analyses of *scripta* research, which are restricted to the written language, can deduce concrete objective findings from the facts present in the underlying documents, a ‘medieval dialectology’ will always remain hypothetical and at best be able to indicate certain probabilities based on indirect data. These probabilities, alongside the medieval and modern writing landscapes, constitute the third point of reference for calculating the unknown side of a triangle. An ‘objective’ method in this procedure can only be approximate and probabilistic, but the degree of probability could indeed be substantiated by a multifactorial statistical analysis. However, one must always bear in mind that statistics can never examine or ‘explain’ what actually happened, but merely quantify the sum of single events. Nonetheless, they can provide a useful framework and starting-point for the detailed philological analysis and interpretation of individual texts.

2. Periods of convergence in medieval Romance

If we consider the development of koineization tendencies and writing traditions in medieval Romance languages as a whole, we can observe certain parallels between the different areas, which enable a classification into different phases to be made (cf. Koch 1988; Krefeld 1988). However, these parallels take on quite different shapes in the individual areas and are sometimes not chronologically identical.

The first phase could be called 'prehistoric' Romance or proto-Romance, lasting up to the appearance of the first written documents of clearly Romance form. At this point, many of the typically Romance characteristics have presumably already been established in the spoken dialects as opposed to written Latin, but Latin, as the relatively uniform written *Dachsprache* ('roof-language', Kloss 1987), conceals these differences from later examination, rendering the postulation of certain convergences or divergences between Romance varieties rather speculative. Though early Romance features are repeatedly found in Latin texts, there are no clearly Romance texts yet. Nonetheless, at least in parts of the Romània, it is possible to reconstruct Romance language areas even for this initial phase, since, first, later linguistic developments allow us to draw inferences, and second, at least a rough division of areas is mirrored in the Latin texts of the third to ninth centuries (cf. Bonfante 1999; Lausberg 1956–62:39f.; Kontzi 1982; Herman 1990; Lüdtke 2005). The sources for this phase are Latin and, in part, Greek texts which enable deductions to be made concerning Romance articulation and certain syntactic phenomena. Since the 1960s (cf. Sabatini 1968), research in this area has increasingly tended to consider the *scripta latina rustica* in connection with the clearly Romance *scriptae*, since it can be observed that more and more Romance elements figure in certain Latin texts (mainly in less formulaic parts of documents) in various areas from roughly the sixth century onwards. These Romance elements in Latin lay the foundations for the emergence of the first Romance texts in the second phase. On all these issues, see also Banniard (this volume, chapter 2).

This second phase could be called that of 'sporadic Romance', in which admittedly rare, but nonetheless available, written evidence and metalinguistic comments indicate some awareness of 'Romanceness'. One might object that the sporadic appearances of Romance texts starting in the ninth century are but isolated instances and that in reality we are dealing with a continuation of the first phase. But we must presuppose a previously unattested awareness of a distinction between Latin and Romance when we examine the

composition of the *Strasbourg Oaths*, the *Eulalia Sequence*, the *Placiti cassinesi*, the *Glosas Emilianenses* or other supposedly isolated Romance texts – short passages written in Romance that appear from the ninth century onwards, mostly in Latin contexts (for further comments, see below) – even if a more or less lengthy period of transition and considerable differences between the various regions and centres must be assumed. This consciousness seems to be based on an apparently paradoxical development, which occurred repeatedly in a similar way throughout the history of the Romance languages: it is generally agreed that a major cause of the consciousness of ‘Romanceness’ was the rekoineization of Latin, namely, the reform of the pronunciation of written Latin, which spread in several waves in the Romance areas at different times. The goal of this reform was actually to achieve uniformity, but as a side-effect, an awareness of the gulf between the claim of unity and the heterogeneous linguistic reality was created. In ninth-century France, it was the Carolingian *Correctio*, initiated by Irish and English monks and with Alcuin of York in a pivotal role, that modified the spelling and pronunciation of Latin texts (Wright 1982; Lüdtke 2005:644–54). In the eleventh century, the Cluniac reform of Latin reached the Iberian Peninsula, whereas in Italy the effects of the reforms were weaker, maybe also due to the smaller distance of the vernacular from Latin (cf. Raible 1993:236; see also Banniard, this volume, chapter 2).

The first known texts resulting from the differentiation between Latin and Romance (Banniard 2006c) display an array of common characteristics (Lüdtke 1964; 2005; Renzi 1985:239; Koch 1993; Selig 2001): they are testimonies to spoken language marked as vernacular for reasons of authenticity, for instance in oaths, records, notes in records (cf. Wunderli 1965; Sabatini 1964:149f.; Petrucci and Romeo 1992:116), vow formulae, lists, commentaries, glosses (Wright 1982; Quilis Merín 1999), or religious texts intended to propagate Christian thought, as the use of writing was generally tied to the monopoly of the clergy. Larger works of literature (such as Occitan *troubador* poetry or early epics such as the *Chanson de Roland* or the Castilian *Poema de mio Cid*) are sometimes attributed to this period of ‘sporadic Romance’ as well, although one must bear in mind that they were passed on in manuscripts that must actually be attributed to the next phase. It is striking that, despite early evidence of written Romance, the ‘sporadic Romance’ phase lasted for a relatively long time: in spite of an attestable consciousness of distinction – or perhaps precisely for that reason – the diglossic situation, in which the written language was almost exclusively Latin, remained stable for several centuries. In the second phase, we can identify certain historical events

concerning koineization tendencies that presumably led to instances of convergence, but we have no direct written evidence. Thus we may assume that certain cities of growing importance at this time (e.g., Pavia, Bologna, Paris, Toulouse, Montpellier, Barcelona, Burgos, Toledo, Lisbon) became centres of development for urban varieties that could subsequently spread (to varying degrees) to their environs. As far as the writing of this sporadic evidence is concerned, it corresponds to a more or less spontaneous attempt to find adequate Latin graphemes to express certain phonetic realities with no established vernacular tradition ('Verschriftung', in the sense of Oesterreicher 1993). On the other hand, spelling traditions for Romance elements had already developed within Latin during the first phase (for example, for the representation of proper names), with certain tendencies towards areas of convergence even in the earliest Romance texts (Sabatini 1968; Hilty 1973).

The appearance of the first series of texts marks the beginning of the third phase: at first it is in legal texts – feudal oaths and other legal documents – that the vernacular appears in the less formulaic parts, mirroring the Romance reading practice of these parts and subsequently spreading to the other sections of the documents, where Latin formulae are more resistant to the vernacular. This happens in the south of France from the beginning of the twelfth century onwards. In the following century, the same process occurs in other Romance areas, where it actually takes place rather rapidly in certain centres after a long period of diglossia, so that one cannot really speak of a gradual development: rather, certain underlying external factors must have initiated this process. These factors may be linked to those salient for the fourth phase, which will be discussed below. Certain religious orders (Knights Templar, Benedictines) and their centres seem to have played an important role in this process by promoting the spread of the Romance writing tradition. The emergence of Romance writing must also be seen in the context of a general, predominantly Latin 'explosion' of text production from the end of the twelfth century onwards (Raible 1993), of which Romance texts are actually only a by-product. A factor of central importance for the emergence of certain Romance writing traditions is their pragmatic context, where Latin-educated scribes or readers transmit information to illiterate speakers or listeners by reading aloud and writing records of legal acts with Romance passages (Lüdtkke 1964; Wunderli 1965; Sabatini 1968; Selig 1995; 2001). The scribes, who were closely tied to the monasteries, developed individual and local traditions, which in part became supra-regional *scriptae*. In certain areas, particularly in northern France, a tendency towards establishing supra-regional language areas has been observed from the first series of

texts onwards (see below, §3.3.1). Generally, however, linguistic heterogeneity appears to have been more widespread in supra-regional communication than in later periods. Multilingualism was the norm in the domain of the monasteries; the monks often did not come from the area of the monastery and would frequently move on to other places.

The fourth phase is characterized by a range of historical and social phenomena which, amongst other effects, also brought about a radical change in the linguistic situation. These phenomena have been subsumed under the term ‘Renaissance of the Twelfth Century’ (Haskins 1927), and include the establishment of intellectual centres, the new education in Classical Latin, the growing significance of historiography, abundant translation activity, and the Renaissance of jurisprudence, science and philosophy, as well as the founding of the first European universities. This Renaissance took place at a time which saw the growing importance of cities, the secularization of society and the emergence of political centralization. The consequences of these radical changes for written Romance are best illustrated in the field of law: here, the most important innovation of the twelfth century was a rediscovered concern for Roman law, especially at the new University of Bologna, which soon became the centre of legal education for all of Europe. The ‘Bologna Discourse’ (Kabatek 2001; 2005a) consisted of a new way of thinking, orientation towards a particular institutional centre and a new and restored knowledge of Latin. It reached monastic and secular centres throughout Europe almost immediately, above all due to its combination of civil and canon law. The Latin texts of the *Corpus Iuris Civilis* were studied in these centres, but, by the end of the twelfth century, Romance texts summarizing the new legal system start to appear, first in southern France, where there already existed an established tradition of Romance documents, and then also in northern France, in the Crusader States and in the Iberian Peninsula, whereas in Italy, Latin still dominated at this point (as far as Romanian is concerned, we have no written sources for this period, as all writings are in Greek and Slavonic until the sixteenth century; see Windisch 1993). These compendia for legal practitioners created a new opening for vernacular varieties and led to a functional linguistic differentiation between the late Classical Latin of legal academia and the Romance texts for practising lawyers. In some places, a vernacular literature was written parallel to the legal texts, and sometimes by the same hand. This had direct connections to the new way of thinking and to Roman law, as observable in the works of Marie de France or in Gonzalo de Berceo’s poetry, where direct allusions to the new legal system can be found. Parallel to its effects on jurisprudence, the

Renaissance also had an impact on other scientific domains. Courtly use of Romance led to the development of genuine linguistic centres in various areas, which probably acted as centres of linguistic gravity in more than just questions of writing. This is particularly evident in the case of Paris and Toledo, where, after a period of consolidation during which different linguistic influences from outside converged in these centres, urban elements actually began to spread in the opposite direction. In the manuscript age, this process initially involved only certain social classes in the surrounding areas who maintained contact with the centre; it is also likely that, to begin with, only the written language and certain contact varieties were affected. The more an obvious asymmetry of power took root, the more a verticalization of influence developed: the *koiné* also spread to groups who had only indirect contact with the centre, via others. However, in the written domain, this verticalization occurred only to a slight extent in the manuscript age under consideration here. It became strikingly important after the Middle Ages, when, following the introduction of the printing press in the fifteenth century, the debate on orthography is triggered, mainly in the sixteenth century. The fourth phase is also the one during which the Romance languages to a large extent broke away from Latin, yet concomitantly underwent a process of relatinization in certain text genres (Raible 1996; Barra Jover 2008; see also Pountain, volume 1, chapter 13). The languages were elaborated during this period; that is, they were used for an ever-increasing range of discourse traditions. After the development of their basic structure and their emergence through centuries of oral communication, the vocabulary and the textual techniques required for certain written texts were developed as part of the process of elaboration (in the sense of Haugen 1966, or in the sense of *Ausbau*, Kloss 1987, or 'Verschriftlichung', Oesterreicher 1993). Yet, whilst the development of the basic structures of individual Romance languages in contrast to Latin arises from oral communication, the fourth phase is characterized by renewed European convergence of certain writing phenomena in pan-European discourse traditions, alongside the delimitation of individual Romance language areas.

Looking at the general question of the evolution of written and spoken Romance areas, it is clear that we are dealing with a process that had its starting-point in a situation with oral dialect diversity, but with Latin as a more or less uniform written language. The next stage is the emancipation of regional vernacular writing traditions based on oral varieties, leading to supra-regionalization and unification of these traditions, coinciding with the emergence of supra-regional *koinés*. Thus, the end of the process resembles

the beginning: supra-regional *Dachsprachen* ('roof-languages'; Kloss 1987) and their corresponding written forms dominate local varieties, with the significant difference that the original uniformity of the *Dachsprache* is now perforated by areas standing in opposition to each other, having different standard languages. The underlying dialect continuum at first remained largely unaffected by these developments. It is only in the course of the following centuries (and particularly from the nineteenth century onwards) that different linguistic boundaries emerge within this continuum as a result of vertical contact with different standard languages. To some extent, this goes hand in hand, in some Romance areas, with the complete or almost complete disappearance of the basic dialects.

3. The Romance language areas from east to west

The following remarks on Romance language areas are not intended to be exhaustive. Rather, the aim is to survey some general aspects, highlighting some of the crucial issues and providing basic bibliographical information. The division into areas is a rather broad one and should not obscure the fact that, on the one hand, we are dealing with a dialect continuum without real divisions, and that, on the other hand, in the Middle Ages, most of the larger linguistic areas are still emerging, apart from very large and already established areas such as those of written Latin and Arabic. If we concentrate on these emerging areas, we shall have to avoid any anachronistic national linguistic history based on subsequently established national borders (Kabatek 2007). In Romance historiography, it is generally assumed that variation is characteristic of the Middle Ages, but for later stages there is a switch to a unified perspective without sufficient consideration of the continuity of variation. The growing importance of certain centres will always leave other areas peripheral, with an inherent potential for emancipation, as turned out to be the case in several regions across Europe in the nineteenth century.

We shall have to leave aside the Balkan-Romance area, dominated by Greek and Slavonic in the Middle Ages, since, in the absence of written Romance texts from this period, only speculative comments can be made on koineization processes at the time we are examining. In opposition to the other examined areas, the Balkan-Romance area does not participate in the common western European tendencies sketched above, and Romanian only exists as a spoken vernacular until the appearance of the first written texts in the sixteenth century. We shall also leave aside the Ræto-Romance area,

where reconstruction is possible only with considerable reservation, owing to the lack of any substantial written texts. Even if sporadic written evidence of what we have called the ‘second phase’ exists, the presence of German and the lack of any major Romansh-speaking urban centre did not permit the stabilization of an independent written Romance language, nor the development of supra-regional *koinés*. Documents such as the *Würzburger Federprobe* (‘Würzburg pen test/*probatio pennae*’, tenth/eleventh century) or the *Einsiedel Interlinear Version* (late eleventh century; Liver 1993) may of course belong to a period of more extensive text production; nonetheless, it does not appear to be the case that genuinely Romansh, Ladin or Friulian *scriptae* could have emerged (cf. amongst others Liver 1995). For reasons of space, Dalmatian will also be disregarded (cf. Tagliavini 1972:467–68). There exists, however, an early testimony of Ragusan (the Dalmatian dialect in Ragusa/Dubrovnik) in an inventory list from the end of the thirteenth century, as well as two letters from Zadar (from 1325 and 1397).

3.1 *Italo-Romance*

In no other Romance language area has the question of the *koiné* given rise to such a prolonged and controversial debate as in Italy. Thus, the Italian notion of the *Questione della lingua* has become a prototypical label for metalinguistic discussion of the (predominantly literary) *koiné* (see also Sanson, this volume, chapter 7). The scientific debate surrounding the *Questione* addresses not only the explicit disagreement concerning the Italian standard language since the sixteenth and in part since the fourteenth century, but also its ‘pre-history’, namely, the question of the uniformity and diversity of Italian dialects in the pre-literary period or the problem of the regional characterization or supra-regionality of the earliest written testimonies. The case of Italian Studies also shows how different the perspective on the language of the Middle Ages appears to be in the Romance sub-disciplines, a fact that makes comparison between different areas, such as Italo-Romance or Ibero-Romance, difficult: there is notable variation when it comes to delimiting what are considered to be Romance linguistic monuments (Frank and Hartmann 1997:I, 36) and, since indisputably Romance series of documents or elaborated written texts appear only relatively late in the Italo-Romance area (Trifone 2006:1167), there is a tendency in Italian Studies to consider as early Romance texts a large number of short inscriptions, marginal notes or fragments which are only partly Romance. In other areas, where extensive vernacular text series are available far earlier, such texts are considered to have only marginal significance.

Within our proposal of different phases of evolution (see §2; cf. also Devoto 1953; Koch 1988; Krefeld 1988), the first phase can be said to end for Italian in 960, since the formulaic oaths of the *Placiti cassinesi* are generally seen to be the first clearly vernacular monuments marking the beginning of the linguistic history of written Italian. Because reconstruction is problematic, it is difficult to determine how far back the pre-history of Romance stretches. If we accept the principle that when a political unit ceases to exist, diversity may be triggered as a result, it can be assumed that an important initial step towards the vernacularization of Italy is the Germanic invasions of the sixth century. The basis for the emergence of the Italian dialects is the relative unity of Latin in the Appenine Peninsula. As noted by Terracini (1956), this unity is manifested above all in the lexical domain; the Magra–Rubicon line, which divides northern Italy from central Italy, seems to correspond to a basically phonetic distinction. The first evidence of Romance forms can be found in texts from the seventh and eighth centuries. Various papyri from Ravenna dating back partly even to the sixth century show characteristics of ‘Romance’ morphology, such as, for example, lack of case distinction with only number and gender markers (Sabatini 1965:979). The existence of Greek interlinear versions even allows us to determine phonetic tendencies (like Lat. *fundi* transcribed in Greek spelling as *fondi*; Sabatini 1978:451). In general, in this first phase, numerous ‘Romance’ elements may be found in Latin legal texts (cf. Raible 1993; Hartmann 1992). Langobard legal Latin, christened *volgare italico* by Sanga (1995), is a mixed form between Classical Latin and elements of the vernacular, which appear in the syntax and in the lexicon, above all as regards proper names. This legal written language appears to be relatively uniform across the entire Langobard kingdom; it is more or less remote from the vernacular according to text type. The question of uniformity and convergence would require closer examination (see Jodl 2003), but far-reaching, supra-regional quantitative studies of *scriptae* in the Italo-Romance area are still lacking (although for northern Italy, see Videsott 2009). In particular, examples of ‘Romance’ elements may be found in report-style notes (sometimes written on the back of deeds) that were used for later elaborations of a document (Sabatini 1965); but they are also found in the text of the deeds themselves. On the threshold between the ‘first phase’ and the ‘second phase’, we find the *Indovinello veronese*, a short Latin text with several Romance elements from the second half of the eighth century (Hausmann 1999). Some scholars have classified this text as Romance, but there seem to be more convincing arguments for classifying it as romanized Latin. On the other hand, the *Placitum capuanum*, dated March 960, includes, with several

repetitions, almost an entire sentence in Romance. It is a typical case of 'sporadic' Romance, where some Romance words are inserted in a Latin context. In this case, this is due to the reproduction of an oral testimony in a legal context. In fact, the Romance sentence is presented with the Latin metatextual statement 'testificando dixit'. Then the Romance passage follows: *Sao ko kelle terre, per kelle fini que ki contene, trenta anni le possette parte Sancti Benedicti* (cf. Migliorini 1961:92). The form *sao* 'I know' (in contrast to the typical Campanian dialectal form *saccio*), was interpreted by Bartoli (1945a) as a first sign of linguistic unity, a very first tendency towards koineization in a text actually employing a dialectally marked vernacular. By contrast, it has also been claimed that the form could be due to analogy, without any necessary supra-regional influence (for discussion cf. Sanga 1995:82 and Bianchi *et al.* 1993: 211–12.). The decline of Langobard unity and the political division of Italy once again led to the penetration of regional elements.

Characteristic of the following phases is the emergence of various written *scriptae* with spoken correlates. Sanga (1995:85f.) distinguishes between the southern *volgare beneventano*, the central *volgare toscano* and a northern *lingua lombarda*, written forms which partially coexisted in time. However, instead of being genuinely uniform and stable written traditions, these regional varieties still show tendencies of sporadic writing. Apart from the studies on the northern *scriptae* by Videsott (2009), detailed analyses of these texts with systematic references to the current dialectal situation are still a desideratum. A striking characteristic of Italian seems to be the fact that the phases sketched in section 2 do not appear in a linear sequence, but rather correspond to different areas: thus, the written traditions associated with the Benedictine monasteries in the south, with Montecassino at their centre (the so-called *volgare beneventano*), display comparatively uniform characteristics, even if these texts (for instance, *Ritmo su S. Alessio*, *Pianta della Madonna*) belong to a relatively early period (between the tenth and twelfth centuries) and should in all likelihood be attributed to the period of sporadic Romance text production (our 'second phase'). In contrast to these monastic traditions, the *volgare toscano*, a written language which emerged from the twelfth century onwards and above all in the thirteenth century, stems almost exclusively from the domain of trade and the application of law (for instance, *Conto navale pisano*, *Libro di banchieri fiorentini*). It is, to a large extent, attributable to the fourth phase of regional consolidation, the third phase remaining predominantly Latin in Italian legal documents, in contrast to other areas. The *volgare toscano* emerged from expanding trade and from the growing significance of cities and of the bourgeoisie and is actually closely linked to the renaissance

of Roman law and the consequent linguistic division between the ‘educated Latin’ texts of the lawyers and the ‘everyday’ practical texts written in vulgar language (cf. Castellani 1982b). A clear division of functions, target groups and languages also seems to be responsible for the fact that Tuscan, in contrast to the monastic written forms of the south, was much more a product of the spoken language. It did not seek to achieve a symbiosis of Latin and Romance elements, but rather showed linguistic independence and a clear-cut differentiation of languages, as was observable in several Romance areas in the aftermath of the Bolognese Renaissance in the thirteenth century. From the thirteenth century onwards, a third focus of written vernacular is manifested in the north of Italy, namely, the *lingua lombarda* or *koinè padana* (see Durante 1981; Grignani 1990; Wilhelm 2011), which is the most obvious continuation of the archaic tendencies of the *volgare italico*. This is a supra-regional literary written language, in which a large proportion of early Italian literary texts was produced. The Latinate and rhetorical background of a courtly author such as Guido Fava lends a certain aura of remoteness to his texts (Koch 1987). An exception in the Italian thirteenth-century tradition is the school of Sicilian poetry, which can also be assigned to the fourth phase and shows direct ties to the University of Bologna. It came into being as part of Frederick II’s Magna Curia and had strong links to the southern French *trobadors*, as a literary phenomenon of rather short duration (1230–60) and was based more on Latin and Provençal models than on the Sicilian dialect, as already observed by Dante in *De vulgari eloquentia*, where he claims Sicilian to be a general name for the language of Italian poetry, not comparable to the dialect of the locals.

In contrast to the archaic tendencies of the north and the south, Tuscany displays its own innovative tendencies, providing a balance between dialect differences and soon coming to play a leading role in the further convergence of the written language. This was not so much the result of the immediate spread of Tuscan, but rather due to the conscious elaboration of the Florentine dialect, where a selection process weeded out elements considered as dialectal, and Latin elements were borrowed in order to create a *volgare illustre* from the end of the thirteenth century onwards. This by no means resolves the *Questione*, which in the Renaissance essentially concerns the competition between Tuscan and the *lingua cortigiana*. However, the *koiné* based on Tuscan becomes clearly predominant and outdoes its competitors in the course of the centuries. This can basically be attributed to the high prestige acquired by Florentine as the language of the *Tre corone* – Dante, Petrarch and Boccaccio – as well as to the central geographical position and the economic

significance of the city of Florence. The dominance of Florentine over both Latin and the other Italian dialects is in fact not achieved until Bembo's grammatical and stylistic prescriptions in the sixteenth century, giving rise to a new conflict which shifted the focus of the *Questione* to the difference between the literary language of the *Tre corone* and the spoken language.

In summary, it can be said that, in the case of Italian, 'Romance' elements can be observed at a very early stage in written Latin texts, but writing in Romance remains the exception for a considerable time, and only towards the end of the thirteenth century does an extensive production of written Romance texts begin. Tuscany plays a leading role in this development as a result of its central position and the significance of the Tuscan cities, but also because of the prestige of its literary output. The dominance of Tuscan is, however, not linked to any constant political and cultural centre of radiation. Therefore, in Italy, the non-existence of a common spoken language linked to the written language continues for a long time.

3.2 Sardinian

In the case of Sardinian, it is particularly clear that external developments are of crucial significance for the history of the written language. The establishment of the jurisdictional areas of Gallura, Torres, Arborea and Cagliari in the eleventh century appears to have been based on ethnic-cultural boundaries and may even have strengthened the significance of these boundaries from a dialectal point of view (Blasco Ferrer 1984:63). Almost no traces of Sardinian are extant from the first phase, which can thus only be reconstructed; the second phase is nugatory (various texts which supposedly could be attributed to this period having been shown to be forgeries; cf., on the so-called *Carta di Arborea*, Frank and Hartmann 1997, I:28–35; see also Blasco Ferrer 1984; 1993; 1995a). There exist some isolated texts, like the *Privilegio logudorese* (also called *Carta consolare pisana*), from the beginning of the twelfth century, as well as some legal documents from the twelfth and early thirteenth centuries. From the thirteenth century, several collections of internal or external administrative documents, the so-called *Condaghi*, as well as a series of charters, are preserved. They are the main source for the reconstruction of medieval Sardinian (Frank and Hartmann 1997, I:34).

From the end of the eleventh century onwards, an 'external' development, namely, the arrival of Benedictine monks from Montecassino, marks the beginning of a strong foreign influence on Sardinian and the almost sudden appearance of entirely Romance texts. The Benedictines initiated an extensive production of texts, which soon spread to local writers. In the Romance

passages of legal documents, differences related to the various local dialects can be observed between a Campidanese, a Logudorese and an Arborean *scripta* (see Blasco Ferrer 1993:114–16), even if in practice the influence of the peninsular models is the dominant one. With the loss of Sardinian autonomy in 1297 and the Catalan–Aragonese conquest between 1323 and 1410, Sardinian again disappeared from legal documents. The lack of political unity and the permanent presence of other prestige languages (Catalan, Spanish, Italian) prevented Sardinian koineization in the following period and kept Sardinian at the stage of dialectal fragmentation.

3.3 *Gallo-Romance*

The Romance language area for which the problem of medieval linguistic convergence has been by far the best studied is Gallo-Romance, the language area stretching from the border with Italo-Romance and Ibero-Romance up to the Germanic language area. It will not be possible to discuss the individual regions in detail; the observations in the following section must be restricted to some general aspects of French and Occitan.

3.3.1 French

In traditional approaches, the medieval history of northern France was considered as part of national history, in the sense that it was seen as a pre-history of literary Classical French. Old French appeared, particularly in university teaching, as a rather unified and standardized language. During the last few decades, a different view has emerged, with precursors in language geography studies since Gilliéron, in scriptological approaches and, more recently, in studies stressing varieties and variation within the medieval languages (amongst others Cerquiglini 1989; Buridant 2000; Völker 2003; 2006; Hafner and Oesterreicher 2007) and applying sociolinguistic terminology and methods to medieval linguistics (for instance, Wright 1982; 2001; Banniard 1992a; Lodge 1993; 2011). In the case of French, the central questions relevant to our topic are the following:

- the question of the dialectal basis for processes of convergence (phase I);
- the question of possible tendencies of convergence in the sporadic Romance texts before 1200 (phase II);
- the emergence of regional written traditions in Romance from the twelfth century onwards, their internal development and their mutual relationship (phase III);

- the question of the role of Paris as a centre of gravity and the emergence and consolidation of the French language (phase IV).

With respect to the issue of the dialectal basis of French, the same factors are traditionally adduced for Gallo-Romance as for other Romance areas. The bibliography in this field is particularly rich and gave rise to several hypotheses concerning the emergence of Romance which subsequently figured in the discussion of other regions. The substrate hypothesis (Brun 1936), according to which the essential dialect areas are already delineated by pre-Romance languages, stands alongside the superstrate hypothesis (Wartburg 1950), which attributes the dialect division primarily to the various Germanic conquests. Moreover, the role of other factors, such as the Roman provincial divisions (Merlo 1941), the medieval boundaries of ecclesiastical administrative areas (Morf 1911) and transport routes (Lüdtke *apud* Kontzi 1982) has also been stressed. All of these factors certainly played a more or less influential role, and it would be wrong to give priority to any monocausal explanation. It can even be said that the various influences determine each other to a certain extent. Thus, the Roman administrative units were supposedly based at least partially on existing geographical and ethnic boundaries, and there is a further link between these, the areas of Germanic settlement and the medieval diocesan boundaries.⁶

The two main factors in establishing an overall division of the Gallo-Romance area are the far more intensive and continuous Romanization of the south and the more profound Germanization of the north; the reconstruction of subareas beyond these is generally a more difficult task. As is the case elsewhere, it seems that conclusions may be drawn about medieval linguistic geography from the dialect situation which currently obtains in certain areas. In other areas, however, owing to the spread of the French standard language, strong substitution or convergence processes occurred, resulting in the partial loss of the dialect contours. This is particularly true of the Île de France and surrounding areas, above all Champagne, Burgundy and Franche-Comté. Closer to the border with Occitan, dialects such as Poitevin are better preserved, the best preserved being those of the north, with the exception of Anglo-Norman, which has completely disappeared. It is no coincidence that the most successful studies of *scriptae* relate to the north and the north-east.

⁶ A good example of this mutual influence is the coincidence in Normandy of the boundaries of the bishopric of Rouen and the border of the second administrative area of the Roman province of Lugdunum.

It regularly occurs in the history of languages that processes which appear to the contemporary observer to be examples of linguistic change are in fact indirect consequences of certain broader historical processes, whereby something which had previously been latent now becomes visible – and French is no exception. Thus, the Germanic conquest is frequently said to be the reason for the dialectalization of France, although it should rather be assumed that the most important consequence of Germanization was the loss of a certain previous superficial unity, which caused a pre-existing heterogeneity to surface (Wüest 1979:343). For the general division of the Gallo-Romance area into a southern and a northern area, Wüest (1979:354–59) cites a series of geographical factors which favoured different tendencies in settlement or certain contact situations. In the north, the dialect areas had already been formed in the Merovingian period (448–751), long before the appearance of the first written documents; namely, Francien in the centre; Walloon, Picard and Norman in the north; Bourguignon, Champenois and Lorrain in the east; and Angevin, Poitevin and Berrichon in the south. There is a dialect continuum that borders Occitan in the south and the Franco-Provençal dialects in the south-east (Lyonnais, Franc-Comtois, Romand, Savoyard and Dauphinois; cf. Vurpas 1995). According to Remacle (1948:141), the basic dialect division was established in the ninth century, and the dialect boundaries were further shaped during the following centuries.

It is debatable whether supra-regionality can already be observed in the first sporadic written evidence (phase II) and to what extent these texts indicate existing Romance written traditions. This issue has been discussed in most detail with reference to the French passage of the *Strasbourg Oaths* of 842, generally considered to be the oldest French text. As in the case of other Romance texts from this period, the *Oaths* are reproduced in a chronicle in their ‘original version’ to reflect the authenticity of the eye-witness report. The main problem with this text is not the question of the authenticity of the copy in the manuscript dating from the tenth century, but far more the fact that it might possibly not be an authentic repetition, but rather a historiographic stylization of the Romance language. The chronicler Nithard, a grandson of Charlemagne, wanted above all to create a kind of mimesis of the vernacular language, and not an authentic report, as has been repeatedly stressed (McKitterick 1991). For this reason, we might question the value of attempts to localize the language of the *Oaths*. Furthermore, the text is marked by formulaic expressions from the language of Latin documents and by the written style of Merovingian Latin (Ewald 1964). Determining the geographical origin of the *Eulalia Sequence* (the oldest literary French text, dating from

the end of the ninth century) has been seen as a much easier task. It originates from the area of Hucbald de Saint-Amand in Saint-Armand-les-Eaux, in the Picard-Walloon area. Hilty (1973), arguing against the common assumption that writers initially simply wrote in their local dialects, attempted to demonstrate that the presence of glide consonants (*voldret* 'he wanted', *voldrent* 'they wanted', *sostendreiet* 'she would put up with') in the *Eulalia Sequence*, as in the *Strasbourg Oaths* (*sendra* 'lord'), represents a central French characteristic, so that even the very first written texts provide us with evidence not only for the locality of a particular dialect but also for supra-regional tendencies.

On the basis of the dialect studies of the *ALF* and all French documents from the period before 1200, Pfister (1973) compares the twentieth-century dialect situation with medieval written documents. For the period between the sixth and the ninth centuries, he assumes that there were various innovations originating from a central French area. These innovations (such as the fronting of *u* and the spread of *a > ae* and *iei > i*) reached some peripheral areas but not others, leading to an isolation of the latter. Pfister doubts that these innovations spread from Paris. It is often assumed that Paris, due to its central position, had developed into an influential centre at an early stage, but a distinction must be made between Paris itself and the surrounding area, the Île de France, where cultural centres had in fact been established from a very early period. Documents from Bourges, Angers, Tours, Paris and Orléans dating from the sixth and seventh centuries are available (Pfister 1973:251), yet the production of texts written directly in Paris appears to be minimal in the Merovingian period, and there are absolutely no written texts from Paris dating from the Carolingian period. Even the Abbey of Saint-Denis near Paris, founded in the seventh century, appears to have acquired supra-regional importance only in the twelfth century. Furthermore, Lodge (1993:102) mentions the strategically unsuitable position which made Paris vulnerable to sea attacks by the Vikings as an argument against the city's having had a leading role early on. After the supposed initial period of central French innovation, the centre of innovation shifted between the ninth and the twelfth centuries to the Picardy-Flanders-Wallonia area, where important cultural centres were settled (Corbie, Saint-Riquier, Saint-Amand, Laon) and important scholars were active. The written innovations of this period also took effect in Paris (spread of the glide consonants *b* and *d*, change of *ei > oi* and of *ou > eu*; Pfister 1973). Only as Paris became an undisputed centre of radiation from the second half of the twelfth century onwards did these innovations spread secondarily. Paris became a source of radiation for both central and northern innovations. The first metalinguistic evidence for the significance of the language of Paris

appears at the end of the twelfth century, when the Picard poet Conon de Béthune reports that he was mocked in Paris because of his provincial accent (Brunot 1966:329). Later on, praise of the characteristics of Parisian language becomes more frequent, although such metalinguistic comments generally lag behind actual linguistic developments. Nevertheless, the actual force of Paris's radiation does not appear to have begun much earlier than the time of Philippe-Auguste (reigned 1180–1223). The metalinguistic comments also show that indirect signs of linguistic developments visible in the spread of certain graphemes might in fact correlate with spoken phenomena, if only in the spoken language of a very specific and restricted class.

It was the twelfth century with its radical changes that enabled Paris to become an undisputed centre. During this century, the city's population grew exponentially, and Paris became the most influential urban centre in northern Europe; trade flourished and important clerical centres were established. The relocation of the royal residence to Saint-Denis in the first third of the twelfth century appears to have been effect rather than cause, but the elevation of Paris to the status of capital city clearly consolidated its central role. Paris also became the seat of Europe's most important university, along with that of Bologna. Thus, the city concentrated the different elements which characterized a 'modern' metropolis in the twelfth century: it was a centre of trade and a political centre, and possessed a sufficient cultural (monastic) basis for the more secular society of the thirteenth century to build upon – the century during which the language of Paris began its expansion beyond the confines of the city.

In the course of a general growth in text production in the thirteenth century, regular Romance text production marks the beginning of the third phase in different areas of northern France (Frank and Hartmann 1997: IV):

- from 1246 in Normandy;
- from the beginning of the thirteenth century in western France;
- from 1241 in the Île de France;
- from the end of the twelfth century in Picardy;
- from 1233 in Wallonia;
- from 1219 in Lorraine;
- from 1228 in Champagne;
- from 1233 in Burgundy.

Before 1250, text production is still sporadic in some areas, whereas in Picardy, Lorraine and Champagne, extensive series of documents are already available. On the one hand, the documents display *scripta* phenomena which appear to

be partially linked to the local dialects. On the other hand, it is striking that even from the beginning of extensive Romance text production, tendencies towards convergence can be observed. In different areas, these tendencies reveal, to varying degrees, alignment of regional orthographic conventions with those of regional centres or with the central written forms of the Île de France, as may be observed from instances of hypercorrection (cf. Remacle 1948; Gossen 1967; Goebel 1970). It has been observed that the convergent tendencies do not appear in the sense of positively attested adoptions of Parisian writing traditions, but rather in an avoidance of writing habits perceived or marked as local (Voßler 1929:27). In this context, it is important to stress the differences in the communicative range of the documents, particularly visible in the case of the contrast between local originals and their copies with supra-regional scope. In the case of the Norman form *rei*/*rey* 'king', in contrast to the central form *roy*, Goebel (1975:184) provides an example of a virtually linear loss of the regional form in original documents and copies alike between 1246 and 1551.

The expansion of Francien is also supported by political events: in 1284, the province of Champagne is incorporated into the area of the crown, and shortly afterwards, the economic importance of the Picard cities begins to decline.

From the end of the twelfth century onwards, and particularly during the course of the thirteenth, substantial and elaborated literary texts belonging to the fourth phase begin to appear. This is initially the case in Picardy, Normandy, England and Champagne, but then, in the thirteenth century, also in Paris, which gradually developed into a literary centre. However, given that this development occurred earlier in other regions, their literary prestige continued to radiate outwards to the Île de France well into the thirteenth century. In this context, we can speak of the Franco-Picard *scriptae*, as even the texts produced in Paris display a Picard bias. The language of Paris appears first to have functioned as a supra-regional standard pronunciation and only later on to have gradually come to be an orthographic norm. Interpretations of the role of the language of Paris which adopt the terminology of variationist linguistics have attempted to shed more light on the relationship between Paris and the surrounding areas (cf. Völker 2001a). Beginning in 1328, the Hundred Years War was to bring about a sea-change in French society, as it led to the collapse of regional feudalism and its replacement by a centralized monarchy and a horizontal political structure; but even at its outset, Paris was already the undisputed centre of the written and spoken French norm. Its supra-regional radiating effect was reinforced yet further as a result of the war and its consequences.

In 1539, the French *scriptae* lose their mixed nature, as the official language of Paris becomes obligatory in all legal documents in France. The language of Paris, which had been a focus of orientation for French koineization since the twelfth century, would repeatedly undergo radical changes in the following centuries. However, Paris remained the linguistic centre of gravity; it is here that the discussions about the standard language in the sixteenth and seventeenth century and the subsequent establishment of *bon usage* as a measure of normative-linguistic orientation would take place; in Paris, the centre of the French *koiné* becomes equated with a particular style of a certain urban class, and the standardization of orthography is equated with the Parisian French Academy – a situation that has continued up to the present day.

3.3.2 Occitan

The impression of a bipartite division of the Gallo-Romance dialect continuum is justified when viewed with hindsight; there are, in fact, some early indications of certain distinctions between the northern and the southern French dialects, while transitions are also revealed. Gascon, which displays some Ibero-Romance characteristics, may be included amongst the Occitan dialects (Limousin, Auvergnat, Béarnais, Languedocien and Provençal). As was the case for the northern French area, we may assume that the most important dialect boundaries in southern France were established relatively early on. There are, however, several fundamental differences with respect to the north. First, Occitan is the oldest Romance written language from a purely chronological point of view – not in terms of sporadic evidence from the second phase, but in terms of the first text series (phase III) and with reference to the first elaborated texts in Romance (phase IV). Second, Occitan is, compared to French, a language with a low degree of dialect differentiation and a high degree of orthographic uniformity. Third, in the Occitan area no clearly recognizable urban political centre acts as a linguistic centre of gravity. Finally, the development of a uniform Occitan *koiné* is interrupted as early as the thirteenth century and has ended permanently by the sixteenth century.

Occitan is found as a written language in documents as early as at the end of the eleventh century (Brunel 1926; Frank and Hartmann 1997, IV:347). The oldest documents originate from military orders, the Knights Templar and the Knights of St John, in Albi and the Rouergue. In the first documents, which are mostly wills or feudal oaths, Occitan may be found in free sections alongside Latin passages. Texts which are Occitan even in their formulaic parts are, however, found very early on. It is striking that even the early documents to some extent employ quite uniform written forms. This can be explained by

the fact that writing was restricted to just a few monastic centres which were in contact with one another and the fact that there were only a small number of discourse traditions. Bec (1986) establishes that the uniformity of the language is greater in the twelfth and thirteenth centuries than in the fourteenth century, and that in the Occitan area, in contrast to northern France, there was increasing dialectalization. This process can be linked to the lack of a centre, mentioned above, and to the decadence of the Occitan nobility following the Albigensian crusades in the thirteenth century. The language of the Occitan documents differs from the texts of Béarn, whose capital was Pau. According to Bec (1986:71), the other areas – Albi, Quercy and the Rouergue – developed an ‘administrative *koiné*’ very early on, although its existence is doubted by others (Gleßgen and Pfister 1995:410). The main characteristic of the old Occitan documents is their dialectalism (cf. also Grafström 1958), which even increased in later periods. The most important urban centre of Occitan is Toulouse, the former Visigothic capital.

Besides being the first Romance language used for legal documents in their entirety, Occitan also provides us with the first literary Romance *koiné*. Supposedly, the language of the *trobadors* corresponded to an ancient oral tradition. The first known testimony of this tradition can be found in the first *trobador*, William of Aquitaine (Pfister 1970). The language of the *trobadors* displays a certain uniformity due to the wide distribution of the texts and melodies (northern France, Germany, Italy) and the fact that the different generations of *trobadors* always referred to one another and that their texts are closely woven together intertextually (Gruber 1983; Paden 1998; Beltrán 2005). However, closer examination reveals internal differences, regional features and, partly, northern French influences in the case of individual *trobadors* (Pfister 1976). The language of the *trobadors* is also the first Romance literary language described in didactic texts as early as the thirteenth century (Swiggers 2011). *Lo donatz proensals* by Uc Faidit and *Las razós de trobar* by Raimon Vidal constitute the oldest evidence of Romance grammaticography. They were used in the teaching of Occitan to give foreigners access to Occitan poetry. The *Donatz* addressed an Italian audience and the *Razós* a Catalan audience. Occitan grammaticography reached its height in the region itself in the fourteenth century, when the Toulousain *Leys d’amors*, an admirable text describing the language, appeared. In contrast to the *Donatz proensals*, which is to a large extent derivative of the Latin *Donat*, this is a genuinely independent work which applies the terminology of Roman law to language description. The *Leys d’amors* contains a detailed description of Occitan, including interesting observations concerning pronunciation. It should also be mentioned that

the work came into being in an urban, bourgeois environment (Coseriu and Meisterfeld 2003:31–49).

The Occitan language area is also in the vanguard of the fourth phase: it was here that the first extensive legal prose text in Romance was written between 1149 and 1170, the legal summa *Lo codi*, which summarizes the Justinian laws of the *Codex Iuris Civilis* in a vernacular version (Derrer 1974). It is generally assumed today that *Lo codi* is not an isolated work but that it stems from an important southern French school of law that flourished in the middle of the twelfth century, with centres in Arles, Valence and Saint-Gilles, and which also produced important Latin works on Roman law, such as the so-called *Summa Trecensis* which served as a model for the Occitan *Codi* (cf. Gouron 1978; 1985). This school is often described as the ‘Valence School’, but its precise location remains unclear (Weimar 1972:24; Gouron 1978:113). The text of *Lo codi* is not the result of a gradual process of replacing Latin, as is the case for the legal documents, but rather it reflects an elaborate process of vernacular writing which came ‘from above’ and could become established only in a highly educated, Latinate legal environment. Pfister (1978) attempted to situate manuscript A of the *Codi*, which dates from the twelfth century, on the basis of scriptological evidence. *Lo codi* is an outstanding example of the European nature of early elaborated Romance writings. In the mid twelfth century, Roman law spread rapidly throughout the whole of Europe, on account of its connection with canon law, as laid out in Gratian’s *Decretum*. Thus, *Lo codi* was created at a time when Roman law was not only an object of study for legal scholars in the newly established university domain, but was also being applied practically as a ‘new’ and appropriate law for the growth in trade, for the cities and centralized power structures. As a result, vernacular versions summarizing the new law were needed in different European regions as early as the end of the twelfth century, and above all from the thirteenth century onwards. *Lo codi* appears to be an isolated case in the history of Occitan; however, the text was translated into numerous languages. Hence, various old French versions of this text are known, as well as several Occitan manuscripts, a translation into Latin originating from Italy, and translations into Franco-Provençal and Castilian. A Catalan version probably existed but is no longer extant. The influences of this text can be seen in various *coutumiers* in northern France and in Romance legal writings from the Iberian Peninsula and Italy (Kabatek 2000a; 2005a). The extensive spread of the text also attests to contact between vernacular writing traditions.

In summary, the Occitan language area can be shown to have played a leading role in the establishment of Romance writing traditions and in their

unification. Yet within the Occitan area, these innovations have no continuity. The Albigensian wars in the thirteenth century weakened the Occitan area not only politically but also with regard to its autochthonous language. The final suppression of Occitan was completed in the sixteenth century with the officialization of French under François I. The lack of certain discourse traditions in Occitan (Gleßgen and Pfister 1995:406f.), as well as the lack of a spoken *koiné*, correlates with the non-existence of a unified literary language. Only in the nineteenth century did the Romantic Occitanist movement make a renewed attempt to unify the literary language (see Sanson, this volume, chapter 7). Such attempts have remained marginal up to the present day. The most significant effect of the Occitan innovations for later periods is seen in other areas; namely, in the way in which Occitan served as a model for written Romance in northern France, in Italy and in the Iberian Peninsula. It is not only the general tradition of writing legal documents in the vernacular that emerges from the Occitan area; one can even observe the export of concrete linguistic elements. The first Romance poetic tradition started here and spread far beyond the south of France, setting a precedent for the emergence of regional literary languages in many places. It is also from this area that the first Romance prose production originates, representing the first stage in the development of an elaborated Romance vernacular tradition.

3.4 *Ibero-Romance*

With regard to Ibero-Romance, some general observations are in order before we look in detail at the three main blocks: Catalan, Castilian (including Navarrese, Aragonese and Leonese) and Galician-Portuguese. A certain internal heterogeneity can be observed in Hispanic Latin, the foundation of the Ibero-Romance languages, and this can be related to the time at which the individual areas were Romanized and to the Latin of the Romanizers, as when we compare the language of the Andalusian patrician colonies with that of the Greco-Roman traders on the east coast (Meier 1930; Penny 2002:8–13). After the various waves of Germanic invasions, the political unification achieved by the Latinized Visigoths in 585, with Toledo as their capital, could probably have provided a basis for linguistic unification, if the Arab conquest of 711 had not led to political fragmentation, with individual Christian nuclei in the north presenting an obstacle to a large Arab-ruled area. These nuclei were the Spanish Marches in the east, Aragon and Navarre, and, in the west, Asturias and Galicia. Asturias soon enlarged its territory and became the kingdom of León, which included the county of Castile in its process of emancipation. For the Iberian Peninsula, new linguistic findings in recent decades have led to a

more differentiated description of the situation of the proto-Romance first phase before the eleventh century (Díaz y Díaz 1978; 1996; 1998). The basic principle applying to the entire peninsula is that of territorial expansion from north to south, which, as part of the Christian reconquest of the Arab areas, gives rise to the southward spread of the northern dialects: Galician to Portugal, Leonese to Extremadura, Castilian to the centre and Andalusia, Aragonese to Murcia, and Catalan down to Valencia (Tuten 2003; Cano Aguilar 2005). Together with the 'axe-shaped' spread of Castilian from north to south during the Reconquest (Menéndez Pidal 1926:513; Quilis Merín 1999; Pons Rodríguez 2010:62; Fernández-Ordóñez 2011:21), the influence of Castilian also 'spread toward the south fanwise' (Vendryès 1925:264; see also Penny 2000:74–128; 2002:14–19). Thus, it was Castilian, the dialect of the centre, which broke up the linguistic similarity between the dialects of the east and west and which, through its relative distinctiveness, substantially contributed to the later picture of the linguistic heterogeneity of the Iberian Peninsula (cf. also Lleal 1990; Torres Montes 2006). In this section, the discussion must be restricted to a few brief observations concerning Catalan and Castilian, as well as Galician and Portuguese. Navarro-Aragonese (Menéndez Pidal 1926:460–72; Alvar 1953; Martín Zorraquino *et al.* 2000), together with Leonese, represents the continuation of a certain linguistic unity from the Visigothic period. Within the Navarro-Aragonese language area, the important Riojan monasteries, above all San Millán de la Cogolla, assume a special position owing to their supra-regional significance. This is where the first Romance texts belonging to the second phase are found. A considerable number of early Latin texts with numerous Romance characteristics can be found in the Leonese area, a fact which led Menéndez Pidal (1926:454–60) to postulate a triglossic situation, in which spoken Romance coexisted alongside 'Leonese Vulgar Latin' and scholarly Latin. Wright (1982) disputes this claim and classifies Leonese Vulgar Latin as written Romance dating from the time before the Cluniac reform of pronunciation and orthography.

3.4.1 Catalan

Catalan, described as a bridging language, owing to its linguistic proximity to both Gallo-Romance and Ibero-Romance (Baldinger 1971:125–60), is, to a greater extent than Castilian, aligned in a Romance continuum covering both sides of the Pyrenees. Throughout the entire Middle Ages, Catalan is closely linked to the south of France and is subject to influences from other Mediterranean areas. Its linguistic proximity to Occitan is attested from the

earliest written evidence onwards. The political centre of Catalonia is the capital city of the county of Barcelona, which has remained a centre of linguistic gravity up to the present day. After the conquest of the kingdom of Valencia, a further centre of gravity was established, whose linguistic rivalry with Barcelona became apparent at a later date, for example in the conflict concerning the name of the language (Colon 1978; Eberenz 1989).

Statements claiming that a large degree of uniformity characterizes even the first written texts within the Catalan language area contrast strikingly with present-day dialect diversity (Veny 1985:31–38), which can hardly be attributed solely to developments since the medieval period. The most important dialect boundary is that between eastern and western Catalan, which corresponds to a line leading northwards from the west of Tarragona, and according to which the dialect of Barcelona and its surrounding area, including northern Catalan in Roussillon and the dialects of Tarragona and the Balearics, is classed as eastern Catalan, whereas the dialects of Lleida and Tortosa up to and including Valencia belong to western Catalan. Eastern Catalan in Sardinia changed in many respects due to contact with Spanish, Sardinian and finally Italian in the centuries after the conquest. The relative uniformity of the medieval literary language does not necessarily contradict the observation that the principal dialect areas were already established in the first phase, namely, in the High Middle Ages. Badia i Margarit (1981) considers a combination of substrate effects and Arabic superstrate effects to be responsible for the dialect division. Western Catalan is said to have been more strongly influenced by the pre-Roman substratum than eastern Catalan, and the Arab conquest is said to have had a considerably lesser effect on northern (or eastern) Catalan than on southern (or western) Catalan. In contrast, Blasco Ferrer (1995b; 1995c) suggests that the main reasons for the dialect differences are resettlement, the different Latin foundations and the diocesan divisions, which defined areas of different cultural development. Instead of uniformity, he identifies two *scriptae* which are largely coterminous with the later dialect areas and which can be clearly differentiated in the texts of the thirteenth and fourteenth centuries on the basis of a scriptological study of non-literary texts. In the case of the first sporadic texts (phase II, for example, Catalan fragments of the *Forum Iudicum* from the first half of the twelfth century or the *Homilies d'Organyà*, at the end of the twelfth century; Moran 2004), some written solutions for the representation of Catalan phonemes indicate an already existing tradition. In the initial phase of the development of regional text series (phase III), restricted areas may still be singled out (Majorcan vs. Rossellonese vs. Valencian vs. central *scripta*), but a differentiation of two main blocks soon emerges, the influence

of Barcelona becoming increasingly evident. In the fourteenth century, standardizing tendencies of a uniform official language can be discerned in Barcelona (above all under Peter III, 1336–87). After the death of Martin the Human (1410), the chancellery moves to Valencia, adopting the unified official language of Barcelona without modifications; older texts are partially brought into line in adaptations, and the abundance of synonyms in earlier texts, which indicated the copresence of different varieties, is reduced as a result of a conscious concern for linguistic purity (Blasco Ferrer 1995b:480–84). Parallel to the development of the language of legal documents, from the thirteenth century onwards, an important legal tradition of prose texts emerges, due to the Catalanization of Roman law (*Costums de Tortosa*, *Furs de Valencia*, etc.) and in the writings of Raimundus Lullus (Ramon Llull). The lexicon and the spectrum of linking techniques is extended, and polymorphism is dispensed with, so that the Catalan written language achieves a high degree of independence with respect to the Latin models (Duarte i Montserrat 1996).

This uniformity of the written language does not last long, however, and has no counterpart in the spoken language or in linguistic awareness. The coexistence of different centres of radiation and the lack of political unity serve to keep the debate on Catalan koineization lively right up to the present day.

Detailed research into Catalan has, to some extent, qualified its linguistic history. Claims made in the context of nation-building from the beginning of the twentieth century onwards that Catalan had already achieved the status of an official or even a national language at the time of James I (1213–76) contradict the clear findings which show that the production of texts in Catalan was minimal in comparison with the production of Latin texts (Philipp-Sattel 1996:10). This is the case in spite of the undeniable political and cultural significance of James I, under whose leadership important, predominantly legal, Catalan texts were produced. It is above all in popular intellectual discussions in Catalonia where we find the claim that Catalan was a ‘completely standardized’ language as early as the thirteenth century. Irrespective of the rather vague definition of the term ‘standardized’, such claims are disproved by historical findings and more often have the aim of legitimizing current demands rather than of accurately portraying the diachronic situation.

3.4.2 Castilian

We single out Castilian when describing the central area of the tripartite division of Ibero-Romance not because of an anachronistic perspective based on hindsight and the knowledge that an established unified Spanish language

would later emerge, but rather because the expansion of Castilian takes place within the entire central area (and in part beyond) in the period under consideration, despite the much greater significance in earlier periods of other dialects, such as Leonese or Navarrese. It has already been noted that the most important distinction in the Iberian Peninsula is that between the dialects in the north from east to west. During the Reconquest, an additional distinction appears, namely, that of the north–south axis, with mutual influence and levelling among the dialects spreading southwards, whilst the archaic forms of the dialects in the northern mountains are less, or even scarcely, affected by these changes (Kabatek 2007; Penny 2000).

Several problems arise when reconstructing the linguistic situation and convergence processes of the central part of the Iberian Peninsula during the Middle Ages. For certain areas, above all those of the archaic northern dialects (whilst they continue to exist), we may assume that there is a close link between modern-day dialect divisions and medieval linguistic geography, as in other parts of the *România*; however, this assumption is much less justified, or even completely unjustified, elsewhere, namely, in the case of the dialects which were eclipsed by the spread of Castilian from the thirteenth century (and above all from the fifteenth century) onwards, or which only acquired their later form as a result of migration and contact in the course of the Reconquest. A further problem arises from the question of the link between regional written production and regional dialect, since the Iberian Peninsula was marked by a high degree of mobility during the Middle Ages, with several linguistic consequences. Thus, Catalan was sometimes spoken and written in Aragon, and there were French colonies along the Way of St James, in which Occitan or Occitan-Spanish mixed forms were written (Beltrán 2005), whilst in the Riojan monasteries there was a massive influence of Mozarabs from the south, whose Romance dialect can only partly be reconstructed (Ariza 2005).

In the case of Spanish, as a result of the studies carried out by Ramón Menéndez Pidal (1926), a precursor of *scripta* research has been in existence since the 1920s, far earlier than for the Gallo-Romance area. This work takes a comparatively modern approach and includes the differentiation of several varieties, a clear-cut division between orthography and pronunciation and, above all, the combination of a geographical-linguistic approach with the rigorous study of non-literary documents, analysis of the historical background and consideration of centres of linguistic radiation. This gave Spanish Medieval Studies a certain methodological lead over other areas. From the perspective of current scholarship, these predominantly qualitative

surveys need to be enhanced by quantitative analyses, the prerequisite for which is already in place in the form of reliable editions of original documents. The region has also been well studied from a dialectological perspective, including regional studies. Certain developments, such as the increasing Castilianization of the neighbouring dialects, have been extensively attested by detailed studies of texts, including consideration of the importance of institutions and centres, which enable us to draw at least indirect conclusions concerning linguistic changes of gravity.

Castilian, the language of the Iberian Peninsula which has the greatest significance from a later viewpoint, grew out of a dialect which was initially only spoken in a small area in the Cantabrian mountains which formed part of the kingdom of León. In the case of Castilian, the close relationship between koineization, unification of writing and the expansion of domains of political power can be seen particularly clearly.

Spanish historiography traditionally exaggerated the particular role assumed by Castile within the areas of the north, when it claimed that Castile had very early, already under Count Fernán González in the tenth century, become the most important power of the Reconquest. This image of the 'revolutionary nature' of Castile was transferred to the language (Menéndez Pidal 2005:359–63). It repeatedly seems as if the particular character of Castilian does indeed display a range of characteristics which distinguish it from neighbouring dialects (diphthongization from *o* > *ue* and *e* > *ie*, as was also the case in Aragonese and Leonese, but in contrast to Galician-Portuguese and Catalan; change of *f* > *h*, etc.). However, it seems highly doubtful that language-internal criteria are responsible for the territorial ascent of Castilian; the explanation is rather to be sought in political and historical factors. Furthermore, the highlighting of a special position for Castile before the eleventh century can be proved to be an *a posteriori* mythification (Márquez-Sterling 1980; Martin 1997); Castile was nothing more than the relatively sparsely populated eastern part of the kingdom of León, marked by territorial battles. The emphasis on its supposed special status in the ninth century actually dates largely from the thirteenth century, a time in which it did indeed enjoy supremacy and a moment when a kind of national history is, to some extent, retrospectively created (Kabatek 1999b).

The first urban centre of radiation of Castilian was Burgos, where different influences from the surrounding area converged (Menéndez Pidal 1926:485–89; Lapesa 1989:182; Tuten 2003:94–144). The most important centres of writing in the eleventh and twelfth centuries are the monasteries along the Way of St James, above all San Millán de la Cogolla, Santo Domingo de Silos

and the Galician city of Santiago de Compostela. Following the conquest of Toledo (1085), an antagonism grew up between this city and Burgos, the effects of which left historical linguistic traces into the sixteenth century. Burgos rapidly gained significance after the Castilian conquest, with the domination of the local nobility and the local dialect. In contrast, Toledo, the former Visigothic capital, had a long and complex tradition and contained a relatively heterogeneous population of Mozarabs, Leonese, Franks, Castilians, Moors and Jews. Linguistically, it was characterized by a mixture of different varieties and soon achieved supra-regional significance. The antagonism between the two centres is evident in the political domain (regional minor nobility in Burgos, royal power in Toledo) as well as in the legal system (common law in Burgos, statute law in Toledo) and can certainly be extended to the linguistic situation (Kabatek 1999a; Tuten 2003:94–144).

After the first sporadic evidence from the eleventh and twelfth centuries (for discussion see Wright 1982; Quilis Merín 1999; Tuten 2003), Castilian enters the third phase of our model in the thirteenth century. From around 1220 onwards, the first series of documents are found (Menéndez Pidal 1919). During the first half of the thirteenth century, a ‘more conservative’, Latinizing current competes with an innovative Castilianizing tendency in Toledo (Wright 2001). Under Fernando III, Castilian becomes the language of the royal chancellery, and Toledo, as the seat of the chancellery linked to the archbishopric, becomes the centre of the Castilian *scripta*, whose influence can soon be observed in documents and legal charters throughout the entire Iberian Peninsula. There has been a lengthy debate on whether the language of Toledo was imposed by decree as a norm for Castile (González Ollé 1996), but a royal decree to this effect supposedly issued by Alfonso the Wise appears to be a later invention. In the fifteenth and sixteenth centuries, as the dialects from Seville and from Old Castile competed to become the Castilian norm, the alleged decree was brought out as proof of the excellence of the Toledan norm.

The fourth phase was achieved in Castile under Alfonso the Wise. Castilian became the language of prose works in various domains (Gómez Redondo 1998), including extensive historiographic, legal and scientific texts. The creation of these works required new clause-linking techniques and extension of the vocabulary. The criteria for this elaboration allegedly stemmed from the personal intervention of the monarch (Solalinde 1915), as is stated in the works. It remains disputed to what extent this is a reflex of stylization or of the inclusion of references intended to give an aura of authority. The criteria for the creation of written Castilian included, amongst others, the rejection of a

Latinizing, Provençalizing or Arabizing language and an attempt to create new expressions by employing the possibilities of Castilian word formation, even in the case of scientific or legal terminology (cf. Niederehe 1975; Bossong 1979; Castillo Lluch 2008). The Alphonsine writing traditions remained the model for the whole of Castile and for supra-regional correspondence throughout the entire Iberian Peninsula right into the fourteenth century. Parallel to this language, a Toledan spoken norm existed, although this lost its significance when the court was moved in the sixteenth century (provisionally to Valladolid and then definitively to Madrid) and the language of Old Castile again took on the leading role (Menéndez Pidal 1962).

3.4.3 Galician and Portuguese

In the field of research on Galician and Portuguese, there is – as is the case for the other Ibero-Romance areas – a lack of extensive scriptological studies comparable to those carried out for France (cf. Monjour 1995). However, in recent decades, considerable advances have been made in terms of the quantity of texts edited and studied. Among others, the studies by Lorenzo (1975), Maia (1986) and Martins (1994) are worthy of particular mention (see the overview by Mariño 2008).

The crucial question about Galician and Portuguese from the Middle Ages until the present day concerns the unity or diversity of the language area: is Galician-Portuguese on the whole a uniform language or are there more or less striking differences between Galician and Portuguese even in this early period? Different phenomena must be distinguished in this context: first, the problem of written uniformity must be differentiated from that of spoken uniformity, and within the written language, the uniformity of literary and non-literary texts must be distinguished. For the spoken language, the question is whether the moment when Galician and Portuguese become separate entities can be fixed, and which factors should be seen as responsible for this separation. An initial diachronic difference can be established between the emergence of the Romance dialects and their expansion due to political factors. Whilst the Romance dialects south of the Minho were overshadowed by Arabic after the Arab conquest of 711, in the north, a primary Romance dialect could develop, which was named Galician as a derivation from the Romance province name *Callaecia*. During the *Reconquista*, the area of Galician dominance spread further south. Following the conquest of Toledo in 1085, as a sign of gratitude for the help of the French knights, and in particular as a tribute to the Abbot of Cluny, Alfonso VI of Castile married his stepdaughter Teresa to Henry of Burgundy, whom he allowed to govern the area south of

the Minho (from 1095 as *Condado de Portugal*). The presumed nephew of Henry, Raimundo, married the king's daughter, Urraca, and ruled Galicia with her. After the death of Alfonso VI in 1109, Urraca's and Raimundo's son became the king of León and Castile. Galicia orientated itself towards the centre, whilst Henry laid the foundation for the separation of Portugal from León, which was finally achieved by his son Afonso Henriques (later Afonso I of Portugal) after several attempts and through skilful diplomacy (particularly towards Rome). In 1131, the monastery of Santa Cruz was founded in Coimbra. Following the battle of Ourique (on 25 July 1140), Afonso Henriques became king, and Portugal, to a considerable degree, independent from León and Castile. The most important date in Portugal's development is 1147, when Afonso Henriques conquered the large and culturally flourishing Mozarabic city of Lisbon. With Lisbon, Portugal gained a new urban centre which paved the way for a linguistic orientation away from the north (cf. Silva Neto 1952:382–95).

There are two different hypotheses concerning the koineization tendencies which mirrored these political events, each of them appearing ultimately to correspond to the historical reality. On the one hand, Lisbon was an important centre whose population included a large proportion of Romance-speaking Mozarabs at the time of the conquest, but on the other hand, the language of the conquerors from the north was Galician. Thus, one can interpret Portuguese as either Mozarabized Galician or as Galicianized Mozarabic. Contrary to the idea that Galician-Portuguese was originally a unified language, Maia (1986) showed in an extensive study that actual differences between the two varieties were already apparent at the time of the conquest of Lisbon, which leads us to conjecture that it was only a question of time before the conquerors' variety, which to begin with probably had diastratic and diaphasic prestige, was 'diatopized' and 'archaized' – in other words, before elements of the urban variety of Lisbon were reinforced and the northern, Galician variety lost prestige in Lisbon. This *desgaleguização* ('de-Galicianization'), to which grammarians refer from the sixteenth century onwards, appears to have been initiated as early as the twelfth century, in the opinion of Serafim da Silva Neto (cf. Silva Neto 1952; 1961; Monjour 1995). During the following period, the indisputable political centre, Lisbon, became the linguistic centre of gravity for the Portuguese *koiné*. By contrast, the north was separated from the south, and the Minho increasingly came to be a linguistic boundary. Galician was more and more influenced by Castilian and has only far more recently undergone its own koineization process (Kabatek 1996).

Although a particular differential present-day perspective tends to highlight the contrast between Galician and Portuguese in the Middle Ages, we are clearly dealing with very slight differences that are at least matched by substantial areas where the varieties are identical. Morphologically and syntactically speaking, Galician and Portuguese are so similar even today that one can assume that the differences in the Middle Ages were predominantly lexical and phonetic in nature. When adopting the language of the north, the Mozarabs of Lisbon would have retained their phonetic habits at least partially, although these can only be hypothetically reconstructed and are not reflected in the written language, since they date from the 'first phase', in which there were no written Romance texts. The first clearly datable, although still isolated, written Romance evidence (phase II) stems from the first half of the thirteenth century. Thus, Galician-Portuguese is a relatively 'late' written Romance language. In the second half of the thirteenth century, Romance *scriptae* came into being in various monastic centres (phase III). The texts from the northern monasteries display, to some extent, differences in comparison to the texts originating from the southern monasteries (Maia 1986; Monjour 1995; Bello Rivas 2001), but only a very vague relationship can be discerned between the orthographic differences and the presumed spoken differences (Börner 1976), especially since the monasteries had very close contact with one another. Towards the end of the thirteenth century, the fourth phase begins, above all as a result of the reception of the Castilian Alphonsine Renaissance brought about by the nephew of Alfonso the Wise, Dom Dinis. With the stabilizing of the chancellery language of Lisbon and the introduction of Provençal graphemes (<lh> and <nh> for [ʎ] and [ɲ], amongst others), the foundations were laid for an independent Portuguese national language. Alphonsine texts, which are in clear contrast to the southern tradition orthographically and in part linguistically, were also translated in Galicia. These texts could not, however, initiate a lasting tradition there, since an increasingly strong Castilianization asserted itself as a result of political dependency.

Rather as in the case of Occitan, a Galician poetic language with supra-regional significance developed early on, largely independent of the documentary tradition. The Castilian king Alfonso the Wise is considered to be the most famous Galician poet. With the *Cantigas de Santa Maria*, he produced a significant literary work, whilst also promoting the dissemination of Castilian prose. This is a further example of the compatibility of different written languages for different purposes that was widespread in the Middle Ages (Beltrán 2005).

Three principal goals remain for research into medieval Galician and Portuguese: first, the edition of extensive collections of unedited medieval documents; second, systematic scriptological analysis of the entire medieval written corpus; and third, diachronic interpretation based on this type of extensive analysis of documents, including references to historical and other additional information.

Contact and borrowing

MARIUS SALA

1. Preliminaries

The subject of contact and borrowing is perennially important in Romance linguistics. After a period in the early years of the discipline when, in order to explain the evolution of the Romance languages, liberal appeal was made to substrate or superstrate influences (and the differences between languages were explained in terms of the substrate–superstrate distinction), the appearance of structuralist methods meant that the scope for internal explanations was gradually expanded. Over recent decades it is perhaps sociolinguistic research which has made the greatest contribution in this domain (e.g., Labov 1994). Unlike structuralism, which accords a major role to internal factors, the sociolinguistic method's point of departure is that language is heterogeneous and variable, and that it is in this variability that the causes of change should be sought.

Indeed, the past half-century has seen the appearance of numerous special studies dealing with the mechanisms of bilingualism or multilingualism which form the basis of linguistic contact and are considered by some (Martinet 1953; Jakobson 1963) as the fundamental problem of all linguistics. Of particular note among these studies are Weinreich (1953; 1968). The most recent synthesis for the Romance domain is Sala (1998a).¹

2. Contacts

The following analysis of the various types of linguistic contact obviously focuses on the situation in Romance, but necessarily mentions research on

1 See also the various studies contained in *LRL* (VII) and references therein; also Posner and Green (1993).

other linguistic areas, especially where they are fundamentally relevant to Romance.

2.1 *Factors and types of linguistic contact (direct and indirect)*

The factors which may stimulate or restrict linguistic contact are varied. Mackey (1976) rates the number of factors which may be at work in contact between languages at over 100. I shall make a maximally simple distinction between two categories: the extra-linguistic and the structural.

It is generally considered that extra-linguistic factors in particular determine and stimulate contact between languages. Some adherents of this idea, among them Thomason and Kaufman (1988), are excessively trenchant in asserting that the sociolinguistic history of speakers is the main determining factor in the linguistic outcome of language contact, and that linguistic factors are relevant but strictly secondary. In their view, the use of linguistic criteria to explain changes caused by language contact fails once simple analysis gives way to prediction. Contact may be within the same territory (direct contact: mixture of populations, cohabitations of varying duration), between different territories (indirect contact: cultural, economic and political relations), or even between different cultural spheres, as in the case of learned Latin influence on Romance. Direct contact was analysed in detail by Uriel Weinreich and Einar Haugen who, taking as their starting-point the descriptive and comparative analysis of the formal elements of which the source form and the borrowing constitute an integral part, go on to analyse borrowing as an interference phenomenon between languages in contact, with regard also to the way borrowings are integrated into the formal structures of the host language. Indirect contact, usually in the written language, does not depend on a state of bilingualism, but on ephemeral bilingual contexts, and appears in educated individuals. This type of contact has been analysed in various works, notably Meillet and Sauvageot (1934), Vidos (1965) and Hope (1971).

The distinction between the two types of language contact, on which Hope (1962–63; 1964; 1965) insisted, is crucial because in the case of direct contact, which leads to the appearance of a stage of bilingualism, the results of language contact are generally much more important than in indirect contact. In the case of bilingualism (by Weinreich's classic definition of 1953 'two languages will be said to be IN CONTACT if they are used alternately by the same person') we observe deviations from the norms of both languages, called interference phenomena. This means that a bilingual who uses two languages, A and B, introduces into the way he or she speaks language A facts which pertain to language B. With time, such phenomena become established and are no longer

a function of bilingualism. Interference phenomena are a matter of the introduction of extraneous elements into the phonology, morphology and syntax of the language, and into some areas of vocabulary, such as terms for kinship, colours and the weather, and they involve a reorganization of the various older distinctive oppositions of the linguistic system in question. At the level of the speech community as a whole, a prolonged state of bilingualism may lead to abandonment of one of the languages (cf. the fate of Occitan in southern France today); in the case of the substrate, bilingualism ends in the disappearance of the 'older' language. Indirect contact involves superficial contact between two languages (usually one of them is a major international language) with consequences for the lexicon and, to a lesser extent, syntax. It primarily affects certain styles (scientific, advertising) of the standard language, with no effects on the system overall, except via stylistic interference.

The distinction between written and spoken language is often mentioned together with that between direct and indirect contact. But the two are not perfectly parallel. Direct contact is generally oral, and indirect contact is generally written, contact, although there are exceptions. In this latter regard we can mention the influence of Greek on Latin (Calboli 2009), while Latin influenced the translation of the Vulgate Bible and other religious texts. The Judaeo-Spanish variety Ladino, used in Biblical translations following a Hebrew model, has been considered a 'calqued' language (*langue-calque*; Séphiha 1972). Appel and Muysken (1987) consider that these influences are restricted to superficial phenomena. Finally, Muljačić (1978) shows how the major standard Romance languages have influenced Slavonic languages via certain 'cultural' loans which do not presuppose the existence of any bilingual populations.

Extra-linguistic factors may confer equal or different status on languages in contact. When their status is different, one of the two languages in contact is generally considered to be in a dominant position, which contributes to its influence over the other. Bilingualism is governed by power relationships in that, because of extra-linguistic factors (economic, cultural, political, or simply numerical, superiority), one of the two languages in contact enjoys a privileged position and prestige (Iordan 1976). The notion of prestige is stressed by the French sociological school, and especially Vendryès (1968). There has been much debate about terminology in this case ('prestige language', 'superior language', 'dominant language'). For us, the distinct *status* of languages is determined by social value, in other words, their capacity to be used as a means of communication. The language used in a greater number of contexts has a greater social value than the other. At the opposite pole we have a

language that is only spoken at home (Istro-Romanian, Judaeo-Spanish, the Friulian used in Romania; cf. Iliescu 1972), and therefore has an 'inferior' status. In the history of all the Romance languages, the Latin from which they originate had a superior status with regard to all the substrate languages. This can sometimes lead to the abandonment of 'inferior' languages. Elizaincín (1981) emphasizes that a distinction needs to be made between (a) the coexistence of high-prestige literary languages (Spanish and English shared between Mexico and the USA, Spanish and Portuguese shared between Uruguay and Brazil, etc.) and (b) the coexistence of a literary language with another language of inferior status. In other cases the social status of languages may be elusive because no unitary and absolute criteria are available. For example, one may say that for a native of the Ecuadorean Sierra, Spanish has a 'superior' position (it is the official language used in many contexts), but not when this individual uses Quechua in his native community – whence the high number of loans from Quechua in the Spanish spoken by natives of the Sierra. Important in this regard are the views of de Granda (1988) on the status of languages in contact in Paraguay (Spanish and Guaraní), where we find languages that influence each other and have opposite values according to context of use. With regard to the languages spoken in Spain, 'interference is almost entirely unidirectional: the influence of Castilian on the other linguistic systems is overwhelming' (Blas Arroyo 1991).²

It follows from the above that not only does the 'dominated' language undergo the influence of the 'dominant', but also the reverse, so that one may speak of reciprocal influences (Zamora 1977). The clearest example is the influence of the substrate languages on Latin in the emergence of Romance. A more recent case is that of the influence of Romance languages in Africa, analysed in various studies by Willy Bal, who stresses the opposite phenomenon, with Romance languages taking 'exotic' loans from African languages (Bal 1979). Matters are more obvious where vocabulary is concerned: Valkhoff (1931) speaks of affectively motivated 'borrowings of convenience' and 'borrowings of necessity', the latter occurring when the host language has no equivalent for terms in the source language. The former can occur even when the source language has inferior status.

There are even rarer cases of contact between languages of identical status. In this case, there are two speech varieties from regions at a linguistic

2 In general, in considering the nature of borrowing, it is important also to bear in mind the distinction made by van Coetsem (1988; 2000) and Winford (2003; 2005; 2010) between borrowing (transfer due to recipient-language agentivity) and imposition (transfer due to source-language agentivity).

boundary which have the same status: both are used in everyday conversation and are regional variants of standard languages, neither of them being used as a state language (see Zawadowski 1961).

The extra-linguistic factors which act as a brake on contact between languages are psychological and social (Récats 1934; Rubin 1968; Alvar 1978). Some works on bilingualism emphasize a linguistic community's attachment to its traditional language which is felt to be symbolic – so-called 'language loyalty' (Fishman 1966). A well-known phenomenon in the history of many languages is puristic rejection or elimination of foreign loans.

Structural linguistic factors are more rarely invoked as stimuli or barriers to language contact. Some linguists, especially in the early history of Romance linguistics, do not even mention them. We owe to Roman Jakobson (1938) the theory that a language accepts foreign structural elements only when they correspond to that language's own evolutionary tendencies. This theory is taken up by Weinreich (1953), who specifies that in this case language contact has the role of triggering or accelerating independently evolving phenomena. Cassano (1976) has repeatedly argued the opposite view. Similarly, it is held that two languages that are genetically related or have closely similar structures are more subject to mutual influence than two genetically or typologically unrelated languages (Sandfeld 1938; Mackey 1976; Thomason and Kaufman 1988). An interesting case is the *dialecto fronterizo* between Spanish and Portuguese analysed by Rona (1959) and Elizaincín (1973; 1976; 1988). On the other hand, Judaeo-Spanish does not represent an extreme case, even if some authors (Weinreich 1956; Wexler 1981; 1988) consider it a 'fusion language' in which the Hebrew element occupies an important place. This theory is criticized in Bossong (1987), Busse (1991) and Sala (1998b), where it is shown that the Hispanic essence of Judaeo-Spanish is conserved despite the Hebrew influence.

Different modules of language are differentially susceptible to contact: although borrowing can affect all subdomains of language, the results are more evident in some than others. Morphology is less liable to change under foreign influence (cf. Meillet 1951; Weinreich 1953; Haugen 1953). The idea appears in a more nuanced form in a series of works (e.g., Mackey 1970), where three levels are established (lexicon, grammar, phonology) showing differing degrees of linguistic interference: the simpler the system, the greater its stability. Zamora (1977) points out that in this connection we must also bear in mind whether one language is the mother tongue (L1) and the other is acquired/learned later (L2). García (1976) reaches the same conclusion from an analysis of contact between Galician and Castilian. As for the weak points of

the system, interference may contribute to a restructuring of less organized aspects of the language. Overbeke (1968) points out that situations of prolonged contact confer a more analytic character on the languages affected. The relative frequency of a given element in speech is also important (Blas Arroyo 1991; Weinreich 1953): the greater the use of a given morpheme or construction, the greater its chances of being transferred into another language.

Linguistic factors tending to hold back the production and spread of interference are obvious. First, there may be resistance due to the bilingual's need to be understood by his interlocutors and to conform to their linguistic norms (Weinreich 1953; Elwert 1960). There are also certain languages (such as Icelandic) which show great difficulty in borrowing (especially in vocabulary), the relevant speech community preferring to use indigenous linguistic material – although in general Romance languages do not display any particular aversion to borrowing.

2.2 A diachronic³ typology of contacts

A distinction must be made between contacts in the period of the emergence of the Romance languages (substrate and superstrate) and later contacts (adstrates such as Slavonic, Arabic, those between Romance languages, Greek, English, German, other Germanic languages, Slavonic, Turkish, Romany, indigenous American and African languages). The former occurred as part of a process of bilingualism⁴ and lasted for many centuries.

2.2.1 Substrates

The Romance substrate represents the totality of the elements which penetrated Latin from the languages of peoples conquered by the Romans who adopted Latin and abandoned their own languages. The concept of substrate proclaims (together with substrate theory) the influence of ethnic factors on the evolution of Latin and the diversification of the languages of the România.

While the influence of pre-Latin languages is mentioned by linguists and philologists of the first half of the nineteenth century, the foundations of scientific enquiry in the Romance domain were laid by Graziadio Isaia Ascoli (1881) in his *Lettere glottologiche*. With the idea of substrate, the ethnic

3 For a description of the various contact situations observable in the modern languages, see also Andreose and Renzi, this volume, chapter 8.

4 The nature of such bilingualism may have varied according to circumstance. Etruscan speakers will have become bilingual in Latin, but it is unlikely that the conquering Romans became bilingual in Etruscan. On the other hand, conquering Germanic speakers will have become bilingual in the more prestigious Latin/Romance, but it is not necessarily the case that Romance speakers became bilingual in Germanic.

factor enters the explanation of the evolution of Latin and the linguistic diversification of the *România*: learned or acquired by numerous and linguistically very different peoples, Latin underwent a series of modifications, which were not the same across the whole Empire; some of these changes have been explained as an effect of interference between the languages of these peoples and Latin, determined by the process of bilingualism or imperfect acquisition and transmission of the external language.

The effects of the substrate are acknowledged in the lexicon, and in toponymy or anthroponymy (linguists may disagree over the etymology of certain terms), but are controversial in phonology: Neogrammarian scholars are sceptical because of the poor attestation of substrate languages and the consequent impossibility of verifying substratist hypotheses. The structuralists deny or play down the influence of the substrate in phonetics and explain linguistic changes through the internal influence of Latin; attempts have been made to reconcile the theory of the substrate with the structural approach, through the interaction of internal and external factors. Substrate explanations in phonetics have appealed to the idea of adaptation of newly learned Latin to the articulatory and auditory basis specific to the learner's native language. Nobody any longer believes in the hypothesis that there are latent tendencies of extinct languages whose effects can be delayed by several centuries, after the disappearance of the prestige of the Latin model.

One of the difficulties in researching the substrate is the fact that none of the substrate languages have been preserved to this day, although some modern languages are related to old substrate languages (the Celtic of the British Isles and Brittany, Basque and, for the Thraco-Dacian substrate, Albanian). No really substantial text survives in any substrate language: in Latin and Greek writers there are attestations of words, placenames and personal names, short and obscure inscriptions, and some glossaries. Note, however, the evidence for Punic influence on the Latin of North Africa and for Hebrew or Aramaic influence on the Latin of Palestine, Syria and Mesopotamia (see Gonzalo Rubio 2009). Comparative Romance evidence is also used for substrate explanation of some linguistic phenomena.

A substrate may be pre-Indo-European: Alpine, Aquitanian, Basque, Mediterranean (Etruscan, Iberian, Ligurian, Sicanian), paleo-Sardinian (Nuragic), Punic, Raetic; or Indo-European: Celtic, Greek, Illyrian, Italic (Oscan-Umbrian, Sicilian), Messapic, Thraco-Dacian, Venetic. A distinction is made between a 'primary' substrate of the Romance languages, comprising the influences exercised on them by the languages of the pre-Latin populations of Italy (these influences were assimilated by Latin and spread throughout the Romance

languages; it is also supposed that these languages had a role in the linguistic fragmentation of Italy), and a 'substrate proper' due to the influence of the languages of the pre-Roman populations in the rest of the Empire.

Not all scholars accept the role attributed to substrates, in that what some hold to be substrate effects others attribute to the superstrate or internal linguistic evolution. The Romance languages for which substrate explanations have most often been invoked are French and Romanian. In phonetics, articulatory tendencies (palatalization in Gallo-Romance attributed to the Celtic substrate, vowel centralization in Romanian attributed to the Thraco-Dacian substrate), leading to modifications of the system through the development of articulations unknown to Latin (Fr. rounded front vowels [y], [ø]; Ro. centralized vowels [ə], [ɨ]) or modifications affecting the norms of realization of the system (lenition of intervocalic occlusives in the western România, the evolution of certain consonant clusters, such as the palatal development of Lat. /kt/ in the western România vs. the labial development as /pt/ in Romanian: cf. Pt. *feito*, Sp. *hecho* 'done, fact' vs. Ro. *fapt* 'fact' < Lat. *factum*). In morphosyntax, reorganizations of the inflectional system (fusion of the genitive and dative into a single form distinct from the nominative-accusative in Romanian), word order (leading to postposition of the definite article in Romanian) or peripheral modifications (traces of the vigesimal counting system in French). In word formation, suffixes (more frequently attested in toponyms). The number of words attributed to the substrate varies from a few dozen to a few hundred (certain, probable, possible), denoting mainly features of the landscape (Ro. *mal* 'bank', Fr. *berge* 'bank', Sp. *vega* 'meadow', Pt. *beira* 'bank'), plants and animals (Ro. *brad* 'fir tree', *copac* 'tree', Fr. *sapin* 'pine', Pt. *carvalho* 'oak', *carrasco* (> *carrasqueira*) 'holm oak', Ro. *șopârlă* 'lizard', Fr. *mouton* 'sheep', Sp. *becerro*, Pt. *bezerro* 'calf'), objects to do with occupations (Ro. *grapă* 'harrow', *țarc* 'sheep pen', Fr. *soc* 'ploughshare'; Sp. Pt. *gancho* 'hook'). There are lexical concordances between Romance languages attributable to concordances between substrate languages (Ro. *ciung* 'one-armed, truncated', It. *cionco*, Egð. *tschung*, Frl. *cionc*, Ro. *ciut* 'hornless', *ciută* 'hind', Cat. *xot(e)* 'bird', 'she-kid'). It is noteworthy that there are substrate words of different form and origin in different Romance languages denoting the same thing, e.g., Ro. *ciocârlie* and It. (*al*)*lodola*, Egð. *alauda*, OFr. *aloue*, Fr. *alouette*, OSp. *aloe*, Sp. *alondra* 'skylark'.

By extension the term 'substrate' has also been used to denote the later influence of languages on which the Romance languages were superimposed, such as Guanche (the aboriginal language of the Canary Islands), the Amerindian languages for Spanish and Portuguese in America, and the indigenous languages for creoles.

There is a long list of languages which have been invoked as substrates acting in various ways on the lexical and structural development of Romance languages. For many of these (e.g., Ligurian, 'Alpine', Sicilian, 'Palaeo-Sardinian', 'Illyrian', 'Messapic', 'Venetic'), the evidence is frankly so tenuous, or dubious, that there is little point in dwelling on them. Very often one has a list of vocabulary items (often toponyms, geomorphic terms or plant names) of no obvious etymology but distributed geographically in an area which corresponds, more or less, to the territory in which the assumed substrate was spoken, so that a substrate explanation seems appealing but cannot be proved. At any rate, the structural – as opposed to lexical – impact of these alleged substrates on Romance is usually nil (see also Loporcaro, volume I, chapter 3, p. 143). Osco-Umbrian (an Italic variety closely related to Latin) is often claimed to have had some phonological effects on central and southern Italo-Romance. Some Latin and southern Italo-Romance words showing /f/ rather than /b/ or its Romance reflexes, such as It. *tafano* 'horsefly' (for Lat. *TABANUS*), or *atufro* 'October' in parts of Campania and Basilicata, do show the effects of a sound change also characteristic of Osco-Umbrian (see also Loporcaro, volume I, chapter 3, p. 111). The 'weakening' of postnasal consonants (/nt/, /mp/, /nk/ > /nd/, /mb/, /ng/; /nd/ > /nn/, /mb/ > /mm/: **kwando* > *'kwanno* 'when', **kampo* > *'kambu* 'field') found extensively in mainland southern Italy is matched by sound changes known to have occurred in Osco-Umbrian. But see, for example, Varvaro (1979), who shows that in the far south, at least, there is no possible appeal to the substrate.

There are two non-Indo-European substrate languages which are sometimes argued particularly to have affected the structural development of the Romance languages in the relevant regions, namely Basque and Etruscan. They are instructive in a negative way: they rather point up the importance, in any appeal to contact as an explanation for language change, of showing: (a) that alleged conditions for the change really did exist in the substrate language; (b) that the change could not have arisen as a result of spontaneous, internal change; and (c) that substrate influence is overall plausible because the phenomenon at issue is one of a cluster of changes safely attributable to that substrate.

To Basque (almost certainly the continuant of the language of the Aquitanians who had spread along the Pyrenean chain between the first century BC and the first century AD), have been attributed a number of structural features of Castilian and Gascon, notably the development of Latin /f/ generally into /h/ (which, in the more recent history of Spanish, has completely disappeared, although it survives in various southern dialects):

e.g., FARINA > *harina* 'flour', FACIT > *hace* 'makes', FEMINA > *hembra* 'female'. There is a long tradition (reviewed for example in Trask 1997:424–29) of ascribing this change to contact with Basque, in which /f/ is known to have been historically absent. The /f/ > /h/ change would presumably have arisen by Basque speakers failing to acquire the pronunciation of /f/, and substituting it with the acoustically closest alternative in their own phonological system, /h/. But this substrate hypothesis also fares badly against various criteria.

First, an /f/ > /h/ change is by no means unique to Castilian and Gascon. A similar change, independent of any Basque influence, has occurred, for example, in some dialects of Sardinian (cf. Barbagia di Ollolai 'hemmina < FEMINA) and in parts of Calabria in southern Italy. What seems to be involved is a fairly unremarkable articulatory relaxation such that the lips no longer make contact with the upper teeth. There are also mismatches between the relevant phonological details of Castilian and those of Basque. Basque simply did not have an /f/, so one would anticipate wholesale replacement of /f/ by an alternative sound. In Castilian, however, the change looks more like an 'internal' sound change sensitive to the phonological environment in which it operates, for /f/ is conspicuously retained before [w] and [r] (e.g., *frente* 'spring', *frente* 'forehead'). For a detailed account of this development, see Penny (2002:90–93). Moreover, Latin words in /f/ borrowed directly into Basque (presumably during the Roman occupation of Hispania and Gallia) actually show replacement of /f/ not with /h/, but usually with /b/: FRONTEM > *boronde* 'forehead', FAGUM > *bago* 'beech', FABAM > *baba* 'bean'. There are, in addition, geographical and chronological discrepancies. There is no evidence that there was ever a significant, stable, Basque-speaking population in the territory that became Old Castile. In contrast, in the kingdom of Navarre, in which Romance speech unquestionably had been superimposed on a Basque-speaking population, there is actually no evidence of aspiration of /f/. Chronologically, the /f/ > /h/ change appears to have been relatively late, occurring long after Romance speech was well established in Castile.

The 'Basque substrate' hypothesis also suffers from 'isolation': there is nothing else in Castilian which can be ascribed with any certainty to Basque influence.⁵ A deft review of the defects of attempts to ascribe various, mainly lexical and phonological, phenomena to Basque may be found in Trask (1997:415–29): in effect, there is virtually nothing in Castilian which can be reasonably argued to be Basque, with the lonely exception of one or two probable loanwords, such as *izquierda* 'left' and *zarza* 'bramble'. As for

⁵ See also Loporcaro, volume I, chapter 3, p. 151.

phonology, Basque is known to have undergone a variety of distinctive phonological changes before the twelfth century, such as loss of intervocalic /n/, a change of /l/ > /r/ and voicing of word-initial occlusives (cf. Trask 1997:427). If /f/ > /h/ originates in Basque, one might expect this whole cluster of traits to be carried over into Castilian, but it is not (cf. MOLINUM 'mill' > Sp. *molino* (not ***morio*), *casam* 'hut' > Sp. *casa* 'house' (not ***gasa*)). On the other hand, certain other phonological developments in Castilian, which are sometimes ascribed to Basque, might just as well be of independent origin (e.g., /v/ > /b/).

The Etruscans lived in the first century BC in an area of Italy whose extent varied over time: from central Italy, where Etruria itself was located (very roughly modern Tuscany), they reached the Po Plain and the Alps, not to mention southward to Campania and beyond. Etruscan was a non-Indo-European language and apparently a linguistic isolate, which may have ceased to be spoken by the beginning of the Christian era, being replaced by Latin. It is preserved in some 10,000 inscriptions (90 percent of them proper names), only fifteen of which are bilingual or involve parallel texts. It had some lexical influence on Latin (e.g., URBS, POPULUS, CATENA), and survives in some Italian place names (e.g., *Modena*, *Chianti*, *Ravenna*, *Todi*), but there is little evidence of influence on Romance languages, despite attempts by some Romance linguists to argue otherwise.

A phonological resemblance between modern Tuscan dialects and Etruscan led various scholars (most notably Merlo 1927) to attribute its origin to the pronunciation habits of the ancient Etruscans. In the relevant Tuscan dialects there is *fricativization/spirantization* of the historically underlying occlusive consonants /k/, /t/, /p/ in postvocalic position: the phonetic details of such spirantization are variable, but a typical set of outcomes is [h], [θ], [ɸ] (e.g., [le ɸa'θaθe 'hɔtɔ] corresponding to Italian *le patate cotte* [le pa'tate 'kɔtɔ] 'the cooked potatoes'). There is indeed approximate correspondence between the geographical extent of spirantization and the ancient territory of the Etruscans, and there is evidence (mainly from the way in which Greek and other loanwords were transliterated into the Etruscan alphabet), that Etruscans tended to *aspirate* voiceless stop consonants (e.g., [k^h], [p^h], [t^h] for /k/, /t/, /p/). However, the substrate hypothesis is open to serious challenges.⁶

The first is the problem of geographical discrepancy: the territories of the ancient Etruscans and the geographical extent of modern spirantization are not coextensive – spirantization is absent, for example, in most of modern

6 For opposition to it, see Hall (1949), Rohlfs (1966) and Izzo (1972). But see also Giannelli (1983).

Lazio and Umbria, the areas most densely populated by the Etruscans in pre-Roman times. This weakens the substratist argument, but is not a fatal objection, because it could be the case that the alleged substrate effect occurred only in some parts of the original domain of Etruscan (perhaps in those furthest from Rome). Second, there is a chronological discrepancy: the existence of spirantization in Tuscany is actually not explicitly attested from any source before 1525, and then only for /k/: we have no direct evidence for spirantization of /t/ and /p/ before the eighteenth century. Moreover, Corsican – essentially a Tuscan variety imported from the mainland in the Middle Ages – shows no sign of spirantization, perhaps suggesting that the phenomenon is more recent. And if the phenomenon had been present in the Latin spoken in Tuscany, why did it not apply to /k/ before front vowels? The Latin *PACEM* /'pakem/ should have yielded **'pahe by spirantization, but in Tuscan it undergoes the general Romance palatalization and affrication of velars before front vowels (yielding 'pafe, or standard Italian 'patʃe), proving that the consonant was not spirantized in antiquity. There is also structural discrepancy: the evidence from Etruscan inscriptions suggests that the changes attested in Etruscan involved *aspiration* (i.e., a delayed release of closure, resulting in a loud expulsion of air after release) not spirantization. Etruscan aspiration has also been shown not to have been restricted to the postvocalic position, unlike Tuscan spirantization. Evidence that we are actually dealing with two structurally different kinds of phenomenon delivers a serious blow to the substrate account. However, there is yet another difficulty. Tuscan, with other central, and southern, Italian dialects, is subject to general 'weakening' of postvocalic voiceless consonants (see, e.g., Maiden 1995:65f.). 'Weakening' describes any kind of attenuation of the maximal blockage of airflow and cessation of vibration of the vocal chords which is inherent in the articulation of voiceless stop consonants. This usually takes the form of voicing, but in parts of Tuscany voicing can be shown to coexist (even in one and the same speaker) with spirantization. In this case spirantization and voicing could simply be alternative diatopic realizations of the same, apparently internally motivated, central and southern Italian 'weakening' process (see further Giannelli and Savoia 1978; 1979–80).

Despite all these difficulties, one might still attempt to argue that spirantization was somehow an imperfect attempt on the part of Latin native speakers to imitate a kind of aspiration being introduced into Latin by Etruscan speakers who were acquiring it: a change from aspirate to spirant articulation is a simple and plausible diachronic change, attested in the history of many languages. If, as is possible, aspiration was a 'tendency' in Etruscan, with variable

realizations (cf. Agostiniani 1983:58), then it might have come to establish itself preferentially in intervocalic position. And the very fact that weakening can be realized as spirantization precisely in Tuscany might ultimately be assigned to an Etruscan substrate. In the end, what we have here is the kind of dead-end discussion typical of appeals to substrate influence: such influence is not downright impossible, but there is no compelling reason to believe in it, and a number of reasons to doubt it, to which we may add 'isolation' – the lack of any other phonological or structural phenomena attributable to Etruscan influence.

There are also two Indo-European languages (or language-groups) from which there has undoubtedly been some lexical influence on Romance, but for which claims of structural influence are far more difficult, namely Celtic and Dacian/Thracian. Peoples speaking Celtic, a language of Indo-European origin, occupied in the second half of the first millennium BC a vast area from Britain to the Danube, Asia Minor and Sarmatia, from the north sea to the Po Plain in Italy and the south of the Iberian Peninsula (in the central and north-eastern part of the Iberian Peninsula the mixture of Celts and local Iberians led to the formation of a people known in historical sources as Celtiberians).

Celtic languages thought to have been particularly important in the evolution to Romance were Gaulish (spoken in Gaul and Cisalpine Gaul, the territory corresponding to modern France, southern Belgium, western Switzerland and the Po Plain of northern Italy), Celtiberian (central Iberian Peninsula) and the Celtic of Raetia. Attestations of these languages are scarce: some inscriptions, toponyms and anthroponyms, occasional references in ancient authors, and what can be inferred indirectly from the study of modern insular Celtic languages (of Ireland, Scotland, the Isle of Man, Wales, as well as Cornish and Breton). On the basis of some fifth- to sixth-century documents and some linguistic evidence (especially placenames), Celtic languages are assumed to have continued to be used even after Romanization in rural areas of France, Spain and northern Italy until the fourth century, surviving in isolated outcrops until the sixth century or later.

The formation of two distinct Gallo-Romance varieties, French and Occitan, has been attributed to Celtic substrate influence, due to the greater density of Celts in the north than in the south – although the number of Celtic words, about 100, is approximately the same in both languages. The contact of Latin with the language of the Celts (the Celts of northern Italy were defeated by the Romans in 175 BC) led to the diffusion of a series of Celtic words into the Latin vocabulary and its spread right across the *România* (BRACAE 'breeches,

trousers', CABALLUS 'horse', CARRUS 'cart'). Other terms spread over most of the western România, following the Romanization of the relevant provinces (betulla 'birch', cattus 'cat', tinca 'tench'), while a third series is restricted to the Gallo-Romance area (Fr. *briser*, Oc. *brizar* 'break'), or just to northern France (Fr. *boue* 'mud', *brai* 'pitch'), or to Ibero-Romance (especially in toponyms). Contaminations of Celtic and Latin words led to Fr. *braire* 'bray' (**brag-* x RAGERE 'low'), *craindre* 'fear' (**crit-* x TEMERE), *orteil* 'toe' (Gaulish *ordiga* x ARTICULUM 'joint').

Various phonological phenomena have been attributed (see, e.g., Di Giovine 2003) to Celtic substrate influence: western Romance voicing of intervocalic consonants (see Loporcaro, volume I, chapter 3, p. 698, n66) and the resolution of Latin /kt/ clusters as (originally) /jt/; Gallo-Romance fronting of /u/ to /y/ and of /a/ to /ɛ/ (Ascoli 1864; 1882; Loporcaro, volume I, chapter 3, p. 138), or syncope of intertonic vowels (cf. Loporcaro, volume I, chapter 2, pp. 61–64). All of these have perfectly plausible internal explanations and there exists no compelling reason to attribute them to Celtic, especially in the absence of evidence that any of the relevant processes were at work in Celtic at the relevant period in the relevant places (see also Ternes 1998). Also sometimes attributed to Celtic influence is the vigesimal counting system used partially in modern French and more extensively in old French (e.g., OFr. *treis-vinz* 'three score (60)', *dix-huit-vinz* 'eighteen score (360)'): on this see the discussion in Bauer, volume I, chapter 10, p. 552, and references therein.

The Thracio-Dacians,⁷ a people of Indo-European origin, occupied in the pre-Christian era a vast territory from the Carpathians to the Aegean Islands. Contact between the Romans and the Thracio-Dacians occurred both to the north and to the south of the Danube, but initially to the south, where the Romans had conquered all the Balkan Peninsula up to the Danube in the first century BC. Intense contact between Dacians and Romans north of the Danube began with the transformation of Dacia into a Roman province, in AD 107. There is very little direct information about Thracio-Dacian: a few glosses (fifty-seven 'Dacian' plant names mentioned in two treatises on medical botany, the Greek Dioscorides in the first century AD and an anonymous author of the end of the third century known as Pseudo-Apuleius), personal names, tribes, deities, human settlements, hydronyms and names of mountains preserved in ancient authors, in Greek or Latin inscriptions, or on coins (1150 anthroponyms, 900 toponyms), and some still undeciphered inscriptions.

7 Whether Thracian and Geto-Dacian are different languages is debated, but the general view is that they are the same.

To identify Thraco-Dacian words, two major methods are used: (a) comparison of Romanian and Albanian (Pușcariu 1940; Rosetti 1968; Brâncuș 1989), the latter being considered the direct descendant of Thracian;⁸ and (b), the reconstruction of some Thraco-Dacian elements on the basis of comparison with a series of ancient Indo-European languages (Hasdeu 1883; Philippide 1925; Russu 1959; Reichenkron 1963; Poghiric 1969; Ivănescu 2000). Method (a) is more reliable, and produces a list of some eighty to one hundred words shared with the list produced by the other method (e.g., *abur* ‘steam’, *argea* ‘loom’, *balaur* ‘ogre’, *balegă* ‘dung’, *barză* ‘stork’, *brad* ‘fir tree’, *brusture* ‘burdock’). Thraco-Dacian substrate terms generally denote flora or fauna specific to the relevant region or have a more specific sense when compared with synonymous terms inherited from Latin (*sâmbure* vs. *os* ‘pip’, *scăpăra* vs. *aprinde* ‘kindle’, *bască* vs. *lână* ‘wool’). In phonology, the Romanian changes of Latin /ks/ and /kt/ to /ps/ and /pt/ (*COXA* > *coapsă* ‘hip’; *PECTUS* > *piept* ‘chest’), and the change of intervocalic /n/ to /r/ in old Romanian texts (*LUNA* > *lu(n)ră* ‘moon’), as well as rhotacism of intervocalic /l/ (e.g., *SOLEM* > *soare* ‘sun’) and various other changes, have been attributed to substrate influence, but are all equally explicable as internal developments (cf. similar developments elsewhere in Romance such as /l/ > /r/ in Ligurian (e.g., *mura* ‘she-mule’)).

A feature sometimes assigned to the Thraco-Dacian substrate (Kopitar 1829; Sandfeld 1930; cf. Graur 1929; 1967) is one that distinguishes Romanian among Romance languages yet is also constitutive of the language’s membership of the Balkan ‘Sprachbund’, namely the postposition of the definite article to the noun (see Ledgeway, volume I, chapter 8, pp. 415, 730), as in *casa*, *casele* ‘the house(s)’, *calul*, *caii* ‘the horse(s)’ (vs. non-definite *casă*, *case*, *cal*, *cai*) – a characteristic shared, for example, with Bulgarian and Albanian. In fact there is no relevant evidence from the substrate to support this view, and the phenomenon has been explained as the result of an internal development: through the use of a reflex of Latin *ILLE* as a ‘linking particle’ between noun and postposed adjective (Gamillscheg 1936), phrasal rhythm (Pușcariu 1937), or a mechanism such that the article originally modified a following adjective but was then reanalysed as enclitic to the preceding noun (Graur 1937; Drăganu 1938).

8 There is no way of knowing, however, whether a term is inherited from the substrate, or borrowed from Albanian.

2.2.2 The Germanic superstrate

Germanic languages are considered a superstrate for western Romance languages, whose influence is manifested to differing degrees in lexical, toponymic and anthroponymic loans. Possible Germanic influence on Romanian is a matter of controversy, ranging from those who deny any Germanic influence to those who admit a number of Germanic words, often bundling together Germanic words which first entered Latin, Slavonic or Hungarian as well as late Germanic loans. There is no Germanic influence in Sardinian (except for words transmitted via Italian or Spanish) and Dalmatian.

There is a first layer of early loans, due to the first Latin-Germanic contacts (second to fifth centuries). The Rhine Valley, a major frontier of the Empire which formed the eastern edge of the Roman province of Germania, was at that time one of the main areas of contact between Romans and Germanic peoples. The latter were also present as mercenaries in garrisons in the territory of the Empire, protecting the frontiers as *foederati*, or established as colonists especially in areas laid waste by earlier raids, or were trading partners of the Romans. For discussion of the criteria for identifying this first layer of loans, see Brück (1913), Gamillscheg (1948) and Wartburg (1950). There is just one pan-Romance word in this class, the reflex of *sapo* 'soap' (attested in Latin in the first century): e.g., Sp. *jabón*, Fr. *savon*, It. *sapone*, Ro. *săpun*. Other early Germanic loans attested in Latin are *taxo* > It. *tasso*, Fr. *tais*, Sp. *tejón* 'badger'; *supa* > It. *zuppa*, Fr. *soupe*, Cat., Sp. Pt. *sopa* 'soup'; *helmus* > Fr. *heaume*, Oc. *alm*, Sp. *yelmo*, It., Cat., Pt. *elmo* 'helmet'; also **uerra* > Fr. *guerre*, It., Cat., Sp., Pt. *guerra* 'war'.

The second phase of Germanic influence dates from after the fifth century, and is a consequence of the settlement of some Germanic tribes on Roman territory (both before and after 476), either with the agreement of the Romans or due to invasions. The consequences of Germanic influence on the linguistic history of the România were of an extra-linguistic nature: the isolation of some Romanized areas (the breaking of the link between Gallo-Romance and Raetia in the sixth century as a consequence of the invasion of the Alamans into what is now Switzerland, and of that between Gallo-Romance and Italy in the sixth to eighth centuries following the Longobard invasions);⁹ the widening of the

9 In some cases the presence of Germanic speakers may have left relatively few reliably identifiable linguistic traces, but may have contributed to the creation of linguistic boundaries. This is the case with the Burgundians (speakers of what was probably an eastern branch of Germanic) who settled around Geneva in 443 founding a kingdom (in 457) which eventually stretched from the south of the modern province of Champagne as far as the *département* of the Alpes Maritimes, with its capital in Lyon. The Burgundians

linguistic differences across the Empire between areas where contact with Germanic was intensive and those where it was limited or nil; the intensification of the differences between Romance languages as a consequence of locally different Germanic influences (a factor considered decisive in the fragmentation of Gallo-Romance) or, conversely, the levelling of incipient linguistic differences (the unifying role of the Visigoths in the Ibero-Romance domain) and the restructuring of linguistic fragmentation, such as the reorientation, from the sixth century, of Tuscan away from central and southern Italy towards northern Italian dialects.

A third layer of Germanic influence (also considered as a superstrate) is that of the old Scandinavian of the Norsemen on French (see, e.g., Ewert 1963:293), largely confined to nautical terms such as *vague* 'wave', *tillac* 'deck'.

Frankish is a west Germanic language considered to constitute a 'superstrate' for French, Occitan, Catalan and Italian. It needs to be said, however, that the historical relation between Frankish and (northern) Gallo-Romance rather calls into question the utility or appropriateness of the notion of 'superstrate'. The Franks were *politically* dominant and prestigious (indeed, they give their name to the country of *France*), but they were relatively few in number and their language was *culturally* subordinate to the infinitely more prestigious Latin/Romance; indeed many Frankish nobles had acquired knowledge of Latin/Romance even before the incursions into Gaul. These considerations constitute an obvious problem for the view, notably developed by Wartburg (1950; [1934] 1971), that the Frankish language imparted to northern Gallo-Romance a number of structural linguistic characteristics which came to distinguish it from varieties south of the Loire. Frankish nobles allegedly introduced their native speech habits into Romance (a situation of language shift with imperfect learning), and these were in turn imitated by native Gallo-Romance-speaking nobles, because of the prestige associated with Frankish speakers; from the nobility the innovations allegedly then percolated into the rest of society. Not only is this claim difficult to prove, but we may ask whether it is even plausible (see Sornicola 1989): how would *imperfect* use of a language more prestigious than Frankish have been so keenly imitated by the great mass of Romance native speakers, especially when it is reasonable to believe

were converted to Christianity and their territory was assimilated into Latinity. They were defeated by the Franks in 534, and thereafter disappear as a historical entity. Burgundian influence has been invoked to explain the emergence of Franco-Provençal dialects distinct from French on the one hand and Occitan on the other. There are a dozen or so words which may be attributable to Burgundian (e.g., *fata* 'pocket' in Swiss Romandy). For possible phonological effects, and other assumed Burgundian lexical forms, see Wartburg (1950), Gamillscheg (1934–36) and Hilty (1968).

that the Franks themselves would have been at pains to learn Latin/Romance as ‘correctly’ as possible?

The first contacts between Romans and Franks occur in the third to fourth centuries, on the eastern boundary of Gaul, where the Franks carry out numerous attacks, settling in groups within the Empire. In the fifth century they are allied with the Romans against the Huns. During the reign of Clovis (in 486), they conquer the Roman province of northern Gaul to the north of the regions under the sway of the Visigoths and Burgundians, and convert to Christianity. In 507 they spread southwards, forcing the Visigoths to migrate into the Iberian Peninsula, and in 534 they conquer the Burgundian kingdom. The colonization begins from the southern part of modern Belgium and is most intense in the far north of France (there are many Frankish placenames in the area of the Picard and Walloon dialects), less so between the Seine and the Loire. Northern France saw the development of bilingualism¹⁰ which appears to have lasted in some areas (the west of the Frankish domain, Neustria) into the tenth century when, even here, Romanization was complete.

There are six to seven hundred lexical loans from Frankish into French,¹¹ grouped into two categories. The first is cultural loans (some 50) imposed by Frankish administration and taken over in the fifth century by late Latin, being diffused in the western Romania by the Latin of the Merovingian and Carolingian administrations, such as *mariscalcus* ‘marshal’, *siniscalcus* ‘seneschal’ (> Fr. *maréchal*, *sénéchal*, It. *maniscalco*, *siniscalco*, Cat. *menescal*, *senescal*), which have survived since the Middle Ages. The second belongs to the sphere of everyday life, with differing geographical extensions. Over the area of Frankish domination as far south as the Loire we find French *houx* ‘holly’, *saule* ‘willow’, *hêtre* ‘beech’ and local equivalents vs. Occitan terms deriving from Latin *RERIFOLIUM*, *SALIX*, *FAGUS*. A further series of terms extends south of the Loire, especially into Occitan, either due to very early borrowing before the Frankish invasion, or to Frankish expansion southward after the sixth century, or to direct borrowing from French: **flaska* ‘bottle, flask’ attested in the Latinized form *flasco* in the sixth century and *flasca* in the seventh; It. *fiasco*, Cat. *flascó*, Sp., Pt. *frasco*.

With the southward expansion of the Merovingian and especially the Carolingian dynasties (sixth to eighth centuries), the annexation of Italy to the

10 Not only were the Franks bilingual in Frankish and Romance, but the minority of nobles in the Frankish kingdom who came from the native Latin-speaking community seem also to have become bilingual.

11 A much longer list and analysis of Frankish lexical loans will be found in Elcock (1975:248–70) and Ewert (1963:292f.).

Frankish kingdom (until 840) and the creation in Catalonia of the Spanish Marches, Frankish elements penetrate Italian and Catalan: for example, It. *dardo* 'dart', *guadagnare* 'to earn', *galoppare* 'to gallop', *orgoglio* 'pride', *schernire* 'to scorn'; Cat. *blau* 'blue', *guanyar* 'to earn', *escarnir* 'to scorn'. Italian *giardino* 'garden' [dʒar'dino] shows the action of a typically northern Gallo-Romance sound change (palatalization of velars before /a/; see Loporcaro, volume I, chapter 3, p. 149), and shows that some Frankish loans into Italian may be more properly described as borrowing of 'naturalized' Frankish words in French.

The French and Occitan lexicon is enriched by military terms, many of which have lost their specific military associations (e.g., OFr. *marchis* Occ. *marques* 'marquis', Fr. *banir* 'banish', Fr. *gant* Occ. *gan* 'glove', Fr. *guetter* Occ. *gachar* 'lie in wait', Fr. *haubert* Occ. *ausberc* 'halbard', Fr. *blessor* Occ. *blessar* 'wound', Fr. *guérir* Occ. *garir* 'heal'). There are also social and political terms (Fr. *alleu* Occ. *alodi* 'freehold', Fr. *fief* Occ. *feu* 'fief', Fr. *trêve* Occ. *treva* 'truce'), and terms from various other domains (hunting, agriculture, cattle-rearing, parts of the body): e.g., Fr. *bois* 'wood', *houx* 'holly', *hêtre* 'beech', *gagner* 'earn' originally 'gather in crops', *blé* 'wheat', *roseau* 'reed', *jardin* 'garden', *haine* 'hatred' and *haïr* 'to hate' (cf. Romance *aimer* 'love'), Fr., Occ. *cresson* 'cress', Fr. *mésange* 'tit', Fr., Occ. *flanc* 'flank', Fr. *échine* 'backbone'). There are also abstract terms such as Fr. *honte* 'shame', Fr., Occ. *franc* 'frank, free', Fr. *riche*¹² Occ. *ric* 'rich', and the names of colours such as Fr. *bleu* Occ. *blau* 'blue', Fr., Occ. *gris* 'grey'.

The influence of Frankish on Gallo-Romance extends notably beyond relatively 'concrete' substantive loans of nouns, and some adjectives and verbs. It affects the 'functional' vocabulary as the source of the French adverbs *trop* 'too' and *guère* 'hardly, not much' (from *waigaro* 'much, a great deal'). The former enters Italian as the quantifier *troppo*. Frankish also affects the lexicon in respect of derivational morphology, providing some derivational affixes such as the negative prefix *mé(s)-* and the derogatory *-ard* (and *-aud*), which have been extremely productive in the subsequent history of French and are freely combinable with Romance words (*mésentente* 'misunderstanding', *vieillard* 'oldster, old man', *salaud* 'bastard, son of a bitch'). In all of these cases the strong expressive associations of these forms ('too', 'hardly', and affixes with negative or pejorative values) may have favoured their absorption into Gallo-Romance, their very 'foreignness' lending them additional expressive force (compare the borrowing of the adverb *very* into English from Norman French *verrai* 'true').

12. Contrast this with Romance *pauvre* 'poor'.

There is also some evidence of lexical calquing: the old French *soi tiers* ‘oneself third’, i.e., ‘oneself and two others’, reproduces a Germanic expression. It is very often suggested that the French indefinite personal pronoun *on* (< *HOMO* ‘person’) is a calque of Germanic *man* (cf. modern German *man spricht* = Fr. *on parle* ‘one speaks’). Nothing excludes this possibility, but it could just as easily be an internal Romance development, as is evidenced by parallel developments in some Italian dialects and Catalan (cf. *hom* ‘one’ vs. *home* ‘man’), and we cannot even be sure that the Germanic construction is not a result of contact with Romance (see also Hunnius 1975:72–74). In any case, this particular development is widely attested across European languages (see particularly Egerland 2010).

Frankish also had a notable effect on Gallo-Romance phonology, but here it is important to distinguish between those influences which are incontrovertibly attributable to Frankish, and those only of allegedly Frankish origin. The former are identifiably Frankish because they are conveyed into Romance *within words of Frankish origin* and involve sounds previously absent from the Romance phonological inventory, most importantly word-initial /h/. This sound survives intact to this day in words of Germanic origin in dialects of Picardy, Wallonia and Lorraine. In French, /h/ has since been lost, but leaves a clear trace both in orthography (e.g., *houx* ‘holly’, *hêtre* ‘beech’, *heaume* ‘helmet’, *hérisson* ‘hedghehog’, *haïr* ‘hate’), and in the fact that such words behave phonotactically as if consonant-initial (e.g., *le hêtre* ‘the beech’, not ***l’hêtre*). The sound /h/ is relatively unusual in the sound inventory of the world’s languages, and is not easy to acquire for adults whose native system lacks it. That it was consistently and correctly borrowed from Frankish into Romance suggests that the borrowing was carried out not by monolingual Romance speakers trying to acquire ‘foreign’ words, but by *bilinguals* with a native (or at least near-native) phonological command of both languages. The word *haut* ‘high’, from Latin *ALTUS* ‘high’, but showing an unexpected initial *h-*, betrays the influence of Frankish *hoh*, of the same meaning, and again appears to be the product of bilingual minds, where a Romance and a Frankish form were both linked to the same meaning. The other phonological innovation from Frankish found in Gallo-Romance is the introduction of /w/ (surviving in some north-eastern Gallo-Romance dialects), which, via a fortition stage /gw/, has emerged as /g/ in modern French: thus *guère* ‘hardly’, *guêpe* ‘wasp’,¹³ *guérir* ‘heal’, *gagner* ‘earn’, *guetter* ‘lie in wait for’ (cf. Italian *guari*

13 As with *haut*, above, we seem to have here a fusion of a Germanic word and a Romance one (< *UESPA*). Similarly *guiche* ‘strap’ seems to represent a cross between Germanic *windan* ‘wind’ and Latin *UTICA* ‘vine tendril’.

‘somewhat, much’ (obsolete), *guarire*, *guadagnare*) – all from Frankish forms in /w/ (*waigaro* ‘much’, *wespa*, *warjan*, *waidanjan*, *wahten*).

The case for regarding other aspects of Gallo-Romance phonology, as the result of the introduction of Frankish articulatory habits, rests on much shakier ground. The pervasive deletion or reduction of unstressed vowels in northern France (e.g., *MATURUM* > *meur* > *mûr* ‘mature, ripe’; *CANTO* > *chant* ‘I sing’; *CAMERAM* > *chambre* ‘(bed)room’) has been attributed (Wartburg 1950) to the effects of a heavy expiratory stress-related accent allegedly characteristic also of Frankish, as has diphthongization of high mid vowels in stressed open syllables, characteristic of northern Gallo-Romance but not southern (Occitan varieties), allegedly the result of an exaggerated lengthening of vowels in stressed open syllables believed to have been characteristic of Frankish pronunciation (e.g., **tela* > *tejlə* > *toile* ‘canvas’, **flore* > *floʊr* > *fleur* ‘flower’). Claims of this kind meet serious problems of geographical and chronological discrepancy: reduction and deletion of unstressed vowels was already under way in Latin; reduction and sometimes deletion of unstressed vowels also occurred extensively in dialects of northern Italy; and within the relevant dialect area, there are pockets with extensive diphthongization of stressed vowels in open syllables. Diphthongization of mid and low vowels is also characteristic of large areas of northern Italy and, for Wartburg’s hypothesis about Frankish influence in Gaul to stand up, we would have to accept his surmise that in northern Italy the same phenomena are due to the linguistic influence of another Germanic language, Longobardic. But we do not really know enough about Longobardic phonology, or the circumstances of linguistic contact between Longobardic and Romance in northern Italy, to know whether this is even plausible. And if the northern Italian case cannot be safely ascribed to Germanic influence, can the northern Gallo-Romance one? In fact, the diphthongization of vowels in stressed open syllables is a phenomenon widely observed across many languages. And ‘exaggerated’, or very distinctive, lengthening of vowels in stressed open syllables – which is in fact the *only* part of the Romance diphthongizations which Wartburg attributed to Germanic – is also widely present in Italo-Romance, well outside the domain of possible Germanic influence. At best we can say that Germanic patterns of pronunciation *might* have favoured and promoted trends already present in Romance. *Perhaps* Romance speakers carried such trends forward, encouraged by the model of Germanic speech patterns, but there are no particular grounds to attribute such changes directly or exclusively to Germanic speech habits.

Claims for Germanic influence at the level of morphology and syntax suffer from similar problems. Often, the best one can say is that existing Romance

tendencies *might* sometimes have been promoted by Germanic parallels. There are no clinching arguments, only circumstantial evidence of geographical correspondence. Old French (cf. Sornicola, volume I, chapter 1, pp. 18–32) retained vestiges of a case system in certain nouns and adjectives, such that a ‘nominative’ (subject) case form was morphologically distinct from other (‘oblique’) forms: e.g., nominative singular *lerre* ‘thief’ vs. oblique singular *larron*. The existence of inflectional case distinctions in Germanic languages has prompted the suggestion (Hilty 1968; 1975) that the old French system owes its survival to Frankish influence. What might favour the hypothesis of Frankish ‘support’ for the existing Gallo-Romance forms is that the two-case system apparently survived longest in northern and north-eastern France as far south as Burgundy, and that the predominance of Frankish proper names among the old French inflectional type *Charles* – *Charlon*, *Hues* – *Huon*, *Eudes* – *Odon*, *Berte* – *Bertain*, *Dode* – *Dodain*, has close inflectional parallels in Frankish (*Hugo* – *Hugun*, *Berta* – *Bertun*) – even if the masculine *-on* and feminine *-ain* inflections are also explicable in purely Romance terms. Arguments against the view that Germanic influence is at work come from the lack of geographical correspondence and from structural discrepancies: the two-case system survived *also in the south* of France beyond the usually assumed area of intensive Frankish influence (Maiden, 2000, even argues for its persistence in early medieval Italo-Romance as well), and in Frankish, unlike French, the inflectional pattern in personal names was stressed on the root in both case forms, whereas in Romance the stress alternates between root and ending.

Germanic languages regularly place adjectives in front of the noun. While this position is also possible in Romance, the norm is for the adjective to follow. Could the fact that, in old French, adjectives denoting *colour* showed a strikingly high frequency of occurrence before the noun reflect Germanic influence on a syntactic possibility independently available in Romance (cf. Rohlfs 1979; Hilty 1975)?¹⁴ The kind of evidence that might lend circumstantial weight to it is the fact that a number of colour terms (e.g., Fr. *blanc* ‘white’, *brun* ‘brown’, *bleu* ‘blue’, *blond* ‘fair-coloured’, *fauve* ‘reddish brown’, *gris* ‘grey’, *sor* ‘reddish brown’) indisputably are of Germanic origin, so that if the expression of colour is a domain in which Germanic influenced Romance lexically, one might ask whether it could not also have done so syntactically. But again one can never definitively prove such a view. We are perhaps on

¹⁴ Interestingly, in her survey of the possible syntactic influences of another Germanic language, modern English, on contemporary French, McLaughlin (2011) finds in her corpus that pre-position of the adjective emerges as a phenomenon which is probably not affected by contact (unlike, say, use of the passive).

stronger ground when we observe that in some modern north-eastern French dialects neighbouring on Germanic ones this adjectival position is generally obligatory (except for adjectives of nationality which follow the noun; see Bernstein 1991).

The emergence of obligatory subject pronouns (in place of the Latin ‘pro-drop’ type in which the subject is normally only expressed through the inflectional morphology of the verb) and the strict positioning, in old French, of the verb as second element in the sentence (e.g., *les deniers prendrons nos* lit. ‘the money will take we’, ‘we’ll take the money’; modern remnants in *Aussi décida-t-il . . .* lit. ‘so decided he’, ‘so he decided’; *A peine eut-il décidé . . .* lit. ‘hardly had he decided’, etc.; cf. Salvi, volume I, chapter 7, §§3.4.1, 3.4.7) correspond to structures also found in Germanic, and prompts speculation that Germanic influence is at work (see Wartburg [1934] 1971; Gamillscheg 1957; Rohlfs 1982). Whether the resemblance is other than coincidental is very difficult to ascertain. Verb-second constructions are encountered extensively in medieval Romance languages, and there is no compelling need to invoke Germanic influence. After all, the word-order possibilities of language are finite and hence it is unremarkable that two languages, whether contiguous or not, might (at some point) share similar or identical word-order patterns. In fact, it would need first to be proved that verb-second constraints generally operated in Germanic at the relevant period, and it would also need to be explained how speakers imported something like V2 from their native language into another. The incidence of obligatory subject pronouns (but with very significant local differences; cf. Hilty 1968:507f.) is indeed found in northern France, Franco-Provençal, Romansh, Ladin, Friulian, northern Italian dialects and, to a limited degree, in Tuscan – all areas for which other Germanic influences have also been invoked. Hilty (1968) argues that a particular structural detail shared between old French and Frankish, the tendency for the obligatory subject pronoun to appear in preverbal rather than postverbal position, might be taken to suggest that Frankish influence promoted a tendency already at work in French. But not the least problem with this account is that (as, e.g., Roberts 1993 shows) the obligatory preverbal subject pronoun does not become fully established in French until the seventeenth century, and is a gradual process affecting different persons of the verb at different rates – making it look much more like a case of internal, spontaneous development. Objections of this kind are raised by Hunnius (1975) (but see also the response from Hilty 1975). Moreover, the overt subject pronoun was very far from being ‘obligatory’ in old French, except in embedded clauses, a detail apparently not shared by Germanic/Frankish,

and the tendency for subject pronouns to appear in preverbal position was no more than that (cf. frequent sequences such as that found in the *Strasbourg Oaths* *si salvarai eo cist meon fradre* 'I will save this brother of mine').

Overall, there is incontrovertible lexical borrowing from Frankish in Gallo-Romance, with some penetration into the 'functional' lexicon. Lexical borrowing brings with it certain phonological innovations. There is much weaker, circumstantial, evidence that the various structural features indigenous to Romance were favoured by contact with Germanic. None of this is consistent with any notion that Gallo-Romance developments have their origin in the 'imperfect learning' of Romance by Germanic speakers, but it points rather to a situation of active bilingualism (or probably 'bilingualism with diglossia'), in which speakers with very good knowledge of both languages may have been influenced by Frankish morphology, syntax and phonology in their use of Gallo-Romance.

Longobard was a western Germanic language, spoken by a population settled in Italy, initially in the Veneto (568) then throughout northern (they gave their name to Lombardy) and central Italy. Some arrived further south, where they formed the duchies of Spoleto and Benevento. The Longobards held sway in Italy for two centuries, until Charlemagne defeated the Longobard kingdom in 774. The last remnant of Longobard domination (Benevento) was occupied by the Normans two centuries later. Longobard linguistic influence is manifest in the Italian lexicon and, to a lesser degree, in that of Friulian and Romansh. Italian retains some 300 words from Longobard, and Longobard is the major Germanic influence on Italian. These words belong to a variety of onomasiological domains: the military (*strale* 'arrow', *briccola* 'catapult', *spalto* 'bastion'), the home (*balcone* 'balcony', *palco* 'floor(ing)', *banca* 'bench, bank', *scaffa* 'shelf', *scranna* 'high-backed chair', *rosta* 'bundle of branches'), domestic instruments (*gruccia* 'crutch; clothes hanger', *spranga* 'bolt, bar', *greppia* 'manger, crib', *trogolo* 'trough', *palla* 'ball'), the human body (*guancia* 'cheek', *schiena* 'back', *nocca* 'nuckle', *milza* 'spleen', *anca* 'hip', *stinco* 'shin'), animals (*stambecco* 'ibex', *taccola* 'daw', *zecca* 'tick'), horse-breeding (*guidalesco* 'withers', *predella* 'rein', *staffa* 'stirrup'), the soil (*melma* 'mud, slime', *tonfano* 'deep part of river'), agriculture (*bara* 'litter, bier', *bica* 'rick', *grumereccio* 'second-crop hay', *stollo* 'wooden pole for haystack', *sterzo* 'steering'), woodlands and the uses of wood (*gualdo* 'wood', *spaccare* 'chop, split', *sprocco* 'sprout', *stecco* 'stick'), verbs indicating concrete, technical and basic actions ((*im*)*bastire* 'baste', (*s*)*gualcire* 'crease', *spruzzare* 'spray', *strofinare* 'rub', *baruffare* 'scuffle', *graffiare* 'scratch'), abstract terms (*scherno* 'scorn', *smacco* 'blow, humiliation', *tanfo* 'stench'), adjectives (*gramo*

‘wretched; sickly, feeble’, *ricco* ‘rich’, *stracco* ‘exhausted’). The majority of Longobard terms are regional: *braida* ‘meadow’, *brera* ‘meadow’, *bro(v)ar* ‘scald’, *godazzo* ‘godfather’, *stoa* ‘mare’ are limited to northern dialects; *bica* ‘rick’, *chiazzare* ‘stain’, *chionzo* ‘thickset’, *federa* ‘pillow case’, *gruccia* ‘crutch; clothes hanger’, *lonzo* ‘droopy’, *strozza* ‘gullet’ to Tuscan; *lefa* ‘female boar’, *luffò*, *uffò* ‘thigh’ to central-southern dialects. Many Longobard terms intimately tied to the Longobard administrative and legal system have effectively disappeared (e.g., *guidrigildo* ‘blood money, wergild’). Common Lombard elements in placenames are: the ending *-engo/-ingo* (e.g., *Pastrengo* – combining a reflex of Latin *PASTOR* with the ending), and *far(r)a* ‘group of families’ (e.g., *Fara d’Alpago*).

Gothic (an east Germanic language) is considered a superstrate for Occitan, Catalan, Spanish and Italian. In 270 the Goths split into the Visigoths (western Goths) and Ostrogoths (eastern Goths). The first contact between Romans and Goths took place in the third century on the eastern frontier of the Roman Empire. The Visigoths besieged and conquered Rome in 410, and established themselves in south-west France from 418 to 507. Having been defeated by the Franks, they migrated to the Iberian Peninsula where they founded a kingdom with its capital at Toledo. They were assimilated into the Romanized population and disappear from history with the Arab conquest of the Iberian Peninsula (from 711). The Ostrogoths settled as *foederati* in Pannonia in 308. In 405, together with other Germanic tribes, they attacked northern Italy, but were beaten by the armies of the Western Roman Empire. Under King Theodoric they attacked Italy again, defeating Odoacer (493). Theodoric was recognized by Zeno, Emperor of the Eastern Empire, as his deputy in the Western Empire. The Goths installed themselves as masters of Italy, founding the Ostrogoth kingdom with its capital at Ravenna – the most powerful barbarian kingdom of the time in western Europe (494–555). Ultimately defeated by the Longobards, they were assimilated to them and into the Romanized population.

The earliest layer of Gothic words penetrated Latin in the initial period of contacts between Latin and Gothic. They are not easy to identify, because at the same time similar words from various different Germanic languages entered Latin. The same problem applies to Occitan, Spanish, Catalan, Portuguese and Italian, because various Germanic influences (especially Frankish) were simultaneously at work on the speech of the Romanized peoples of western Romània, introducing words having the same etymological base. However, the restriction of some term to Ibero-Romània, southern France and Italy may constitute an argument in favour of its Gothic origin. The extent of Gothic influence on the occupied territories is limited: some fifty in Occitan, and some

seventy (fifty of them limited to dialects) in Italian, reflecting the brevity of Latin–Germanic bilingualism, and about thirty in Spanish and Portuguese, because by the time they reached the Iberian Peninsula the Visigoths were partially Romanized.

Gothic influence is limited to the lexicon, placenames and personal names. It usually has a concrete nature, referring to aspects of everyday life (home, means of transport, trades). In Occitan, Spanish and Portuguese we find terms referring to military life and administration which are lacking in Italian. Words attributed to Gothic are: *bandwô*, Latinized as *bandum*, pl. *banda* > It., Occ., Sp., Pt. *banda* ‘band’; **gasalha* > Occ. *gasalha* ‘companion’, Sp. (a)*gasajar*, Pt. *agasalhar* ‘to receive as a friend’; **haspa* > Sp., Pt. *aspa*, Gsc. *aspo*, It. (n)*aspa* ‘reel’; **taikn* > Occ. *tana*, *tacon*, Sp. *taco* ‘swear word’, It. *tacca* ‘stain, mark’. Words existing only in the Iberian and Italian areas are: *triggwa* > Sp. *tregua*, Pt. *trégoa*, It. *tregua* ‘truce’; **grims* > Cat., Sp., Pt. *grima* ‘sadness, horror’, Lombard *grim* ‘angry’. Limited to Iberia are: **snôbô* > Sp. *eslabón* ‘link’, **glova* > Sp. *lua*, Pt. *luva* ‘glove’.

2.2.3 Slavonic

Romance languages in general show some lexical borrowings from modern Slavonic languages, usually as a result of cultured and written transmission (e.g., Fr. *mazurka*, It., Sp. *mazurca* from Polish; Ro. *balalaică*, It., Sp. *balalaica*, Fr. *balalaïka* or Ro. *stepă*, Sp. *estepa*, It. *steppa*, Fr. *steppe*; Ro. *bolșevic*, Pt. *bolchevique*, It. *bolscevico*, Fr. *bolchévique*, etc., all from Russian). In the case of Romanian there were two periods of direct contact with Russian speakers. The first period (1829–53) saw the borrowing of some words (*cazon* ‘harsh, militaristic’, *pojarnic* ‘fireman’, *polcovnic* ‘colonel’); some Russisms belonging to administrative language (e.g., *pameșnic* ‘proprietor’) rapidly disappeared. In the same era some French neologisms penetrated Romanian through the intermediary of Russian, for example with the unstressed ending *-ie*: military terms (*artilerie*, *cavalerie*, *infanterie*), or administration (*administrație*, *constituție*). Russian also transmitted to Romanian some Asian words (*ceai* ‘tea’, *hoardă* ‘horde’) and names of countries (e.g., *Anglia*, *Bulgaria*, *Norvegia*). The second phase, after 1944, brings a few words (*agregat* ‘agrégé’, *combinat* ‘industrial complex’, *cursant* ‘student’, *tovarăș* ‘comrade’) but also many phraseological calques (*activ de partid* ‘party activist’, *examen de stat* ‘state examination’, *gazetă de perete* ‘wall newspaper’), or names of firms formed acronymically using (usually) the first syllables of the component words (see Bauer, volume I, chapter 10, p. 561f.) (e.g., *sovrom* – a kind of joint Soviet–Romanian company set up after the Second World War). In the Romanian of the Republic of

Moldova (formerly the Soviet Socialist Republic of Moldova) there are more numerous loans and calques (*maladeț* 'young man, lad', *seiceas* 'now') and syntactic calques.

Economic and political contact between Romanians (and especially Moldovans) and Poles occurred particularly in the fifteenth to eighteenth centuries – when many Moldovan boyars studied in Poland, and many Poles migrated to Moldova. Polish terms in Romanian belong to rather restricted domains: the army (*husar* 'hussar', *zamecă* 'fortress'), social life (*pan* 'gentleman', *seim* 'parliament'), fauna (*clapon* 'capon', *dulău* 'hound'). Polish terms are not popular, and many (outside the Moldovan domain) have disappeared. Contact between Ukrainian and Romanian began in the twelfth and thirteenth centuries (according to some, eleventh and twelfth centuries) in the north of the Romanian territory of Maramureș and in old subcarpathian Russia and continues to this day, most intensely in northern Moldova. Some Ukrainian elements (*boroană* 'harrow', *scripcă* 'fiddle') are hard to distinguish from Russian. The main influences occur in the domain of household objects (*horn* 'chimney', *prostire* 'sheet'), agriculture (*hrișcă* 'buckwheat'), fauna (*hulub* 'dove'), food (*horelcă* 'cheap brandy', *hrib* 'boletus mushroom'); most words are regional (*balie* 'washing tub', *harbuz* 'watermelon'). There may be some phonetic influence on the Maramureș dialect ('hard' pronunciation of the affricates /tʃ/ and /dʒ/ principally manifested in a centralizing effect on following front vowels) and in the dialect of Țara Oașului a velarized pronunciation of /l/ (see Urițescu 1984:393f.).

Contact between Romance and Slovenian occurred due to the fact that in the Middle Ages many Slovenian colonies appeared in north-eastern Italy because of the resettlement of wide areas devastated by natural calamity and invasion. Slovenian words thus found their way into Friulian and Venetian (e.g., Frl. *coșe* 'type of easel', *britule* 'knife', *komat* 'horse collar', *cespe* 'plum').

Contact with Bulgarian is limited to Romanian (and to a lesser degree the Judaeo-Spanish of Bulgaria). The Bulgarian people, who appeared after the proto-Bulgars were assimilated by the Slavs, have been in permanent contact with the Romanian people. The first phase of this contact coincides with contact with old Slavonic. Direct relations between the Romanian lands and Bulgaria continued intensively after the eleventh and twelfth centuries, the *terminus ante quem* for the penetration of words from old Slavonic. After that time there were continual Bulgarian migrations to the north of the Danube. To the south of the Danube the Aromanians and the Spanish-speaking Jews (from the sixteenth century) were in permanent contact with the Bulgarians even before the appearance of the Bulgarian state in the nineteenth century.

The chronology of Bulgarian elements in Daco-Romanian is based on phonetic and geographical criteria (Bulgarian words are widespread in the Muntenian subdialect of southern Romania). From the domain of agriculture: *clacă* ‘corvée, collective agricultural work’, *cobilă* ‘yoke’, *plută* ‘poplar’, *rapă* ‘rape’, *răsadniță* ‘hotbed’; social relations *rudă* ‘relation’; verbs *prăși* ‘to hoe’, *risipi* ‘to scatter, waste’, *scrobi* ‘to starch’. The Bulgarian influence on Aromanian and Judaeo-Spanish is limited to a few dialects. For contact with other south Slavonic varieties (Serbian, Croatian) – notably observable in the Banat subdialect of Romanian – see Neagoe (1984:270f.). For the influence of Croatian on Istro-Romanian, see below. Contact with Macedonian is largely limited to Megleno-Romanian (see Andreose and Renzi, this volume, chapter 8, p. 312), and after the First World War was continued intensely in Yugoslavia (from which the Republic of Macedonia subsequently emerged), but not in Greece. The Megleno-Romanians of Macedonia are all bilingual, Macedonian often being preferred in the family because children learn it at school. Some Macedonian words are general in Megleno-Romanian (*cular* ‘cartwright’, *iscra* ‘spark’, *lud* ‘mad’), other more recent ones (*bolniță* ‘hospital’, *danuc* ‘tax’, *marcă* ‘postage stamp’, *peglă* ‘flat iron’, *voz* ‘train’) are limited to Macedonian Megleno-Romanian.

The intensive, and historically enduring, influence of Slavonic is found only in some Romance languages geographically and culturally closest to Slavonic varieties, and first and foremost Romanian. The oldest contacts (seventh to eleventh centuries) between Romanian and old Slavonic (the latter as superstrate) took place both to the north and to the south of the Danube. When this began and how long it lasted is debatable. The old Slavonic loans in Romanian show phonetic characteristics of Bulgarian dialects from the period before the eleventh century. The influence is most evident in the lexicon, old Slavonic words being widespread in many onomasiological domains: the human being (*gleznă* ‘ankle’, (*a*) *iubi* ‘(to) love’, *obraz* ‘cheek’, *trup* ‘body’, *drag* ‘dear’, *groază* ‘terror’, *jale* ‘misery’, *vesel* ‘merry’), clothing (*cojoc* ‘sheepskin coat’, *izmană* ‘drawers’, *poală* ‘skirt’), home and household objects (*bici* ‘whip’, *clește* ‘tong’, *coș* ‘basket’, *pod* ‘loft, bridge’), family (*babă* ‘old lady’, *nevastă* ‘bride’), agriculture (*brazdă* ‘furrow’, *coasă* ‘scythe’, *grădină* ‘garden’, *greblă* ‘rake’, *plug* ‘plough’), nature (*bolovan* ‘boulder’, *izvor* ‘spring’, *luncă* ‘meadow’, *peșteră* ‘cave’), fauna (*bivol* ‘buffalo’, *cârțiță* ‘mole’, *lăstun* ‘housemartin’, *lebedă* ‘swan’, *rac* ‘crab’), flora (*hrean* ‘horseradish’, *mac* ‘poppy’, *ovăz* ‘oats’), social life (*bogat* ‘rich’, *muncă* ‘work’, *rob* ‘slave’, *sărac* ‘poor’, *scump* ‘dear (in both affective and economic senses)’, *drag* ‘dear (only in the affective sense)’). An example of a Slavonic calque on a word inherited from Latin (Meyer-Lübke

1898; Niculescu 1965) involves Latin SUPER > Ro. *spre* not only in the original sense of 'over, above', but also 'towards', influenced by the semantics of the Slavonic preposition *na*. There was also ancient contact with Church Slavonic, the literary variant of old Slavonic, which was the language of culture and administration in the Romanian lands between the tenth and the seventeenth centuries (it had the same status in Russia, Bulgaria and Serbia). Slavonic has a similar role in Romanian as Latin had in the western Romance world. Contact with Slavonic began in the ninth to eleventh centuries, but most Slavonic words entered Romanian in the fourteenth and fifteenth centuries via the church and the chancelleries. Slavonic was the conduit for the introduction of Greek words referring to the church (*apostol* 'apostle', *călugăr* 'monk', *icoană* 'icon', *mănăstire* 'monastery/convent', *popă* 'priest', *psalm*). It is often difficult to distinguish between Slavonic words and old Slavonic words which have entered Romanian in popular usage, for lack of rigorous criteria. Using the onomasiological criterion, terms considered to be Slavonic from the religious domain are *diacon* 'deacon', *vlădică* 'bishop', from the administrative domain *stolnic* 'high steward', *vornic* 'magistrate, headman', *zapis* 'deed, document', and from the cultural domain *bucoavnă* 'ABC book, book', *zbornic* 'collection of religious texts'. There are also phonetic criteria for identifying Slavonicisms: Slavonic words have final /h/, while popular forms have /f/ (*duh* 'spirit' – *vârf* 'peak, tip'); vocalization of old unstressed super-short vowels (or 'jers') in Slavonic (Ro. *săvârși* 'finish' vs. the popular loan *sfârși* 'finish'). Most Slavonicisms belong to old Romanian, but some persist to this day in popular speech (*duh* 'spirit', *vai* 'woe', *iad* 'hell'). In old Romanian there were prepositions (e.g., *bez* 'without', *na* 'to, on', *ot* 'from') borrowed from old Slavonic in translated texts or in the chancellery style, where Slavonic was used. Slavonic borrowings affected the old Romanian phonological system. Originally, [dʒ] (from Latin /j/ or /dj/ + back vowel, or from /g/ + front vowel), stood in complementary distribution with /ʒ/, the result of further evolution of /dʒ/ followed by a back vowel (thus /dʒem/ 'I groan' < GEMO; /ʒok/ 'I play' < IOCOR). Contact with old Slavonic introduced a series of words in which /ʒ/ appeared in other environments (e.g., *jalbă* /'ʒalbə/ 'supplication' < Slv. *žaliba*, *jale* /'ʒale/ 'sorrow' < Slv. *žalī*), thereby phonologizing the distinction between the two sounds (cf. ModRo. *geană* /'dʒanə/ 'eyelash' < Lat. GENA). In old Romanian the sequences /kl/ and /gl/ were absent as a result of sound change (e.g., OC(U)LUM > Ro. *ochi* 'eye', UNG(U)LAM > Ro. *unghie* 'fingernail'). Slavonic loanwords also introduce the clusters /kl/ and /gl/ into Romanian: e.g., *clopot* 'bell' < Slv. *klopotū*; *glas* 'voice' < Slv. *glasū* (see also Loporcaro, volume I, chapter 2, p. 83, for the effect of Slavonic loans on Romanian stress patterns).

Old Slavonic also contributes to derivational morphology through prefixes (*ne-*, *pre-*, *răs-*) and productive suffixes (*-an*, *-aș*, *-ean*, *-iște*, *-iță*, *-eală*). *Ne-* is a negative prefix which percolates into the indigenous grammatical system as an obligatory marker of negation of participial forms: cf. *Nu a văzut* ‘He has not seen’, with *nu* (< Lat. *NON*), but *nevăzut* (not ***nu văzut*) ‘not seen’, *nevăzând* (not ***nu văzând*) ‘not seeing’. *Pre-* and *răs-/răz-* may be loosely defined as intensifying prefixes which are freely combinable with indigenous roots: e.g., *a prelungi* ‘to lengthen, extend’, *a răsfrânge* ‘to refract’ (cf. *a frânge* ‘break, shatter’). Many of the suffixes are also combinable with Romance roots, for example *fruntaș* ‘leader, corporal’ (cf. *frunte* ‘forehead, fore’ < Lat. *FRONS*) or *îndoială* ‘doubt’ (*a îndoii* ‘to doubt’ < *doi* ‘two’ + *-eală*).

As mentioned, Slavonic is a major early source of a phonological innovation, namely the sound /h/, which was wholly absent from all proto-Romance varieties. Rather in the same way that Germanic affected French, /h/ appears in loanwords from Slavonic (e.g., *hrean* ‘horseradish’, *duh* ‘spirit’, *hrană* ‘food’; see also Loporcaro, volume 1, chapter 3, p. 142). Indeed, the vast majority of incidences of /h/ in Romanian vocabulary are located in loanwords, whether from Slavonic, Turkish (e.g., *han* ‘inn’, *habar* ‘clue, idea’), Hungarian (*hotar* ‘boundary’), German (*halbă* ‘beer mug’) or learned words of Greek or Latin origin bearing orthographic ‘h’ in other European languages, such as *hipocrit*, *hepatită*.

Slavonic impact on Romanian inflectional morphology is minimal, but an exception is the vocative ending, a form optionally used in addressing individuals by their name or function (cf. Petrucci 1999: 101–9). In the masculine singular there is a vocative ending *-e*, attached to the root (e.g., *stăpân* ‘master’, *stăpâne!* ‘o master!’) or, in monosyllabic roots and personal names, to the definite article (e.g., *om* ‘man’ – *omule!*), while singulars in unstressed *-a* or *-ă* (overwhelmingly, but not exclusively, feminines) attach *-o* to the root (e.g., *Maria* – *Mario!*). The masculine singular form may continue the Lat. second declension vocative singular *-e*, or may be from Slv. *-e* – or, indeed, the Slv. form may simply have favoured survival of the Latin form.¹⁵ Feminine *-o*, however, is indisputably a Slavonic feminine vocative ending. This vocative ending is, arguably, a superficial borrowing. As a direct address form, it is an aspect of Slavonic grammatical structure to which Romanians would have been exposed in face-to-face discourse, without their necessarily having a deeper grasp of Slavonic morphology.

¹⁵ Sandfeld (1930) considers that this vocative is of Slavonic origin, while others (e.g., Rosetti 1968) believe, without adducing any evidence, that the Latin vocative was reinforced by the Slavonic vocative in *-e*. The most widely held view is that *-e* is preserved from Latin and does not need explanation by borrowing.

It is an easily segmentable form with a transparent form–function relationship, similar to a derivational affix. Romanians have extended the ending *-o* into the morphology of one verb, the 2SG imperative *vin!* ‘come!’, which suggests that the ending was perceived as a marker of direct address, and that its introduction was not a matter of language shift on the part of Slavonic speakers using Romanian (e.g., Petrucci 1999: 107–9). The creation in Romanian of a *plural* vocative, out of purely Romanian morphological materials (the genitive–dative plural of the definite form: e.g., *domnilor!* ‘gentlemen!’ = ‘of/to the gentlemen’, *doamnelor* ‘ladies!’ = ‘of/to the ladies’) and without precedent in either Latin or Slavonic, is striking. The details of its emergence are obscure and problematic, and it certainly has no structural counterpart in Slavonic languages, but it appears to reflect an attempt to create a plural equivalent to a structure which, in the singular, is at least partly of Slavonic origin.

There is no reason to invoke Slavonic influence (cf. Graur 1963:35f.) to account for the preservation of Romanian nominal case inflection (cf. Salvi, volume I, chapter 7, p. 321) inherited from Latin (e.g., feminine nominative–accusative singular *casă* ‘house’, *vulpe* ‘fox’ vs. genitive–dative (*a unei*) *case* ‘(of a) house’, (*a unei*) *vulpi* ‘(of a) fox’). The so-called Romanian neuters (those nouns which have masculine gender in the singular and feminine gender in the plural; see Maiden, volume I, chapter 4, p. 174) have been attributed to Slavonic (Graur 1954) or to substrate (Sandfeld 1930; Nandriș 1961) influence. But they are also considered an internal Romanian creation (Rosetti 1957, who claims they arose from the need to distinguish animate from inanimate). Many accept the hypothesis that they simply continue the Latin neuter (Meyer-Lübke 1895:§12; Ivănescu 1957; Maiden, volume I, chapter 4). Fischer (1975) explains the phenomenon as a reorganization in late Latin whose results survived in Romanian and were slowly abandoned by the other Romance languages (traces of a similar state of affairs exist in old Italian and Sardinian, and parallels persist in southern Italian dialects).

Romanian compound cardinal numbers between eleven and nineteen (e.g., *unsprezece* ‘eleven’, *doisprezece* ‘twelve’, *nouăsprezece* ‘nineteen’), and the numbers for tens (e.g., *douăzeci* ‘twenty’, *treizeci* ‘thirty’, *nouăzeci* ‘ninety’), have a structure not found in other Romance languages (see Bauer, volume I, chapter 10, p. 551f.). The numerals for eleven, twelve and the teens have the structure ‘one/two, etc. + preposition *spre* + ten’, while the tens have the structure ‘two/three, etc. tens’. It must be noted that they are formed of Latin elements, although the apparently underlying construction *UNUS SUPER DECEM ‘one on ten’ is unattested in Latin, where *spre* (< Lat. SUPER) has a sense which it does not have in modern Romanian (where it means ‘towards’).

It is usually thought that the model for these constructions lay in Slavonic,¹⁶ where there are certainly parallel constructions. But these constructions could equally emerge independently of the situation in other languages. The numbers between eleven and nineteen could have as their basis the system of notches on tally sticks: ‘eleven’ is marked by making a notch above the first ten (it is known that shepherding was the basic occupation of the Romanized population in the region where the Romanian language formed, which supports the notion that the construction was extended to neighbouring populations).

One effect of contact between Slavonic and Romanian is most immediately apparent to the eye, in the fact that Romanian was written in the Cyrillic alphabet until well into the nineteenth century, when the Roman alphabet began to be adopted. A special case is represented by the Romanian of the Republic of Moldova (former Soviet Socialist Republic of Moldova), where contemporary Russian influence was manifest throughout much of the twentieth century in the imposition of a Cyrillic writing system identical (with certain modifications) to that of modern Russian.

Some of the most striking structural effects of Slavonic on Romanian occur in dialects spoken south and west of the Danube, and particularly in Istro-Romanian. The long coexistence of this tiny linguistic minority alongside Croatian (all Istro-Romanian speakers are bilingual, also speaking Croatian) has led to a highly asymmetrical relationship. Istro-Romanian influence on Croatian is minimal and superficial, being largely restricted to the lexicon of livestock rearing (cf. Kovačec 1968:80f.). In contrast, Croatian phonology, lexicon and grammatical structure have encroached profoundly on the structure of Istro-Romanian: the phonological system of Istro-Romanian has come to conform almost exactly to that of Croatian – a fact which has entailed wholesale neutralization, for example, of morphological number distinctions in masculine nouns and adjectives, originally expressed by phonemic oppositions alien to Croatian. Istro-Romanian has lost the opposition, under the influence of surrounding Croatian dialects, between palatalized and non-palatalized consonants, thereby neutralizing a distinction that played a crucial role in morphology (cf. Ro. SG. *lup* ‘wolf’ vs. PL. *lupi* [lupʲ]; ISG.PRS *rup* ‘tear’ vs. 2SG.PRS *rupi* [rupʲ]). In the village of Šušnjevića, the opposition between /s/ and /ʃ/ (with the same morphological functions – e.g., M.SG. *gras* vs. PL. *graʃ* ‘fat’) has also been lost, this time under the influence of Croatian dialects of the Lubin region, which in turn have lost this opposition under Venetian linguistic

¹⁶ Or in a Thraco-Dacian substrate (cf. Brăncuș 1973; Fischer 1985).

influence. In the noun this kind of neutralization is in part ‘remedied’ in the village of Žejane by the redeployment to animates of the originally inanimate Istro-Romanian plural endings *-e* and *-ure* (distributed, by the way, according to a pattern dependent on the number of preceding syllables and exactly reflecting an inflectional distinction made in Croatian; see Kovačec 1966:61, 63). Istro-Romanian stressed /a/ has also acquired the articulation /ɑ/, characteristic also of the Croatian dialects of that area.

Istro-Romanian has acquired the relatively free word order of Croatian (yet, unlike Croatian, lacks case marking on nouns and adjectives: cf. Kovačec 1966:63). Much of this influence involves relatively straightforward lexical replacement or structural calquing: where once there were (presumably) native Istro-Romanian forms and structures, now there are Croatian ones. But this is by no means always the case, and the focus here will be on what one might term ‘structured accommodation’, whereby the penetration into Istro-Romanian of a Croatian word or grammatical phenomenon is systematically attached to particular semantic or structural contexts, giving rise to distributional patterns which are native neither to Istro-Romanian nor to Croatian, but a product of the encroachment of the latter.

Istro-Romanian shows various examples of lexical encroachment from Croatian where the native term is preserved (usually in a rather more concrete, ‘domestic’, sense) and the incoming term has a complementary (and usually wider) semantic sphere. The result is often a semantic nuance not previously lexicalized in either language (cf. especially Kovačec 1963:37). Table 6.1 shows the standard Romanian form corresponding to the Istro-Romanian word, to exemplify the (presumed) more general meaning in the earlier history of Istro-Romanian.

Table 6.1 *The partial semantic encroachment of Croatian on Istro-Romanian*

Ro.	IRo.	CROATIAN WORD IN IRo.
<i>începe</i> ‘begin’	antʃepa ‘begin something consumed in the home (bread, sugar, wood)’	poʃni ‘begin generally, begin to do something’
<i>verde</i> ‘green’	verde ‘unripe, fresh, still growing’	zelen ‘green’
<i>rău</i> ‘bad’	rev ‘bad (of dog or land)’	təmən ‘bad’
<i>scoate</i> ‘remove’	skote ‘rescue from fire or water’	poteʏni/spasi/zvuʦʃi ‘remove’
<i>greu</i> ‘heavy, difficult’	grev ‘heavy’ (dialect of Šušnjevića)	teʒək ‘difficult’ (in Croatian this word also means ‘heavy’)

Table 6.2 *Romance and Croatian numerals in Istro-Romanian (Croatian forms boxed). The numbers '1' and '2' alternate for gender*

I	2	3	4	5	6	7	8	9	10	11
ur	doi	trej	patru	ʦɨntʃ	ʃase	ʃapte	opt	devet	deset	jedanašt
o	do			pet	ʃest	sedəm	osəm			

The novel lexicalization of sense distinctions as a concomitant of lexical borrowing is a well-known feature of language contact: native English *swine*, *calf*, *cow* as against *pork*, *veal*, *beef*, of Norman French origin, is perhaps the classic example. What needs to be stressed is that, even in circumstances where the lexicon of one language is being 'eroded' by that of another, the result may be something 'new' – a lexical distinction not previously made. But Istro-Romanian shows a more elaborate type of 'structured accommodation', in respect of numerals.

At first sight, the encroachment of Croatian numerals into Istro-Romanian seems to follow a clearly defined path, such that higher numbers are more susceptible to replacement than lower ones. Specifically, numbers above ten are Croatian, as are ten and/or nine in some dialects; from eight downwards they are usually Romance. However, from five to eight Croatian numerals coexist with Romance ones (Table 6.2).

The fact that the Croatian numerals do not percolate below 'five' in Istro-Romanian is reminiscent of a major division in the Croatian numeral system (although one of a quite different kind), such that plurals up to 'four' select the genitive singular, but those from 'five' upwards select the genitive plural. But the *formal* manifestation of this division in Istro-Romanian is quite different from anything in Croatian (or in Romanian dialects).

The copresence of Romance and Croatian forms from 'five' to 'eight' is by no means a matter of 'free variation'. Rather, the Croatian forms *must* be used in 'lexical measure phrases' (phrases expressing characteristic units of measurement, such as time, weight and distance); moreover, they *must* be combined with a Croatian noun, where one is available, showing Croatian noun morphology. Use of Romance numerals and nouns in such cases (e.g., Romance ʦɨntʃ zile for 'five days') is reportedly rejected by many Istro-Romanians. Moreover, the noun usually preserves an archaic Croatian genitive plural form, characterized by a lack of inflectional ending (Kovačec 1966:65f.;

Table 6.3 *Romance and Croatian numerals in lexical measure expressions*

Romance numerals 1–4			
ur ɔn '1 year'	doi ɔn '2 years'	trej kile '3 kilos'	patru metər '4 metres'
o ura '1 hour'	do zile '2 days'		
Romance numerals 5–8			
tʃintʃ omir '5 people'	ʃase vatʃ '6 cows'	ʃapte kase '7 houses'	opt frats '8 brothers'
Croatian numerals and nouns from 5–8 in measure expressions			
pet dən '5 days'	ʃest ur '6 hours'	sedəm let '7 years'	osəm kil '8 kilos'

1968:99f.; Petrovici and Neiescu 1964:191), and this is another respect in which the construction is neither 'Romanian' nor 'Croatian', in that the modern Croatian Čakavian dialects of the Istrian peninsula have replaced such genitive plurals with forms bearing inflectional endings (Table 6.3).

Apparently, no other Romance language, nor any Slavonic language, has such a specialized marking of lexical measure phrases. Its existence in Istro-Romanian is a product of the entry of Croatian vocabulary into Istro-Romanian. Kovačec (1963:25) surmises:

To begin with, these constructions with the genitive plural were felt as belonging to an alien system, but today, in Žejane, they are the only ones possible. Croatian numerals above 5 [sic] were not translated by the corresponding Romanian terms [...] to avoid syntagmatically hybrid formations. Bilingualism, at the high level at which we find it among the Istro-Romanians, tends to find for each form of a language (the more prestigious one) a corresponding different form in the other language, even if this means increasing the number of forms in that language.

Istro-Romanian plural nouns do not inflect for case (case is marked on their determiners) so the Croatian genitive plural form would have had no structural equivalent in Istro-Romanian: in other words it was a distinctively Croatian form and perhaps particularly for this reason would have selected a Croatian numeral. But the historical consequence is, in effect, the conversion of synonymous forms belonging to different languages into suppletively alternant numerals distributed according to semantic (and morphological) characteristics of the accompanying noun.

The morphological marking of *aspect* constitutes a major typological difference between Romance and Slavonic languages. While Romanian, like other

Romance languages, limits distinctions of aspect in its inflectional morphology to the past tense ('imperfect' vs. 'perfect' tense forms: e.g., *veneam* 'I was coming' vs. *am venit* 'I came'), Croatian has a verb system organized around distinctions of aspect, and particularly that between perfective and imperfective aspect. Indeed, Croatian possesses multiple morphological devices (prefixation, infixation, conjugational class, stress, stress and tone combined) to differentiate aspect (cf. Kovačec 1963:25–28, 37; 1966:70f.; 1968:108f.; also Hurren 1969). Istro-Romanian is typologically distinguished among Romance languages in that it has acquired, through contact with Croatian, a fully-fledged system of morphological distinction between perfective/imperfective in the verb (and in non-finite as well as finite forms). In many cases both the perfective and the imperfective form are borrowed: e.g., *skotʃi* / *skakɛj* < Crt. *skočit/skakat* 'jump'. But in other cases aspectual pairs are formed from indigenous Istro-Romanian verbs: in the typical case, the normal Istro-Romanian verb-form is treated as the imperfective, while a perfective form is constructed by various devices, such as addition of Croatian affixes (e.g., imperfective *tortʃe* / perfective *potortʃe* 'spin', imperfective *latra* / perfective *zalatra* 'bark').¹⁷ However, in some cases the aspectual pair is created by treating the indigenous Istro-Romanian verb as imperfective, and deploying a Croatian perfective form to provide the perfective. Table 6.4 shows infinitive forms.

The emergence of a full morphological aspect system in a Romance language under the influence of a more dominant Slavonic language is striking (and well known in the literature on language contact), but its real theoretical interest lies in the type of 'bilingual' aspect marking illustrated in Table 6.4. At one level, all that has happened is that a typically Romance system of limited aspect marking has been effaced by the more extensive Slavonic one, but what is remarkable is that speakers have, in a sense, 'grammaticalized' the difference between the dominant and the recessive language, by effectively expressing perfectives in one language and imperfectives in the other. The result is a considerable increase in the incidence of suppletion in the verb – perhaps the kind of morphological phenomenon that one would not ordinarily expect to encounter under circumstances of language death. In general, it serves as further evidence in favour of the observation made by Maiden (2006), that contact between languages can play a major role in the genesis of suppletion.

¹⁷ Croatian infixes are also used to form iteratives: imperfective *durmi* / iterative *durmivej* 'sleep'.

Table 6.4 'Bilingual' signalling of aspect in Istro-Romanian verbs

	ROMANCE IMPERFECTIVE	CROATIAN-BASED PERFECTIVE
'drink'	bɛ	popi
'eat'	məŋka	pojdi
'spin'	tortʃe	spredi
'grind'	matʃira	zmeʎi
'plough'	ara	zori
'measure'	mesura	zmiri
'wet'	muʎa	zmoʃi
'dry'	uska	osuʃi
'hurt'	durɛ	zaboli
'dig'	sapa	skopej
'sleep'	durmi	zaspi
'bake'	kotʃe	spetʃi
'shoe [a horse]'	fareka	prikuji

Neither Romanian nor Croatian has an inflectionally distinct category of adverbs. Romanian adverbs are (generally) identical in form to the masculine singular adjective. Masculine and feminine adjectives are given in example (1), and the corresponding adverbs in (2):

- (1) a. *Cântecul* *e* *frumos.*
 song-the.M.SG is beautiful.M.SG
 'The song is beautiful.'
- b. *Povestea* *e* *frumoasă.*
 story-the.F.SG is beautiful.F.SG
 'The story is beautiful.'
- (2) a. *Băiatul* *cântă* *frumos.*
 boy-the.M.SG sings beautiful
 'The boy sings beautifully.'
- b. *Fata* *povestește* *frumos.*
 girl-the.F.SG narrates beautiful
 'The girl narrates beautifully.'

Like Romanian, Croatian lacks morphologically distinct marking of adverbs. Adverbs are identical in form to the neuter singular nominative–accusative of the adjective, usually characterized by inflectional -o: e.g., adjective M *dobar* 'good', F *dobra*, NEUT *dobro* adverb *dobro* 'well'; adjective M

Table 6.5 *Inflectional marking of adverbs in Istro-Romanian*

M	F	ADVERB	
siȳuran	siȳurna	siȳurno	‘certain’
vesel	vesela	veselo	‘happy’
ȳrev	ȳre	ȳrevo	‘heavy’
əntreȳ	əntreȳa	əntreȳo	‘whole’
plin	plina	plino	‘full’
tərd	tərda	tərdo	‘hard’
at	ata	ato	‘other’

običan ‘usual’, F *obična*, NEUT *obično* adverb *obično* ‘usually’. However, Istro-Romanian acquired the Croatian neuter inflection *-o*, but primarily and originally in the *adverbial* use (there is also some use of the *-o* form as an attributive adjective with the ‘neuter’ pronouns *tȳa* ‘that’ and *tȳasta* ‘this’: *tȳa* a fost buro ‘that was good’; there is also innovatory use of agreement in *-o* with borrowed Croatian neuter nouns in *-o* (Kovačec 1963:33f.; 1966:67f.; 1968:87, 90; 1998:267). Affixation of *-o* is also productively used to form adverbs from native Istro-Romanian, as well as from Croatian, adjectives (Table 6.5).

Thus, for example, *ku kosiru se tȳaȳa mai drobno e ku sekura mai ȳroso* ‘With the billhook one cuts more finely and with the axe more roughly’; *a durmit tərdo* ‘He slept deeply’; *s a tot fino vezut* ‘One has seen everything well’.

The result is that Istro-Romanian, unlike other Romanian dialects, other ‘eastern Romance’ languages, and Croatian, has a morphologically distinct class of adverbs. Occasionally, it is not only the *-o* ending but an entire Croatian neuter singular lexeme which is borrowed into Istro-Romanian, resulting in suppletive alternation between adjective and adverb:¹⁸ adjectives *tare* ‘strong’, *ȳrev* ‘heavy, difficult’ vs. adverbs *jako* ‘strongly’, *teȳko* ‘heavily, with difficulty’ (Kovačec 1968:85). We are indeed witnessing the structural ‘erosion’ of Istro-Romanian by Croatian, but we are also seeing the genesis of something typologically unique among Romanian dialects, an inflectional marking of adverbs. In the literature on ‘grammaticalization’ the emergence of morphological marking of adverbs in the history of the Romance languages is a classic chapter (see Bauer, volume I, chapter 10, pp. 552–56; also Bauer 2001;

¹⁸ This expands a phenomenon for which there was originally but one example: *bur* ‘good’ vs. *bire* ‘well’ (cf. Romanian *bun* vs. *bine*).

Table 6.6 *Aspectual pairs in Megleno-Romanian*

IMPERFECTIVE	PERFECTIVE
mən'kari 'eat'	nəmən'kari 'eat up'
dur'miri 'sleep'	zədur'miri 'fall asleep'
'twartsiri 'spin'	du'twartsiri 'finish spinning'
vikə'jiri 'shout'	vik'niri 'give a shout'
Imperfective	Iterative ¹⁹
tur'nari 'turn over'	printur'nari 'turn over and over'
mu'triri 'look'	zəmu'triri 'keep looking'

2003): the idiosyncratic and lexical variable adverb marking of Latin was largely lost, leaving much of southern Italo-Romance and Balkan Romance without distinct morphological marking of adverbs, while in 'western' Romance languages reflexes of Latin *MENTE* '(with a) mind' were reanalysed as adverbial affixes (e.g., *FELICE MENTE* 'with a happy mind' > It. *felicemente* 'happily'). The IRO. data not only call for a minor modification in the usual typological division of Romance languages, but they also show that inflectional marking of a grammatical category can be a product of the structural encroachment of one language on another, even when neither language previously made such a distinction.

Another branch of Daco-Romance in which Slavonic influence has penetrated the grammatical system is Megleno-Romanian. Like Istro-Romanian (although apparently to a lesser degree; see Atanasov 2002:226), Megleno-Romanian creates aspectual distinctions by means of prefixation or infixation, using Slavonic (Macedonian) affixes. For example, see Table 6.6.

Finally, Megleno-Romanian is cited by Weinreich (1953:32) in support of his claim that close typological fit favours deeper structural effects under contact. Specifically, Megleno-Romanian dialects display inflectional *-u* and *-i* respectively in the 1SG and the 2SG present (*aflu aflu*). These endings are present in Megleno-Romanian to varying degrees according to dialect and phonological environment (see Atanasov 2002:237), but they are subject in some subvarieties (particularly of Lunđiń and Oşîń) to the influence of the similar Macedonian 1SG and 2SG endings *-um* and *-iş* (*aflum afluş*). In syntax, there is clear calquing in

¹⁹ Interestingly, Atanasov (2002:227) observes that oppositions between iterative and non-iterative in the perfective can only be made between verbs of Macedonian origin.

some dialects of a Bulgarian/Macedonian construction in the use of inversion of participle and auxiliary to form an *evidential* form of the perfect: e.g., əḷ kən'tatə 'he has sung' vs. kən'tat əḷ '[apparently/allegedly] he has sung' (see Atanasov 2002:243f.).

2.2.4 Arabic

Arabic is a Semitic language, wholly unrelated to Romance. Its influence on Romance languages is second only to that of Germanic, and is mainly lexical, Ibero-Romance varieties (followed by Sicilian and Occitan) being those principally affected. Romanian and Raeto-Romance have no *direct* loans from Arabic.

The different exposure of Romance varieties to Arabic has a historical and social explanation. In 711, the Moorish invasion brought Islam and the Arabic language to the Iberian Peninsula. The whole peninsula (including the Balearic islands) was soon under Arab dominion, with the exception of Galicia, Asturias and Cantabria, in the north. It remained so for many centuries, until the 'Reconquest' of territories from Moorish dominion by Christian Romance-speakers from the north, which began in the eleventh century and ended in the expulsion of the Moors from Granada in 1492. The Arabs also entered France, but were definitively pushed south of the Pyrenees in 759. They ruled Sicily until it was conquered by the Normans, in 1072.

The Arabs were bearers of a superior civilization, but the Romanized population which found itself under them in the Iberian Peninsula (Mozarabs) generally kept their possessions, religions and language, while adopting the lifestyle of the Arabs.²⁰ The Mozarabs borrowed not only terms for institutions and objects associated with the Arabs but also terms for more basic objects. In Sicily, over two centuries of Arab domination led to some modification of toponyms and the borrowing of numerous words. Contacts between southern France and Arabs took place during the Crusades and, later, with the states of northern Africa.

Most Arabic words are nouns denoting concrete realities (we do not find terms for feelings, emotions,²¹ vices and virtues – probably because these were domains occupied by Christian, hence Latin/Romance, terminology – and there are very few verbs and adjectives). Well represented are:

²⁰ There was extensive Arabic–Romance (Mozarabic) bilingualism, both on the part of the indigenous community and on the part of Arabs. For Mozarabic, see Galmés de Fuentes (1983).

²¹ But note the borrowing of the interjection *wa ša llāh* as Sp. *ojalá*, Pt. *oxalá*, 'would to God' (lit. 'to Allah'). It seems unlikely that this term would have survived in Christian Spain and Portugal had its component etymological meanings been properly understood.

- The army and military life: Pt., Sp., Cat. *adarga* 'shield'; Pt. *alforge*, Sp. *alforjas* 'saddlebags'; Pt. *refem*, Sp. *rehén* 'hostage'; Pt., Sp. *alfange*, OIt. (c) *angiarro* 'sabre'; Sp. *alcaide*, Cat. *alacai*, OFr. *aucaise*, Occ. *aucais* 'commander'.
- Administration: Sp. *alcalde* 'mayor'; Pt. *alguazil*, Sp. *alguacil*, It. *alguazzino* 'bailiff'.
- Agriculture and foodstuffs: Pt. *azeite*, Sp. *aceite* '(olive) oil'; Pt. *almóndega*, Sp. *albóndiga* 'meatball'; Pt. *açafrão*, Sp. *azafrán*, Fr. *saffran* 'saffron'; Pt. *acéquia*, Sp. *acequia*, Cat. *sequia*, Sic. *saya* 'irrigation channel'; Sp. *azúcar*, It. *zucchero* 'sugar'.
- Industry, technology: Sp. *alquimia* 'alchemy'; Sp. *azogue* 'quicksilver'; Sp. *alambique* Occ., Cat. *alambic*, It. *lambicco* 'still', etc.
- Science and learning: Sp., Pt., It. *algebra*; Sp. *cero* 'zero'; It., Sp., Pt. *nadir*; Sp., Pt. *alcohol*, It. *alcol*.
- Commerce, trades and professions: Sp. *aduanas*, Cat. *duana*, It. *dogana* 'customs'; Sp. *ahorrar* 'save money'; Sp. *albañil* 'builder'; Sp. *alfarero* 'potter'; Sp. *tarifa*, It. *tariffa* 'tariff'.
- (Exotic) flora: Pt. *algodão*, Sp. *algodón*, It. *cotone* 'cotton'; Pt., Sp. *arroz*, Cat. *arros* 'rice'.
- (Exotic) fauna: Pt. *alacrão*, Sp. *alacrán* 'scorpion'; Pt. *javalí*, Sp. *jabalí* 'boar'.
- House building and decoration: Sp. *azulejo* 'tile'; Sp. *alfombra* 'carpet'; Sp. *taza* 'cup', etc.

Among the few adjectives are Sp., Pt. *azul*, It. *azzurro* 'blue'; Sp. *mezquino*, Pt. *mesquinho*, Occ., Cat. *mesquí* 'mean, paltry'. Arabic also provides the generic form Sp. *fulano* 'somebody or other, so-and-so'. There is one, isolated, case of borrowing of a preposition: Sp. *hasta*, Pt. *até* 'until' from Arb. *ḥatta*.

Loans into Spanish and Portuguese, in particular, often bespeak an imperfect knowledge of Arabic, since they are taken over complete with the Arabic definite article *a(l)*- analysed as an integral part of the word. This is less true of Catalan and Occitan, and particularly of Italo-Romance (cf. Sp., Pt. *aduanas*, Cat. *duana*, It. *dogana* 'customs'). A rare borrowing of an element of morphological structure from Arabic is the ethnonymic suffix *-í* which has continued to be productive in Spanish (e.g., *marroquí* 'Moroccan', *alfonsí* 'alphonine', *israelí* 'Israeli'). There is also general phonological adaptation of Arabic words to Romance, again suggesting an imperfect knowledge of Arabic on the part of Romance speakers. Thus Arabic laryngeals were taken into old Castilian as /h/ and sometimes /g/, /k/ or zero: *ḥinna* > *alfeña* > *alheña* 'henna', *ḥarruba* > *algarroba* 'carob bean', *ḥayla* > *ola* 'wave'. Arabic dental fricatives had no exact counterpart in Romance, and were replaced by the affricates /ts/ and /dz/ (later /θ/): *safunariya* > Sp. *zanahoria*, Pt. *cenoura*

‘carrot’. Arabic /w/ is adapted variously as /gw/, /β/ or /w/: *ʔustuwân* > Sp. *zaguán* ‘hallway’, *muga:wir* > Sp. *almogávar* ‘frontier guard’, *diwân* > Sp. *aduana* ‘customs’. Arabic loans swelled the originally very small number of Ibero-Romance words with stressed final vowels: e.g., Pt. *javalí*, Sp. *jabalí* ‘boar’.

There is some rare evidence of Romance speakers having deeper awareness of the meanings of Arabic words in cases of semantic borrowing from Arabic, where a Romance word which shares a basic meaning with an Arabic word acquires other senses originally unique to the Arabic. For example, Arb. *ga:wara* ‘run’ also had the meaning ‘to raid, depredate’, a sense which it acquires in old Spanish *correr* and also *corredor* ‘depredator’; Spanish *adelantado* or Pt. *adiantado* equivalent to the Arabic participle *almuqaddam* ‘placed before’, acquires the additional Arabic sense of ‘chief’, ‘magistrate’, ‘authority’; similarly *alcalde*, where the holder of the office in Arab society was both ‘mayor’ and ‘judge’, a usage continued in OSp. Other semantic calques are Sp. *infante* with the meaning ‘son of a king or nobleman’, *casa* ‘house’, with the additional meaning of ‘town’. As Lapesa (1980:151–54) indicates, there is simply nothing in the syntax of Spanish that can be unambiguously assigned to Arabic influence and which does not have parallels in other Romance varieties.

2.2.5 Other languages

It is not possible here to give a comprehensive survey of the effects on Romance of the very many, European and non-European,²² languages with which it has come into contact.²³ We suggest that the most interesting cases are those where the effects of borrowing go beyond mere lexical transfer to affect linguistic structure, or where borrowing affects the most basic areas of the vocabulary.

Turkish has had much stronger influence on Romanian (and Judaeo-Spanish) than on other Romance varieties, because the Romanian lands (Wallachia and Moldavia)²⁴ were under Ottoman suzerainty from the fifteenth to the nineteenth centuries. Romanian shows two layers of loanwords from Turkish, one from the fifteenth to the seventeenth centuries, from which many terms entered popular usage, and a later and short-lived one in the Phanariot period (eighteenth century). Turkish words appear in various lexical domains:

22 For the lexical influence of non-European languages on Romance, see also Jones and Pountain, this volume, chapter 10.

23 For the special case of the effects of Latin on Romance, see particularly Pountain, volume I, chapter 13; also Dworkin, volume I, chapter 12, p. 601.

24 Turkish terms are correspondingly absent from modern Transylvanian dialects.

e.g., *balama* 'hinge', *dulap* 'cupboard', *tavan* 'ceiling', *cafea* 'coffee', *halva*, *telemea* (a type of cheese), *dovleac* 'marrow', *dud* 'mulberry', *para* (a type of coin), *teighea* 'counter, bench', *fudul* 'haughty', *șiret* 'sly', *tembel* 'slothful', *berechet* 'plentifully, galore', *belea* 'trouble, mishap', *chef* 'high spirits, wish, party', *huzur* 'ease, comfort'. It is interesting to note that Turkish loanwords in Romanian that are ultimately of Arabic origin sometimes have cognate counterparts in Arabic loanwords in Spanish: e.g., Ro. *giuvaer* 'jewel', Sp. *aljófár* 'pearl'. Turkish is also the source of some suffixes which have become productive in Romanian: *-giu* (*cafegiu* 'lover of coffee, keeper of coffee house'), *-iu* (*ruginiu* 'rust coloured'), *-lic/-lâc* (*șiretlic* 'ruse', *crailâc* 'philandering'). Romanian has an inflectional type not found in other Romance languages and which emerged following a purely phonetic development, the treatment of the Latin geminate LL before *-a* and *-e*. This affected feminine nouns whose plural ended in *-le* (STELLA STELLAS > Ro. *stea stele* 'star/s'). This was an unproductive type which became productive as a result of Turkish influence, when many words in stressed final vowels (especially stressed *-e* yielding Romanian *-ea*) were added to the language and analogically received the plural ending *-le* (*cafea* – *cafele* 'coffee/s'). The process extended to words of Greek origin (*canapea* – *canapele* 'canapé/s'), and French words (*șosea* – *șosele* 'chaussée/s, highway/s', *bezea* – *bezele* 'baiser/s, meringue/s, blown kiss/es') which had a final stressed vowel.²⁵

Hungarian influence is also largely limited to Romanian (or to those varieties north of the Danube, the Hungarians having penetrated the Pannonian Plain in 896, and Transylvania between the eleventh and the thirteenth centuries, where they encountered a compact population made up of small Romanian states). Some basic Romanian words that are of Hungarian origin are *fel* 'kind, way', *oraș* 'town, city', *marfă* 'goods', *mereu* 'always', *gând* 'thought', *seamă* 'account'; likewise the Romanian suffixes *-eș*, *-ău*, *-șag*, *-șug*.

The influence of medieval (Byzantine) and modern Greek on the Romance languages is largely lexical: 192 medieval Greek words are recorded in western Romance languages, while in Romanian there are twenty-two direct loans and another 254 transmitted via Slavonic (Kahane and Kahane 1970–76; Mihăescu 1966). Some structural effects are observable in Aromanian, many of whose speakers are bilingual in Greek. Aromanian shows the development of the syllable-final glide [ɥ] as a fricative or occlusive under modern Greek influence (e.g., *alavdu* 'I praise' = Ro. *laud*, *caftu* 'I seek' = Ro. *caut*). In Aromanian there is also borrowing from Greek of the conjunctions *ama* 'but' and *ică* 'or'.

25 For a discussion of the possible consequences of this development for the native Romanian diminutive feminine suffix *-ea* (< Lat. *-ELLA*), see Maiden (1999).

It is a matter of controversy whether the Greek spoken extensively in the far south of Italy in the Middle Ages is a direct descendant of the ancient (mainly Doric) Greek varieties imported into *Magna Graecia* as early as the eighth century BC (a view promoted by Rohlfs, for example 1924; 1977) or whether, as argued particularly by Battisti (1927),²⁶ it is a more recent import dating from the Byzantine period of domination between the sixth and eleventh centuries. In any case it is clear that at the beginning of the second millennium AD Greek was still widely spoken as a native language in north-western Sicily, Calabria and Puglia. Today it survives in only a handful of remote villages of the Aspromonte of southern Calabria and, to a greater degree, in a cluster of villages in the Salento peninsula south of Lecce.

The influence of Greek on the Romance varieties is incontrovertible. It extends far beyond the merely lexical (e.g., such widespread southern Italian Hellenisms as χίμαρος > *zímmaro* /-u 'billy-goat', νάχη > *naca* 'cradle', απαλός > *ápulo* /-u/-e 'soft, without a shell', γαστρα > (g)*rasta* 'vase, pot', and λαγανα > *lågana* 'lasagna-style pasta'), or isolated grammatical features. In the area of morphology, there is: widespread absence of the adverbial suffix *-mente*, whose functions are typically covered instead by the simple bare adjective (see Ledgeway 2002; 2011); generalized use of the *passato remoto* in large parts of the far south as the sole past perfective tense (e.g., the Salentino of Aradeo *mo-moi catíu* lit. 'he just fell over' = 'he has just fallen over');²⁷ the use of the so-called *dativo greco* ('Greek-style dative') in large parts of southern Calabria such that, on a par with the merger of dative and genitive during the Middle Greek period (cf. Joseph 1990:160) the genitive preposition *di* 'of' has extended its functions, under specific conditions (Trumper 2003:232f.), to include the marking of dative arguments (Bagaldi, Calabria *nci lu scrissi di mè frati* 'I wrote it to my brother').

Greek syntactic influence is most evident in the dialects of the far south (central-southern Calabria, north-eastern Sicily, and Salento).²⁸ Here, for example, and as in Greek, finite clauses are generally employed at the expense of the infinitive (example (3)),²⁹ the imperfect indicative is employed in the protasis and apodosis of unreal hypothetical clauses (example (4)),³⁰ and first names, both male and female, are regularly preceded by the definite article (example (5))³¹ Thus:³²

26 For an overview of the linguistic evidence, see Horrocks (1997:304–6).

27 See Rohlfs (1968:312; 1972a).

28 See Ledgeway (2006) for an analysis of alleged Greek influence on the southern Italo-Romance complementizer system.

29 See, e.g., Meyer-Lübke (1899:607–8, 615–22), Rohlfs (1972b; 1972c) and Calabrese (1993).

30 See Rohlfs (1972a). 31 See Rohlfs (1977:181). 32 Examples from Ledgeway (2006).

- (3) *Àiu 'u vâiu' 'u vîju duv' àiu 'u vâiu ôja* (Sant'Andrea, Catanzaro)
I-have that I-go that I-see where I-have that I-go today
'I have to go and see where I have to go today.'
- (4) *iva si non chivîa* (Mèlito di Porto Salvo, Reggio Calabria)
I-went if not it-was-raining
'I would go if it were not raining.'
- (5) *fazzu 'i trasi 'u Mariu* (Bianco, Reggio Calabria)
I-make that enters the Mario
'I make Mario enter.'

As may be seen from the foregoing, the plausibility of attributing any one of these developments to Greek influence is reinforced by the fact that it belongs to a *cluster* of characteristics, lexical, morphological and syntactic, all finding counterparts in Greek.

3. What is borrowed and when?

The foregoing typology of contact, and the illustrations of its effects, point to some clear conclusions for Romance languages about what can be borrowed, and under what circumstances. The surest, and most uncontroversially identifiable,³³ cases involve lexical borrowing. The Romance languages differ with regard to the number of words borrowed. Romanian developed in contact with many different non-Romance languages from which it borrowed many words, so much so that it presents particular difficulties for Romance linguists reading a Romanian text, especially one written before the nineteenth century, when Romanian eliminated many older borrowings and began borrowing extensively from other Romance languages. There are no exhaustive studies on lexical borrowing from other Romance languages or from a particular Romance language, but there are some noteworthy works which deal with a number of loans, without being exhaustive (e.g., Pellegrini 1972; Hope 1971).³⁴

³³ But there are in addition thousands of borrowed words of uncertain etymology. The *Französisches Etymologisches Wörterbuch* (Wartburg 1922–2003), the best etymological dictionary of a Romance language (French), indeed dedicates three volumes (21–23) to words of uncertain or unknown origin. Many words previously assumed to have been borrowed have been explained as actually having been inherited from Latin.

³⁴ For 'learned' lexical borrowing between Romance languages and from Classical Latin, see Pountain, volume I, chapter 13, pp. 628–43.

Lexical borrowings are usually unambiguously identifiable because they involve an arbitrary matching of sound and meaning originally specific to a particular language. A type of lexical borrowing which is less safely attributable to contact involves semantic ‘calquing’, effectively translation of the meaning of the component elements of words or expressions from one language into another, without borrowing of form. No complete theory of the phenomenon yet exists; indeed, there are few problems in linguistics so fraught with contradictions. These are even greater if we also consider that, sometimes, what was thought to be the result of a semantic calque can just as easily be, as in syntax, the result of parallel semantic development (cf. the alleged Celtic influence on the Gallo-Romance counting system in §2.2.1 or the alleged Slavonic influence on the Romanian counting system in §2.2.3.), unless the meanings involved are particularly idiosyncratic (cf. the old French *on* ‘one’ vs. *soi tiers* discussed in §2.2.2). Coseriu (1977a; 1977b) points for example to some semantic changes in Spanish wrongly attributed to Arabic influence, and to developments in Romanian considered to be calqued from Balkan languages. These may in fact be semantic developments which also exist in other parts of the Romance world. For example, Sp. *ojo* (*de agua*), Pt. *olho-de-água*, literally ‘eye of water’, meaning ‘spring, source of water’, is reminiscent of the Arabic use of the word ‘*ayn*’ to mean both ‘eye’ and ‘water’, but the existence of parallels in Turkish, in regional varieties of Romanian, and elsewhere, points to ‘polygenesis’ rather than an Arabic origin. Similarly, Romanian *a pierde* ‘lose’, used in the sense ‘destroy, ruin’, has been claimed to be calqued on Slavonic *pogubiti*, which has both meanings, while in fact such use of the verb ‘lose’ is attested in Latin and other Romance languages. A comparative knowledge of the situation across the Romance varieties and in Latin can therefore prevent what are actually normal developments being wrongly attributed to external influences.

The number of borrowed words or semantic calques is greater at the level of dialects that have found themselves in contact with different languages. The Romance linguistic atlases and dialect monographs contain numerous examples of this kind. Basic notions (human beings and their anatomy, physiology, affective and intellectual life, family life, the elements of material civilization, natural phenomena, the earth and waters, nearly all cultivated plants and the majority of non-cultivated ones, fauna,³⁵ actions, gestures) are expressed by terms inherited from Latin (see also Stefenelli, volume I, chapter 11). The set of grammatical function words is of Latin origin; in the case of the various

35 But not names of fish in the case of Romanian.

terminologies illustrating economic and social life there are major differences: agriculture and livestock-raising have terms inherited from Latin down to the finest details, as opposed to certain more complex occupations (e.g., metal-working and wood-working), culture, administration, and religion, which are more open to borrowing.

There is relatively little borrowing of function words, such as prepositions (but see §2.2.2 for some contributions from Germanic in the form of intensifying adverbs; borrowing of prepositions into old Romanian from Church Slavonic in §2.2.3), and borrowing of conjunctions is even rarer, although the syntax of these might be borrowed despite the fact that the lexemes are supplied by Romance morphemes (cf. *mi/ma/mi* and *cu* replacing the infinitive in the dialects of the far south of Italy). In educated Romanian there is *or* < Fr. *or* 'now' (as discourse marker), and in Sardinian *sigumente* 'since, as', from Catalan, is the only consecutive conjunction. In Aromanian, in addition to loans of conjunctions from Greek (see §2.2.5), there is *ma* 'but' from Italian. Megleno-Romanian has some conjunctions from Macedonian (*ācu* 'if', *āmi* 'and', *āmā* 'but'), two from Turkish (*em*. . . *em* 'and', *ia*. . . *ia* 'maybe') and one from Greek (*i*. . . *i* 'either. . . or').

Morphology is generally resistant to borrowing (see below), but an area in which lexical borrowing and morphology rather overlap is word formation, and here borrowing is frequent and easily identifiable. Derivational affixes (especially prefixes) and compounding elements are often like full lexical words, in that they usually involve language-specific arbitrary matchings of sound and meaning, being easily segmentable within the words in which they appear, and easily associable with particular functions or lexical meanings. Malkiel (1978) considers that about 10 percent of Romance affixes are loans,³⁶ pointing out that word formation is 'a kind of bridge between grammar proper and lexicology'. There are many borrowings of suffixes between Romance languages, French being the most obvious source, e.g., the French nominalizing suffix *-age* (< Lat. *-ATICUM*) > Ro. *-aj* (*abataj* 'mine working'), It. *-aggio* (*erbaggio* 'herbage'), Sp. *-aje* (*bestiaje* 'stock of cattle'; see also Bauer, volume I, chapter 10, p. 535). The Latin system of prefixes has been reorganized: in place of prefixes which have disappeared the Romance languages have borrowed numerous prefixes directly from Latin or from other Romance languages. Loans widespread across the Romance languages are prefixes of Greek origin, which are very productive in the literary languages, especially in

36 For affixal borrowing from classical languages, see Pountain, volume I, chapter 13, pp. 640–42.

scientific and technical terminology (e.g., *anti-*, *para-*, *hyper-*/(*h*)*iper-*). Besides Romance, Latin and Greek prefixes, Romanian has of course also borrowed from Slavonic, which is why Romanian is the Romance language richest in prefixes: eighty-six elements, of which fifty-six are loans from Latin (twelve are inherited from Latin). Of the fifteen prefixes of Slavonic origin, only three are of any importance (see §2.2.3). The rise of prefixation, beginning in the nineteenth century, is an essential component of what is known as re-Romanization or modernization of the Romanian language.

Where phonological, syntactic and morphological borrowing is concerned, Romance linguists have come to learn particular caution in attributing phenomena to foreign influence, and there are nowadays relatively few situations for which foreign influence is unanimously accepted. After the advent of structuralism, the number of controversial cases increased and the closer study in the last century of geographical and historical variations showed that some phenomena previously attributed to a particular substrate or superstrate actually recur in other Romance varieties having a different substrate or superstrate, and may therefore be attributable to purely internal development. An alleged borrowing must also be plausible in structural, historical and geographical terms: it must be structurally closely similar to a corresponding feature in the alleged source language, it must be clearly shown to have occurred during a historical period of contact between languages,³⁷ and its geographical distribution should match that of the assumed source language, and not occur extensively outside it.³⁸ If these criteria are not met, then the most likely explanation involves spontaneous, purely internal, development, unmediated by contact. For example, weakening of postnasal consonants in southern Italy; /f/ > /h/ in Spanish attributed to Basque; the alleged Etruscan origins of the Tuscan spirantization of intervocalic voiceless stops; alleged Celtic effects on the vowel and consonant system of Gallo-Romance; rhotacism of /r/ and /l/ in Romanian; postposition of the definite article in Romanian; various phonological processes attributed to Germanic influence on Gallo-Romance.³⁹

Syntax is the domain most open to external influence after the lexicon, but here too it is difficult to distinguish clearly between syntactic borrowing and

37 On this point, see also the example of retroflexion in southern and central Italy and Sardinia, discussed in Loporcaro, volume 1, chapter 3, p. 143.

38 See also Vrabie (1992) and Sala (1998a).

39 Loporcaro, volume 1, chapter 2, p. 85f., gives an example of a pattern of stress-retraction in Franco-Provençal which seems plausibly attributable to contact with neighbouring Alemannic dialects.

parallel independent development, as has often been stressed (Sandfeld 1930; Coseriu 1964), and as we have seen in sections 2.2.2 and 2.2.3. The difficulties are even greater where substrate influence is involved (Hubschmid 1978). In phonology there are phenomena which are certainly borrowed,⁴⁰ and others whose explanation is controversial as to whether they are borrowed or a matter of internal development. The most spectacular cases appear in regional Romance varieties and particularly where there is active bilingualism (see our discussion above on the introduction of the segment /h/ into Gallo-Romance from Frankish (§2.2.2)⁴¹ and Romanian (§2.2.3), effectively circumscribed to loanwords; of /ʒ/ into Romanian (§2.2.3); of neutralization of palatal distinctions in Istro-Romanian (§2.2.3); or /(g)w/ into Spanish from Arabic (§2.2.4).

In inflectional morphology borrowing usually plays an insignificant role, and when it occurs (see, e.g., the discussion of the Romanian vocative ending, or the borrowing of verb endings in Megleno-Romanian, in §2.2.3) what is involved are usually easily segmentable inflectional endings, again involving clear and systematic arbitrary matchings of form and meaning. Under conditions of intense and enduring bilingualism – as in the cases of Istro-Romanian and Megleno-Romanian (§2.2.3) – morphology, too, may be radically affected. There do not appear to be cases of borrowing, for example of patterns of stem allomorphy from one language to another, but we do find instances of borrowing of nominal desinences, especially plural markers (although these are usually contained in loanwords). In Aromanian, there are plural endings in *-ate* (< Grk. *-ata*) in some words of Greek origin (e.g., sg *gramă* PL *gramăti* ‘letter’), and from Greek comes the plural ending *-dzi* which has extended to words of Turkish origin stressed on the final vowel (sg *café* PL *cafedzi* ‘coffee’) and even to a word inherited from Latin (sg *dumnidză* PL *dumnidzazi* ‘god’). In Aromanian there are isolated cases of the Turkish plural desinence *-lár* and *-ân* in personal names; the latter is used even for words of non-Turkish origin.⁴² However, the morphological structure of the Romance languages continues that of Latin, and the modifications which have occurred are overwhelmingly not the result of contact (see also Maiden, volume I, chapter 4). There are very few inflectional elements in the standard Romance varieties unanimously accepted to be loans, but these are more evident in

40 For some details of the phonological ‘re-Latinization’ of Romance produced by borrowing from Classical Latin, see Pountain, volume I, chapter 13, pp. 630–35, 639f.

41 See also Loporcaro, volume I, chapter 3, p. 142.

42 Similarly, Judaeo-Spanish has two endings of Hebrew origin for the plurals of some Hebrew words: *-im* (masculine) and *-ot* (feminine). The only words of Spanish origin in *-im* are *ladronim* ‘thieves’, *ermanim* ‘brothers’, *aznim* ‘asses’, *ratonim* ‘mice’.

dialectal varieties, especially linguistic islands surrounded by other languages, which are not subject to linguistic norms. Yakov Malkiel (1978) has given a succinct and instructive definition of loans in morphology: 'Now, in the inflectional system – and this is what distinguishes it fundamentally from that of derivation – one finds almost no direct trace of loans; rather, one tracks down isolated examples of calques and particularly mutually supportive changes.'⁴³ There is nothing that can be certainly attributed to superstrate or substrate influence. Of the Ibero-Romance verb paradigm, Malkiel declares that it is free of Iberisms, Basquisms, learned forms and even Hellenisms. Romanian is the language showing the greatest number of phenomena normally considered loans, but there is only one in inflectional morphology, the vocative singular suffix *-o* of nouns in *-ă* (see §2.2.3 above). Malkiel says that there are many cases 'of support lent by contact, in a context of bilingualism'; one such involves a Romanian inflectional type not found in other Romance languages, namely the alternation in feminine nouns between stressed final *-ea* (in the feminine nominative–dative singular) and *-ele* (in the genitive-dative singular and in the plural), the result of a historically regular sound change: e.g., *stea*–*stele* 'star' (see §2.2.5). Another example from Romanian nominal inflection involves invariable adjectives. From Latin only one adjective of this type survived, namely *ferice* 'happy' (< Lat. *FELICEM*). This class was swollen by words borrowed from various languages, such as *sadea* 'pure, simple, unalloyed' from Turkish, *atroce* 'atrocious', *eficace* 'efficacious', *gri* 'grey' from Romance languages, and modern borrowings from English, such as *live*, *sexy*.

In general, the number of borrowed elements, and the extent to which borrowing penetrates into grammatical structure, increase if we refer to Romance dialects spoken in regions where there is bilingualism of the kind defined by Malmberg (1990:82f.), who states that: 'someone who gets by more or less in a second language is not a bilingual. In my opinion a speaker is bilingual only if he speaks equally and with the same ease, indeed so perfectly that he is considered in two speech communities as belonging to each.' Such bilingualism exists in some parts of Latin America (the Guaraní of the Andes where Quechua is spoken or the Yucatán region of Mexico, analysed by, among others, de Granda, 1988, and Lope Blanch 1987) or in the case of some sub-Danubian Romanian dialects, the typical example being borrowing of aspect marking in Istro-Romanian.⁴⁴

⁴³ In the original French 'actions solidaires'.

⁴⁴ See also de Granda (1988) for borrowing of aspect marking from Guaraní into the Spanish of Paraguay.

The Romance languages in the Renaissance and after

HELENA L. SANSON

I. Introduction

This chapter¹ is concerned with the external history of the Romance idioms.² It covers the period that runs from the sixteenth century – when various national and literary languages firmly established their prestige alongside Latin – until the end of the nineteenth century,³ when Romanian and Italian eventually acquired the status of national languages.

Although in many cases the rise of literary and national standards had already begun in the later Middle Ages, it was during the early modern period that the state came to be regarded as a linguistic domain and language came to be associated with nationhood. The Romance-speaking areas, as well as western Europe more generally, were crossed by parallel and often interdependent movements of emancipation, codification and standardization of modern languages.⁴ These movements were supported by powerful factors of a political, economic, social and literary nature. According to the oft-cited jocular definition attributed to Max Weinreich, a language is ‘a dialect with an army and a navy’; in other words, a language is a dialect used officially by a state.⁵ The

1 All translations into English are mine, unless otherwise specified.

2 For the fields of interest of the external history of a language, I refer to Saussure (1916:41f.). See also Berschin (2001).

3 Periodization issues (on which see Hupka, 2001, and Wright, this volume, chapter 3) will not be the subject of discussion in this specific context, but, although it is clear that linguistic change and evolution cannot be restrained within strict and arbitrary temporal limits, for practicality’s sake a division by centuries has been adopted.

4 On the emergence and standardization of national languages, see Scaglione (1984), Chiappelli (1985), Coulmas (1985) and Joseph (1987). On standardization in Romance, see also Muljačić (1993). On the notion of ‘standard’, see Lepschy (2002).

5 The word and notion of ‘dialect’ was taken from Greek (where it designated accepted literary varieties, such as Attic, Ionic, Doric and Eolic) and introduced into Italian, and from there into modern European culture, by Italian humanists between the fifteenth and sixteenth centuries (cf. Alinei 1981; Trovato 1984; Lepschy 1996).

creation of a national language is indeed in most cases correlated with the rise of a central political power, with the ensuing consolidation of the nation-state and its machinery.

Spanish, Portuguese, French, Italian and Romanian are the Romance varieties that established themselves, at different historical moments, as the national languages of their respective states, and this chapter pays most attention to them. Nevertheless, care has been taken, as far as possible, to take into consideration a number of other Romance varieties that fared less well on the path to the recognition of national status. Geographically speaking, this chapter will limit its considerations to the external history of the Romance languages within the European continent and will not deal with the spread and development of these idioms overseas.⁶

Traditionally, analysis of the external histories of the Romance varieties tends to focus predominantly or exclusively on the institutional and more officially recognized agents of codification and standardization, and when the history of a language and the history of its literature tend to blend 'it is customary to consider only the prestige of individual authors as creating or spreading the use of a standard language' (Hall 1974:127). This is in itself perhaps an inevitable consequence of the fact that, until the end of the nineteenth century, investigation into the external history of a language had to draw predominantly on written sources. Nevertheless, alongside better-known figures and institutions that had a more directly acknowledged impact on the codification of a language, this chapter will also consider the everyday linguistic reality of the less privileged and the less learned (including women).

2. The sixteenth century and the codification of the vernaculars

The sixteenth century saw the first concerted efforts to analyse and regularize the vernacular languages alongside the development of European national literatures. At the same time, we also witness the emergence of movements for the defence of the vernaculars as opposed to the classical languages, in particular Latin. Long after the Roman Empire had ceased to exist, Latin had continued to enjoy the status of the leading language of culture and was universally accepted and employed in literature, philosophy, theology,

6 On this see, for instance, Silva Neto (1979:427–42, 513–46, 581–634), Cardona (1994), Caudmont (1994), Elia (1994) and Massa (1994) for Portuguese; Lapesa (1980:534–99), Rivarola (1990; 2001; 2005), Penny (2002:22–30) and Sánchez (2002) for Spanish; Brunot (1966–79: VIII (2–3), 1035–194) for French.

history, medicine and other intellectual fields, as well as in law and administration. Latin benefited from a period of renewal and reawakening thanks to the process of rediscovery, establishment, codification and interpretation of the documents of ancient literature carried forward by humanism. This, in turn, sparked an increased interest in ancient, as opposed to medieval, Latin, and improved the need for more sophisticated working aids in textual philology and grammar. From this new approach stemmed the best-known Italian humanist grammars, such as Lorenzo Valla's *Elegantie* of 1449 (first printed in 1471), and the works of Gasparo Veronese (before 1455), Pomponio Leto (before 1467), Niccolò Perotti (1468, first edition 1473), Giovanni Sulpicio Verulano (c. 1470, first edition 1490) and Aldo Manuzio (first edition 1493).⁷ Greek, too, became for the first time since antiquity a major element in intellectual life, and its entrance 'into the humanist system opened up some space for the recognition of the vernacular and its variants by breaking the monopoly of Latin as the language of culture' (Tavoni 1998:46).⁸

In the sixteenth century, the role of Latin as the vehicle of culture, law and government came under threat as the vernacular was increasingly viewed as a possible rival and successor. Paradoxically, the growth in prestige of the vernacular must also be viewed as a result of the humanist attempts to bring fifteenth-century Latin closer to its classical purity, which, as a consequence, made it less suitable to express the requirements of modern life: the Latin language 'was, so to speak, too good for day-to-day concerns' (Rickard 1989:85).

The regularization of the vernaculars further threatened the role of Latin. The long-standing tradition of the prestige of Latin, with its apparent regularity and immobility, contrasted with the variable and rapidly changing vernacular, and it soon became clear that, in order to be able to compete with Latin, and its stability and dignity, the vernacular, too, had to be fixed and regulated. Grammatical codification played a crucial role in ensuring that the vernacular had the resources to contend with its classical antecedents. Inevitably, being the only model they could refer back to, Latin grammar had a strong ascendancy in the minds of early grammarians, with points of divergence in morphology, and to a lesser extent syntax, attracting particular attention.⁹ The vernacular

7 On the Latin grammatical tradition in Europe, see Padley (1976), in particular pp. 5–57 on the humanist tradition. See also Tavoni (1998:2–14).

8 On this point, see Dionisotti (1970). For the birth and development of Greek grammar in western Europe, see Pertusi (1962).

9 On grammatical codification in western Europe in the Renaissance, see Kukenheim (1932); for the period 1500–1700, see Padley (1985; 1988) and Ahlqvist (1987). On the emancipation of vernacular languages in general in western Europe, see Tavoni (1998:14–44). Also very useful is Colombat and Lazcano (1998).

progressively consolidated and strengthened its position and came to be increasingly used in new domains. This does not mean that it triumphed altogether over Latin, but rather that it firmly succeeded in establishing its own literary territory. Indeed, Latin was to remain for at least two more centuries the vehicle for most cultural communication, the usual language of international diplomacy and the medium of instruction in schools and universities, as well as the language of the Catholic Church (except for sermons) (see Waquet 2002).

The newly discovered art of printing was an aid and clear incentive to the dissemination of knowledge in the vulgar tongue. The invention of printing with movable type revolutionized the transmission of knowledge. The consequences of the introduction of the printing press in the last decades of the fifteenth century were profound and wide-ranging:¹⁰ it was in the printer's own interest to cater for as large a public as possible and indeed books became available in much larger quantities than ever before. The increased availability meant a clear cost reduction and therefore an increased affordability of printed books, not to mention the fact that printing had the further advantage over manuscripts that multiple copies of texts could be disseminated more quickly and more widely. It followed that publishers gave to the presses an ever-growing number of works in the vernacular and of translations from Latin and Greek (and to a lesser extent Hebrew), as well as from other European languages. Books in the vernacular catered also for those – the less learned, and among these even women – who had neither the leisure nor the means to access classical studies. With literacy rates growing, the standardization of the vernacular beyond local variations – in its orthography, morphology and syntax – became a pressing matter in order to promote better and wider book circulation, and thus ensure higher profits. The need for a more homogeneous language therefore also encouraged the process of metalinguistic reflection and grammatical codification.

The Reformation, with its demand for the laity to have direct access to the Bible, was another important factor that contributed to the codification of some Romance varieties. The Scriptures had to be made available to all and that meant translating them into a vernacular accessible to everyone. In a politically divided country, Luther played a crucial role in the codification of the German language with his translation of the New Testament in 1522, and

10 On the invention of printing in Europe, see Hirsch (1967). On the revolutionary impact of the printing press on the transmission of knowledge in early modern Europe, see Eisenstein (1979). On printing in Italy, see Richardson (1994; 1999); in France, Claudin (1900–15) and Chartier (1987); in Spain, Norton (1966) and Martín Abad (2003); in Portugal, Anselmo (1981); in Romania, Andreescu (2002).

then of the entire Bible in 1534, for which he adopted the language of the politically prestigious chancery of Saxony. Similarly, with Calvin's *Institution de la religion chrestienne* (1541; originally published in Latin in 1536), French made a resounding entrance into the field of theology, the most heavily guarded bastion of Latin. In Italy, too, many religious works became available in translation in the first part of the sixteenth century, among them even the Scriptures, in Nicolò Malerbi's 1471 version of the entire Bible and in 1530 and 1532 with the New and Old Testaments respectively by Antonio Brucioli. But later, in the spirit of the Counter-Reformation, the Index of Pope Paul IV, in 1559, prohibited vernacular Bibles being printed, read or kept without the permission of the Holy Office, an interdiction which was followed by even stricter prescriptions by the Indexes of Sixtus V in 1590 and Clement VIII in 1596.¹¹

In Italy, favourable cultural, economic and social circumstances meant that Gutenberg's invention found a particularly fertile ground in which to prosper (Richardson 1999:3–5).¹² In the peninsula, though, there was no real social or political force, outside literary society, that could act in favour of linguistic unification, and not surprisingly the 'language question' developed there some decades earlier than in other European countries and soon took on enormous weight. By *Questione della lingua* we refer to the linguistic controversies that sprang up from the first half of the sixteenth century regarding the written literary language to be adopted (and its definition):¹³ Which form of the many vernaculars in use in the peninsula could aspire to be the new language of culture? Against the upholders of contemporary Tuscan and a supra-regional, eclectic, but still Tuscan-based, 'lingua cortegiana'¹⁴ (the language of the courts), the line that prevailed was the vernacular classicism espoused by the Venetian humanist Pietro Bembo (1470–1547). Bembo transferred to the vernacular the theories on imitation of the classics (so dear to

11 See Barbieri (1992), Rozzo (2000) and Waquet (2002:47–50). On the prohibition of vernacular Bibles and the severity of ecclesiastical censorship at the end of the sixteenth century, see Fragnito (1997). For the Italian language in the Church, see Librandi (1993).

12 But on the continuing use of scribal transmission in Renaissance Italy even after the arrival of printing, see Richardson (2009).

13 There is a considerable literature on the *Questione*. For a first account, one could consult, in English, Hall (1942) and Migliorini and Griffith (1984). In Italian, see Vitale (1984) and Marazzini (1993b; 1994). Specifically on the sixteenth century, see Bruni (1984:36–66), Koch (1988:350–54), Marazzini (1993a) and Trovato (1994). On women's involvement in the *Questione* debates, see Sanson (2010).

14 On the 'lingua cortegiana', see Drusi (1995), Giovanardi (1998) and Tesi (2001). For an illustration of the tenets of this theory, see the dedicatory letter and Book I, pp. xxviii–xxxviii, of Baldassar Castiglione's *Il Libro del Cortegiano* (1528).

some humanist men of letters in their approach to Latin), only this time directing his admiration not to Cicero or Vergil, but rather to the great Florentine writers, in particular Boccaccio for prose and Petrarch for poetry (and partly also Dante). In 1525, in his *Prose della volgar lingua*, he expounded his ideas in the form of a fictional Ciceronian dialogue (in three books, of which the third is a non-systematic grammar of the literary language), exalting the role of fourteenth-century Florentine as the best-regulated language. Dismissing other competing theories, Bembo believed that it was inappropriate for a literary language to be too close to everyday speech because in so doing it lost gravity and greatness (*Prose* I, 18). In favouring fourteenth-century Tuscan as the basis of the literary language, rather than contemporary Tuscan, Bembo might seem backward-looking. In fact, he defended a solution that was politically neutral and *super partes*, free from any subservience to any of the rival political powers of the peninsula, and which had the advantage of providing a well-defined and homogeneous linguistic model to refer to, supported by an indisputable literary prestige and wide appreciation.

The drawbacks that Bembo's views implied were nevertheless evident and were to mark the linguistic, literary and social development of the peninsula throughout the following centuries. This was a 'language created by members of the élites, for the élites' and which 'tended by its nature to be socially exclusive' (Richardson 2002:13). By promoting and fixing as the literary language a variety from over two centuries earlier, Bembo was in practice inevitably creating a gulf between the literary medium on the one hand and everyday speech, people's own local vernacular, on the other.¹⁵ In literary production, though, the appeal of Bembo's views was undeniable and saw its first clear results soon after the publication of the *Prose*. The poet Ludovico Ariosto, from Ferrara, for instance, decided to revise the language of his chivalric poem *Orlando Furioso* so that the third edition of this work (of 1532) conformed more closely to Bembo's taste and linguistic principles. Similarly, the Florentine Francesco Berni felt the need to rewrite Count Matteo Maria Boiardo's fifteenth-century chivalric poem *Orlando innamorato*, composed in the Padan *koinè*, the ennobled vernacular of the Po Plain, in a Tuscan conforming to Bembo's taste (c. 1531; published posthumously in 1542). Berni's *Rifacimento*, that is, his recasting of the poem, was extremely popular in the later sixteenth century, becoming better known than Boiardo's original.

¹⁵ On sixteenth-century observations on spoken language, see Maraschio (1977). On the discussions on spoken Tuscan, its qualities and defects, and the influence it could have on the literary language, see Richardson (1987).

The *Questione* saw a wide-ranging, rich and detailed production of texts, which did not leave any linguistic issue or polemic untouched.¹⁶ In Italy, as in the rest of Europe, the standardization of the vernacular came up against the question of orthographic irregularities:¹⁷ the fact that the vernacular had been used for a long time in writings of a practical nature, often reflecting local usage or even the scribe's own habits and linguistic preferences on the issue, meant that these texts had been 'reproduced in an endemic state of formal variability' (Tavoni 1998:18). With the spread of printing, these practices had to be curbed, in favour of a more uniform and rationalized system.¹⁸ The movement for spelling reforms in Italy, for instance, saw among its main participants Giovan Giorgio Trissino, from Vicenza (author also of the *Castellano*, a central text in the *Questione* debates and of the brief *Grammatichetta*, both of 1529). In November 1524 Trissino published the *Epistola* [...] *de le lettere nuovamente aggiunte ne la lingua Italiana* to defend the new graphic system used in his tragedy *Sophonisba*, published in September of the same year. According to Trissino, new letters were needed, because the traditional alphabet was inadequate to represent the sounds of the vernacular. Among the new graphemes introduced, Trissino used Greek letters to distinguish between open and close *e* and *o*, respectively ε/e , ω/o , only to switch, from 1529, the value of ω/o , with ω now being used for the close *o* and the Latin letter for the open (no changes were made to ε/e) (see Castelvechi 1986:xiii–lvii). His suggestions, though, were met with great hostility and controversy by other men of letters, such as, for instance, Ludovico Martelli, Claudio Tolomei and Angelo Firenzuola, and in the end failed to catch on.¹⁹

As for grammatical works, the first vernacular grammars in Italy were produced by the elites for the benefit of those who were 'fortunate enough to have a good education in Latin', namely 'a small minority of males and a tiny minority of females' (Richardson 2002:15). After the short manuscript grammar of Florentine written by Leon Battista Alberti in the fifteenth century (1437–41, subsequently forgotten for centuries and only printed for the first time as an appendix to Trabalza 1908), grammar codification was led

16 For an anthology of primary texts on the *Questione*, see Pozzi (1978; 1988).

17 The codification of punctuation alongside that of spelling was clearly important in relation to the spread of the printing press. On the history of punctuation in Europe, see Mortara Garavelli (2008).

18 On the Italian language and print in the sixteenth century, see Trifone (1993).

19 For a detailed discussion of sixteenth-century orthographic issues in Italy, see Richardson (1984), which also reproduces all treatises from the period 1524–26 (Trissino, Firenzuola, Martelli, Tolomei, Liburnio).

by non-Tuscan authors for the benefit of a learned readership, who had to approach the language of the great fourteenth-century authors with the same dedication and the same efforts needed to acquire Latin.²⁰ The first printed grammar of the vernacular was Giovan Francesco Fortunio's *Regole grammaticali della volgar lingua* printed in Ancona in 1516, followed in the first decades of the century by, among others, Nicolò Liburnio's *Le vulgari elegantie* (1521) and *Le tre fontane* (1526), the *Grammatichetta* of Trissino, *La grammatica volgare* (1536) of Alberto Accarisio, *Le osservazioni* (1539) of Francesco Alunno, the *Regole* (1545) of Paolo del Rosso, the *Regole grammaticali* (1548) by Jacomo Gabriele and the *Osservazioni* (1550) by Lodovico Dolce. After the 1530s, with the political crisis and the breaking down of courtly ideals, came a widening of literary society and the erosion of its hierarchy, which is reflected also in the intended readers of the first vernacular grammars: these texts ceased to be works only for scholars and tried to extend their reach also to beginners, foreigners and to the less learned, including women. Gaetano Tizzone's *La grammatica volgare* of 1539 is dedicated to a noble woman, Dorotea Gonzaga, marchioness of Bitonto, and Rinaldo Corso's *Fondamenti del parlar thoscano* of 1549 was composed not only for foreigners but first and foremost for the woman he was to marry, Lucrezia Lombardi, Hiparcha in the text. In his grammar he explained how he wanted to guide her in the appreciation of the language and works of the great authors of the past, thus structuring his *Fondamenti* to meet the needs of his envisaged, privileged, reader (Sanson 2005; 2007). Girolamo Ruscelli intended the third book of his *Commentarii* (in seven books, published posthumously in 1581) to be a brief and clear grammatical exposition for the benefit of 'women, children, young people and foreigners and all those who did not know Latin' (Ruscelli 1581:377).²¹

If literary works had crucially assisted the establishment of a standard by their prestige, there was no corresponding unity in terms of 'spoken' language. Throughout the sixteenth century, members of the social and cultural elites normally used, in informal conversation, 'the dialect of their region, or at least a type of language that was very close to it' (Richardson 2002). Clearly, this is even more the case among the lower classes: with their non-existent or very limited literacy, they had no easy access to the literary language and were confined to their maternal dialect. Indeed, Italian remained for centuries too

20 On the Italian production of grammars, see Trabalza (1908), Quondam (1978), Swiggers and Vanvolsem (1987), Poggi Salani (1988) and Patota (1993).

21 On the democratization of grammar production in sixteenth-century Italy to include women, see Sanson (2011:83–114).

remote from everyday life and too formal to become popular.²² A condition of diglossia (Ferguson 1959) prevailed in the peninsula at least since the Renaissance, with literary Italian, used almost exclusively in writing, as the High variety, and the local dialects, non-literary varieties, as the Low variety (Lepschy *et al.* 1996:71). By the second half of the century a clear distinction was therefore in place between the archaic Tuscan used in formal writing, accepted across the peninsula, and spoken Tuscan, merely regional, which normally was not meant to invade the realm of spoken language outside the borders of Tuscany (Richardson 1987:104).

In France, Spain and Portugal, discussions on language followed those that took place in Italy and were often initiated by scholars and men of letters who had lived in Italy and had personally come into contact with the *Questione*. More favourable political conditions in these countries meant that the debates were concerned less with the issue of which vernacular to adopt as the language for literature and rather more with ennobling their respective vernaculars by showing how they were either as prestigious as the classical languages or could be made to be so.

In France, for instance, unlike in Italy, there was a capital, Paris, and a court that acted as a centripetal force for language codification. Among the various medieval dialects of the north and centre, known collectively as the *langue d'oïl*, and the Occitan dialects of the south, known as the *langue d'oc*, the dialect of the Île de France (also known by the modern name of Francien) was the one that developed into modern standard French.²³ By the end of the fifteenth century, it was the only literary medium in the north and the spoken language was to an increasing degree based on the language of Paris and other cities such as Orléans, Tours and Chartres, with regional pronunciation being widespread and not carrying any particular social stigma (Rickard 1989:82). The adoption of Francien as the state language from the sixteenth and

22 But in the sixteenth, seventeenth and most of the eighteenth century, Italian was used as the chief diplomatic language between Europeans and Turks in the Ottoman Empire and, in the seventeenth century, as the lingua franca of merchants and seafaring people around the Mediterranean. See on this Cremona (2002).

23 A third group is formed by the Franco-Provençal dialects, identified in 1873 by the Italian linguist Graziadio Isaia Ascoli in his essay *Schizzi franco-provenzali*, comprising a smaller intermediate area in the mountainous regions of south-eastern France (Savoy), together with the varieties originally spoken in western parts of Switzerland (Vaud, Neuchâtel, Valais), as well as over the Alps in the Val d'Aosta (Valdôtain) in the north-west of Italy. There has never been a koineized version or any substantial literary output (but on Franco-Provençal literary production before 1700, see Tuaillon 2001). For an outline of the external history of these dialects, see Martin (1990).

seventeenth centuries was a purely political move, supported not only by the establishment of a fixed royal court in Paris, but also by the development of an educational and legal system centred on that city, and the fact that the nearby abbey of Saint-Denis was the spiritual centre of the kingdom. Latin and Occitan, too, were used until then in administration and for written records, but the strengthening of state functions required the use of a common language. This is why, with the royal Ordonnance of Villers-Cotterêts of 1539, under François I, the *languaige maternel francoys* ('French mother tongue'), that is, the *langue d'oïl*, was officially recognized and imposed as the language to be used in the courts of law, as well as in public and private contracts, to the exclusion not only of Latin but also of other vernaculars (but see Fiorelli 1950:283–88). Nevertheless, by that time, the precise details of the *languaige francoys* were far from fixed. And French, unlike Italian, having evolved quite rapidly in previous centuries, lacked equally prestigious literary models to refer to. Works on language in France tended therefore to focus rather on the merits of French as opposed to Latin and, to a certain extent, to Italian. In the course of the sixteenth century, Italian indeed attracted widespread admiration for its refined Petrarchist poetry and came also to be internationally recognized as the language of elegant conversation (see Richardson 2002).

Joachim du Bellay, one of the seven poets who founded the *Pléiade* (named after the constellation of seven stars), in his *Deffense et illustration de la langue françoise* (1548; modelled very closely on Sperone Speroni's *Dialogo delle lingue*, 1542), aimed to legitimize the French vernacular against both Latin (and to a lesser extent also Greek) and Italian. The core argument of his treatise was that French was not inferior to other tongues and that authors should not feel uneasy in writing in this language. For French to be cultivated, improved and rendered illustrious (*illustration* means 'giving distinction'), translation was not enough and original works were needed too, always with an eye to imitation of the ancients. Similarly, in the *Traicté de la conformité du langage françois avec le grec* (1565), Henri Estienne's aim was to demonstrate that, if Italian was related to Latin, then French was superior to Italian on the grounds of its closeness to Greek, this being, in his view, an even superior language. Estienne wrote at the time when the Italianized court of Catherine de' Médici (who had married Henry II in 1533) seemed to exacerbate the feeling of cultural inferiority to Italy among the ruling elites. This persisted for the rest of the century, so that in 1579 Estienne felt the need to reiterate his belief that Italian was clearly inferior to French in his *De la precellence du langage françois*.

In the 1530s the first printed grammars of French, of a descriptive nature, appeared:²⁴ *Lesclarcissement de la langue francoyse* (1530) by the Englishman John Palsgrave, and the etymological grammar *In linguam gallicam Isagoqe* (1531) by Jacques Dubois. *An introductorie for to lerne to rede to pronounce and to speake Frenche trewly* (1532) by Gilles du Wes and the *Gallicae linguae institutio* (1550) by Joannes Pilotus were practical grammars of French, based on the teaching experience of their authors. In the second half of the century, the main grammars of French were Robert Estienne's *Traicté de la gramraire françoise* (1557) and Pierre de la Ramée's *Gramerę* (1562; revised in 1572).

In France the orthographic issue was a serious one, given the etymological nature of the spelling system, with the presence of numerous graphic Latinisms and superfluous letters, coupled with the stronger phonetic evolution of the language away from Latin. A controversy developed between those who aimed to preserve the status quo and those who pushed for innovation (see Catach 1968:29–50; 2001:97–164). In 1529 the printer Geoffroy Tory in his *Champ Fleury* suggested – but did not actually use (Catach 2001:106–9) – the introduction of diacritics in spelling, such as the use of the apostrophe, of accents and of the cedilla (ç), which, for the first time in France, violated the seemingly untouchable Latin writing system (Principato 2000:24). Tory's work was followed by, among others, the anonymous *Tres utile et compendieux traite de l'art et science d'orthographie gallicane* (1529), the *Isagoqe* by Dubois mentioned earlier, and the *Accents de la langue francoyse* (1540) by Étienne Dolet. Then, in 1542, the humanist Louis Meigret, from Lyon, printed the *Traité touchant le commun usage de l'écriture françoise*, giving the orthographic debate a completely new dimension. Meigret preached the need for an urgent reform of the French spelling system that had to be based on phonetic criteria, hence promoting the loosening of the grip of Latin etymology on French. He put his principles into practice in his 1550 *Trehtë de la gramme françoëze*, the first grammar of French written in French by a French author and a valuable source of information about sixteenth-century pronunciation. His suggestions for reform met with considerable opposition and ultimately had little impact, not least because they were set against the principles applied in Robert Estienne's highly influential *Dictionnaire francoislatin* of 1539–40, where traditional etymological spelling had been employed. At a time when printing

24 The bibliography on French grammars is vast. For an outline of the French production of grammars from the sixteenth to the twentieth century, see Swiggers (1990). See also Chervel (1982) for the period 1800–1914. Brunot (1966–79) is also an invaluable source of information.

conventions were being set, those in favour of a phonetic spelling system based their views on the fact that writing had to reflect the spoken language rationally. In the end, though, French spelling remained etymological in nature, with all the complications it brought with it: a system with etymological spelling criteria and other arbitrary irregularities meant, for instance, that literacy was more easily restricted to the learned.

In Spain, in the twelfth and thirteenth centuries, the dialects of Aragon, Asturias, León, Castile and Galicia were about equal to each other in prestige for administrative and literary use, but already under the reign of Alfonso X el Sabio (Alfonso the Wise, 1252–84) Castilian had replaced Latin as the official language of the chancellery, thus marking an important step in its affirmation as a language of culture. The Arab occupation and the strong presence of Jewish settlements had formed, throughout the centuries, a multilingual society where deference towards Latin was not as strong as in other European countries and where Castilian, accepted as the common language by both Arabs and Jews, was then promoted under the initiative of the court at an early date (Tavoni 1998:17).²⁵ With the *Reconquista*, Castilian took over most of the peninsula, as it swept southwards. A series of alliances, as well as military successes, meant it engulfed first (from the thirteenth century) the Leonese and then (at the end of the fifteenth century) the Aragonese territories and the south, with the taking of Granada, the last Moorish bastion, in the year 1492. At the end of the fifteenth century the promotion of Castilian as a national language, worthy of a kingdom and fit to spread through the country, was strengthened by a series of factors: the establishment of the *Reyes Católicos* ('Catholic Monarchs', a title conferred by Pope Alexander VI in 1494), that is, Isabella I of Castile (1474) and Ferdinand of Aragon (1479) – with the queen encouraging the study of literature and the use of Castilian in all kinds of fields and contexts – the discovery of America in 1492 and the political aspirations and holy vocation of Charles V to make Spain the centre of a vast Catholic empire that would withstand the pressures of the Muslim faith. In 1526 Charles V declared Spanish to be the language of diplomacy and the common language of all Christianity.²⁶ The foundation of the influential order of the Jesuits, the *Compañía de Jesús*, in 1540 by Ignacio de Loyola, further reinforced the political standing of Spain at an international level. In 1535 Juan de Valdés

25 See on this also Niederehe (1975).

26 It is worth pointing out that the imposition of Castilian as the official language of the state dates from only the early eighteenth century, with the ascent of the Bourbons to the Spanish throne.

observed that Castilian was spoken across the country by the populace and nobility alike, thus testifying to the fact that by then it had become firmly established as a national language (Lapesa 1980:298). From the sixteenth century onwards, the more inclusive expression *lengua española* progressively starts to become dominant alongside the more traditional *lengua castellana* (Alonso 1968; Lapesa 1980:299; Cano Aguilar 1988:227–29).

As for the production of grammars, the Spanish tradition can boast the first printed vernacular grammar in the western world, but, compared to other Romance languages, the remainder of its production was very late and infrequent in the course of the century.²⁷ The *Gramática sobre la lengua castellana*, by Elio Antonio de Nebrija (1444–1522), ceremoniously presented to Queen Isabel in 1492, is the expression of the humanist interest in the description and codification of the vernacular following the Latin grammatical model, as well as an interesting work of political propaganda in which the author linked his reflection on language to the ambitious project of the Catholic Monarchs. But Nebrija's grammar remained for decades an isolated (and soon forgotten) case and it was only in 1555 that another grammar of Spanish, the *Util y breve institution para aprender los principios y fundamentos de la lengua hespañola* (published anonymously in Louvain, but generally attributed to Francisco de Villalobos), appeared, followed by the *Gramática castellana* by Cristóbal de Villalón (Antwerp, 1558) and the anonymous *Gramatica dela lengua vulgar de España* (Louvain, 1559). All grammars of Spanish in the sixteenth century were published outside Spain and were written mostly for the benefit of foreigners (Tavoni 1998:39).²⁸ They were of a practical nature and aimed at a public interested in acquiring a good pronunciation and knowledge of Spanish orthography and morphology.²⁹ The influence of the Spanish language and culture across Europe, recognized with the Peace of Cambrai (1529) and the Congress of Bologna (1529–30) and later strengthened by the Treaty of Cateau-Cambrésis (1559), encouraged the publication of a

27 For a general outline of the Spanish production of grammars from the fifteenth century onwards, see Neumann-Holzschuh (1992).

28 See, for instance, *Il paragone della lingua toscana et castigliana* (Naples, 1560) by Giovan Mario Alessandri and the *Osservationi della lingua castigliana* [...] *ne' quali s'insegna con gran facilità la perfetta lingua spagnola* (Venice, 1566) by Giovanni Miranda, as well as Richard Percyvall's *Bibliotheca Hispanica* (London, 1591) and César Oudin's *Grammaire et observations de la langue Espagnolle* (Paris, 1597). Only in the seventeenth century did the production of grammars return to Spanish territory, with the *Arte de la lengua española castellana* by Gonzalo Correás (Salamanca, 1625) as the most important grammar of the *Siglo de Oro* (Swiggers 2001:498).

29 On the Spanish production of grammars in the sixteenth and seventeenth centuries, see Ramajo Caño (1987), and for the period 1500–1800 Niederehe (1994; 1999; 2005).

growing number of grammars of Spanish for foreigners, and of bilingual dictionaries, in Italy, France and England in the following century (see Gallina 1959; Steiner 1970; Niederehe 1987; 1988; 2005; Alvar Ezquerro 1992, esp. 638–42).³⁰

In Spain the Church was in favour of the use of the vernacular in religious matters until the Counter-Reformation period, but after 1559 and the publication of the Spanish Index, the *Catalogus librorum qui prohibentur*, the situation changed. Philip II's linguistic policy in the second half of the sixteenth century counterbalanced this lack of favour for the vernacular by the Church. Latin was the language of schooling and the educational system was very much in the hands of the Jesuits, but in 1583, at the time of the foundation of the *Academia de Matemáticas*, Philip proposed that instruction should take place in the common language and a few years later, in 1588, promulgated a regulation according to which an examination had to be passed by those who taught Spanish and only *cartillas*, that is, basic texts to teach children, published with the permission of the Royal Council, should be employed to teach reading and writing (Hernández González 1992:409).

Literary old Catalan, very close to old Occitan, culturally and linguistically, had enjoyed great prestige in the fourteenth and especially the fifteenth centuries.³¹ After 1412, as a result of the Compromise of Caspe, Catalonia was ruled by the Castilian dynasty of the Trastámaras, in the person of Ferdinand I, and Castilian therefore became the familiar language of the court. Later in the century, in 1474, following the union of Castile and Aragon, the court withdrew even further from the Aragonese territories, and the Catalan aristocracy, as a result, was attracted more and more by the Castilian-speaking court: the 'royal power felt itself to be Castilian, linguistically speaking, and the nobility and men of letters imposed the Castilian language' (Joan *et al.* 1994:123). With the disappearance of the court and the chancellery in an epoch when literature and the literary language were essentially courtly, and with the mercantile classes failing to create their own culture (the centre of commerce and trade moved from the Mediterranean to the Atlantic after the discovery of America), Catalan literary production entered a period of decline, usually known as *La Decadència* (see Comas 1978), which set in after the death of the humanist Joan Roïç de Corella (1497) (Joan *et al.* 1994:137–45; see Nadal and Prats 1996:472–512).

30 For political reasons, relations between Italy and Spain became particularly intense from a literary and linguistic viewpoint and even more so during the seventeenth century; see Croce (1917) and also Beccaria (1968). On the influence of Spanish on Italian, see also D'Agostino (1994).

31 For an outline of the external history of Catalan from the Middle Ages to the twentieth century, see Lüdtke (1991).

From that moment on, until the second half of the nineteenth century, there were 'no major writers in Catalan and until about a hundred years before this, the literary scene was one of almost unrelieved mediocrity' (Terry 2003:53). Besides, with Valencia, where bilingualism had deeper roots, drawing away from the influence of Barcelona and its writers going over almost entirely to Castilian, after 1500 Catalan saw a weakening of its linguistic norms and an increasing fragmentation into dialects (Casanova 1991:220; Terry 2003:54). The decline was one in standard, in terms of sophisticated writing, and not in quantity, however, because popular literature, by contrast, with satirical poetry, popular songs and ballads, remained a genuine and vital source of literary expression throughout the period (Joan *et al.* 1994:159f.; Terry 2003:54f.). As for the spoken language, if the aristocracy and the ruling classes adopted Castilian as the idiom considered most prestigious and distinguished, the vast majority of the population continued to be monolingual speakers of Catalan until the end of the nineteenth century (Joan *et al.* 1994:127, 144f.).

In the sixteenth century, the European territory of Portugal was (except for Olivença)³² identical to that of the modern country and coincided almost exactly with the contemporary area of extension of the Portuguese language. Its borders had been established in the second half of the thirteenth century and by 1500 Lisbon had already been for a long time the centre of political, social, religious and economic life in Portugal. The main religious centres of Alcobaça and Santa Cruz de Coimbra, with their monasteries, were located in the same region, and the university (founded in Lisbon in 1288 or 1290) was definitively transferred from Lisbon to Coimbra in 1537. The linguistic norm of Portuguese was established therefore in this central-southern region, where political power, economic supremacy and social prestige coexisted. The linguistic usage of Lisbon has been preferred by common consensus ever since (Teyssier 1994:462). By the middle of the thirteenth century, Portuguese was already used as the language of many public and private documents, and towards the end of that century it was officially adopted as the written language, instead of Latin, quickly replacing it also in many ecclesiastical deeds (de Oliveira Marques 1972:104). Not surprisingly, then, the 'language question' never took on the proportions or intensity it had, for instance, in Italy.³³

32 Olivença was Portuguese until 1657 and then again between 1668 and 1801. It is now part of Spain, although Portugal does not recognize the latter's sovereignty over it.

33 On the 'language question' in Portugal, see Stegagno Picchio (1959) and Carvalhão Buescu (1984a:217–36). On the Portuguese language in the sixteenth century, see Silva Neto (1979:443–511).

In Portugal, the first grammars of the language were Fernão de Oliveira's *Grammatica da lingoagem portuguesa* (1536), which devoted particular attention also to phonetics and orthography, followed by the *Grammatica da lingua portuguesa* of João de Barros (1540), with a clear normative stance and particular attention to the relationship between Latin and Portuguese.³⁴ Later in the century, we find the *Regras que ensinam a maneira de escrever e orthographia da lingua portuguesa, cum hum Dialogo que a diante se segue em defensam da mesma lingua* (1574) by Pero Magalhães de Gândavo and the *Orthographia da lingoa portuguesa* by Duarte Nunes do Leão (1576).³⁵ All grammarians of sixteenth-century Portugal involved with the codification and renovation of the language were preoccupied with the issue of orthographic reform, with two coexisting forces pulling towards a phonetic tradition on the one hand and an etymological one, of humanist influence, on the other. But whereas Fernão de Oliveira and João de Barros, active at the time of John III, were concerned with both orthography and the origin of the Portuguese language and its relation to Latin (as well as Greek and Hebrew), dealing also with syntax, morphology and pedagogical methods to teach the language, Magalhães Gândavo and Nunes do Leão focused more specifically on orthography. As with other Romance languages, there was a desire to link the vernacular to the Latin model, but there was also the need to introduce new graphemes for phonemes (/ʒ/, /v/, /f/, /k/, /ɲ/, etc.) absent in Classical Latin, and to abandon those letters that had become redundant in Portuguese (*k*, *y* and, except in digraphs, *h*). Indeed, a universally accepted orthographic reform was yet to be devised and accepted and by the end of the century texts still presented many incongruities and archaisms.

Moving away from the realm of linguistic codification to consider instead basic literacy in Portuguese, it is worth mentioning the role of the *cartilhas*, at that time called *cartinhas*. In print since the end of the fifteenth century (and in use until the end of the eighteenth), the *cartilhas* were generally composed of two parts, namely the alphabet and the very first rudiments of language followed by texts of Christian doctrine, occasionally illustrated by basic notions

34 Together with the first and second edition (1785) of the *Gramática* was also published the *Diálogo em louvor da nossa linguagem*, in which Barros attributed to Portuguese some essential qualities, such as gravity, majesty, richness of vocabulary, conformity with Latin, and he rejected the theory of the superiority of Castilian over his own language (on Barros, see Carvalhão Buescu 1984b:29–108).

35 For an outline of the Portuguese production of grammars, see Woll (1994). For a more detailed discussion, see Carvalhão Buescu (1984b) on the sixteenth century and Schäfer-Priess (2000) for the period 1540–1822.

of hygiene and behaviour (de Oliveira Marques 1999:447f.; Mendes Drumond Braga 2001:467f.). In this way, religious teachings were inculcated at the same time as basic literacy. The *cartilhas* were in use in those schools, which started to appear and spread from the end of the fifteenth century onwards, where rudimentary literacy skills were imparted to children who learnt to read and write in the national language (de Oliveira Marques 1999:468).

Portuguese, though, was not the only language of sixteenth-century Portugal. Latin, for instance, was used daily, both in speaking and in writing, by thousands of members of the clergy (de Oliveira Marques 1999:452). It still maintained the usefulness and prestige it had enjoyed for centuries and was a necessary acquisition for those who embarked on a political, diplomatic, academic or bureaucratic career, as well as those who wanted to consider themselves educated. Many grammars of Latin were published in Portugal in the course of the century, with the work of the educator and Jesuit father Manuel Álvares (from Madeira), the *De institutione grammatica libri tres* (Lisbon 1572), acquiring great prestige throughout Europe for centuries (600 editions, of which 25 were Portuguese), thanks to its new methodological approach. Castilian, too, for a long period, between the middle of the fifteenth century and the second half of the eighteenth, was another language of culture, with peaks of stronger influence at specific times, such as between 1580 and 1640, when the king of Spain ruled both countries. But the Spanish language had imposed itself strongly even before that time, because of the frequent marriages of Spanish princesses or *Infantas* into the Portuguese royal family. Between 1500 and 1517, for instance, during the reign of Maria of Aragon, the second of the three Spanish wives of Manuel I, Castilian was strongly encouraged in writing by the queen, who did not speak Portuguese well (de Oliveira Marques 1999:454). Spanish–Portuguese bilingualism during those periods was a social reality and Spanish undoubtedly a language of prestige. It was commonly in use in the highest social circles, where the fashion was set and from where it spread, but progressively lost its impact moving down the social scale: among the lower strata it was basically non-existent (Teyssier 1994:466). Portuguese literature from this period reflects this bilingualism and the majority of the great writers of the time used both. This is the case, for instance, with the poets of the *Cancioneiro Geral* (1516), of Sá de Miranda, Gil Vicente and Luis de Camões (and in the seventeenth century with Francisco Manuel de Melo).³⁶ Jorge de Montemor even went as far as Hispanizing his

³⁶ For a more detailed discussion of Spanish–Portuguese bilingualism at the time, with specific reference to Gil Vicente, see Teyssier (1959).

name to Montemayor, giving up Portuguese altogether. Others simply refused to use Castilian, like António de Ferreira.

Among western Romance languages not elevated to the status of national or literary languages, the Ræto-Romance varieties were in use in areas which never came to form a single administrative unit or a homogeneous linguistic or cultural entity.³⁷ Among the three principal subtypes that are often described as making up the Ræto-Romance group, the one spoken in the Swiss canton of Grisons (comprising the Romansh dialects of the Rhine valley, of which the best known is Surselvan, and the Engadinish dialects of the Inn valley) has fared better than Friulian and the varieties in use in the Dolomites (Ladin), because it benefited from attaining literary status with the Reformation.³⁸ Apart from a few scattered attestations ascribed to the tenth and eleventh centuries, the earliest major material in Romansh was indeed brought into being by the religious situation prevailing in the sixteenth century. The earliest printed works are in the Protestant Engadine variety: they were Jachiam Bifrun's catechism *Una cuorta et christiauna Fuorma da introduider la giuventiuna* of 1552 and his translation of the *New Testament* in 1560, *L'og Nuof Sainc Testamaint, miss in Arumaunsch* (Holtus 1989:858–59).

As for eastern European Romance, the only linguistic standard created in the area is Daco-Romanian,³⁹ of which the variety spoken in Bucharest (the main centre of Muntenia, in Wallachia) is the accepted modern standard.⁴⁰ The antecedent of Romanian has its earliest attestation in a letter, in the Cyrillic alphabet, which dates from 1521: written by the boyar Neacșu of Câmpulung (Mușcel), it was addressed to Hans Benkner, the mayor of Brașov in the Carpathians, informing him of a movement of Turks from

37 On Ræto-Romance, see Arquint (1964), Rohlfs (1975), Haiman and Benincà (1992) and Liver (1999).

38 For an outline of the external history of Romansh, see Holtus (1989), on its production of grammars, see Lutz (1989), on its lexicography, see Dazzi and Gross (1989). For Ladin, see Heilmann and Plangg (1989), Kramer (1989) and Belardi (2003). For a sociolinguistic history of Ladin, see Belardi (1991). For an external history of Friulian, see Marcato (1989a), and on its production of grammars and lexicography, Marcato (1989b).

39 Daco-Romanian derives from the Latin spoken in the Roman province of Dacia, north of the Danube. 'Daco-Romanian' is a learned coinage introduced by linguists and historians to distinguish the northern branch of eastern Romance from the sub-Danubian ones, that is, Istro-Romanian, Megleno-Romanian and Aromanian (Macedo-Romanian). Daco-Romanian has in turn two principal dialects, Moldovan and Muntenian (Wallachian), whereas Transylvanian is considered a transitional area between them rather than a third dialect (Harris 1988:22).

40 On the external history of Romanian, see Rosetti (1973) and Arvinte (1989). On its production of grammars, see Turculeț (1989), and on its lexicography (from the nineteenth century onwards), see Winkelmann (1989). For an overview of its main features, see Rosetti (1938–41) and Tagliavini (1972:356–74).

Sofia along the Danube (Rosetti 1973:97; for a transcription of the letter, see p. 112). In the three Romanian principalities of Moldavia, Wallachia and Transylvania, the language of written culture in use was Church Slavonic, as well as, because of the continuous ties with western Europe, Latin and Greek.⁴¹ Monasteries and churches were important centres of Slavonic writing and so were the princely courts. Latin writing developed in Transylvania, since it was the official language of the Kingdom of Hungary and of the Catholic Church. Here Latin culture flourished in the cities, the episcopal sees, the monasteries and the princely chancellery. By contrast, the vernacular was employed for the oral creations of the rural world, made up of lyric and epic folk literature, in the form of legends and fairy tales, alongside a court oral literature with heroic and historical ballads.

Politically speaking, Wallachia and Moldavia had been independent from the Hungarian kingdom since the fourteenth century, although they were linked to it by a system of vassalage and alliances. In the fifteenth century, from 1417 and 1456 respectively, Wallachia and Moldavia had agreed to pay tribute to the Ottoman Empire, the leading military power at the time, adopting a policy that was a compromise between resistance and conciliation. The tribute was a ransom for peace that allowed the two principalities to keep their independence from the Sublime Porte, as well as their traditions, institutions and laws. As for Transylvania, between the eleventh and thirteen century it had gradually been incorporated into the Hungarian kingdom, within which it nevertheless enjoyed a certain degree of autonomy as the only province recognized as a voivodeship. In the sixteenth century, though, at the height of the Ottoman power under the Sultan Suleiman the Magnificent (1520–66), at times the three Romanian lands found themselves in danger of coming under direct Turkish rule. In 1526 the Turks seized Budapest after the Battle of Mohács and then, in 1541, they wiped the Kingdom of Hungary off the map of Europe by turning it into a pashalik. In the same year, Ottoman suzerainty was also extended over Transylvania, which from this moment on became, like Moldavia and Wallachia, an autonomous tributary principality and was to remain so for the next 150 years. From the second half of the sixteenth century, Ottoman suzerainty became progressively more oppressive and duties towards the Sultan increased.⁴²

41 On the Slavonic features of Romanian, see Petrucci (1999).

42 On the 'oriental' (principally Turkish) influence on Romanian, see Șăineanu (1900), Wendt (1960), Rosetti and Cazacu (1961:333–49), Rosetti *et al.* (1971:408–22) and Rosetti (1973:128–31).

With the conflicts between the Habsburgs and the Ottoman Empire and then the disintegration of Catholic Hungary, the ideas of the Reformation had begun to penetrate into Transylvania, especially among the local Saxon, Hungarian and Szeckler populations,⁴³ but less so among the Orthodox Romanians. The propagation of the Reformation encouraged the translation and printing of religious books in Romanian, and acted as a first step towards the use of the vernacular in ecclesiastical, cultural and political life (Densușianu 1938:4–7; Niculescu 1981:89–90). After the introduction of the printing press, the Romanian principalities had been among the first countries in eastern central Europe to employ the new medium. The beginning of a real printing activity in Romanian is connected with the name of Deacon Coresi, a skilled printer from Târgoviște, in the north of Wallachia. During the second half of the sixteenth century, he embarked upon printing ten Romanian or bilingual Romanian–Slavonic church books (see Rosetti 1946:94) in the town of Brașov (in the south of Transylvania), which became one of the most important centres of Romanian culture.⁴⁴ These were all important steps in the gradual evolution of Romanian into a cultivated language and in laying the basis of the literary language.⁴⁵ Nevertheless, most works in Romanian were still written using the Cyrillic alphabet and only the Calvinists attempted to introduce Roman letters (Tagliavini 1972:546f.).

3. The seventeenth century and the strengthening of the literary and national languages

In Italy, the beginning of the seventeenth century saw the publication of the *Vocabolario degli Accademici della Crusca* (1612; with a further two seventeenth-century editions in 1623 and 1691). It was a monumental lexicographical work and one that was to act as a prestigious model for subsequent dictionaries of European languages, notably that of the *Académie française* (1694, then 1718, 1740, 1762, 1798).⁴⁶ The idea for a dictionary of this kind had originally come

43 To consolidate its rule in Transylvania and weaken the Romanians, who constituted the majority of the population, the Hungarian crown had encouraged Magyars, as well as Szecklers and Saxons, to settle in the region, granting them land and special privileges.

44 Deacon Coresi's earliest printing in Romanian, the Lutheran *Catechism* (*Întrebarea creștinească*), is dated 1559. Note also his translation of the four Gospels, the *Tetraevangelul* of 1560–61. For an outline of sixteenth-century Romanian texts, see Niculescu (1981:92–95).

45 On the literary language between 1532 and 1640, see Gheție (1997a).

46 For an outline of Italian lexicography, see Della Valle (1993).

from the man of letters Lionardo Salviati, who in the early 1580s had given a more philological imprint to the *Accademia della Crusca*, and it was taken up after his death by other members of the *Accademia*.⁴⁷ A self-financed project, the resulting dictionary was the epitome of the primacy of literature in which all fourteenth-century authors, even very minor ones, were deemed of equal worth.⁴⁸ Since it was a puristic dictionary of the literary language with only a remote link with ordinary, everyday speech, criticisms and polemics against what was perceived as the tyranny of the *Crusca* and of fourteenth-century Tuscan soon ensued. A strong disagreement was voiced, for instance, by Paolo Beni in his *Anticrusca* of 1612 and, later in the century, by Daniello Bartoli's *Il torto e il diritto del non si può* (1655, expanded in 1668). The publication of the *Dizionario* overshadowed the production of grammars in the first part of the century so that it can be said that the 'self-confident age of humanist discovery and that of the mature and serene equilibrium of the Renaissance were followed by one of stagnation' (Migliorini and Griffith 1984:259). Seventeenth-century grammars could hardly compete in richness and originality with what the previous century had offered in such abundance. Around the middle of the century, though, new vigour was gained with the *Osservazioni della lingua italiana* by the Jesuit Marcantonio Mambelli (known as Il Cinonio), published in 1644 (volume II) and in 1685 (volume I, posthumously),⁴⁹ and with Benedetto Buommattei's philosophical grammar *Della lingua toscana* (1643), which anticipated the new direction adopted by the theorists of Port Royal.

In scientific writing, Latin still offered an obvious advantage over the vernacular inasmuch as it was an international medium. Besides, when specialized knowledge was involved, certain elite groups were very reluctant to give up their professional cryptolalia (Hall 1974:154). But new scientific approaches and new discoveries with their practical applications also had linguistic implications, not only in terms of the need to create new words (and in this respect Greek and Latin were to continue to play a crucial role), but also in terms of which language to employ to make the newly acquired knowledge more widely known. Galileo Galilei's case is emblematic of such developments. After having established his position in the world of scholarship with the publication of the *Sidereus nuncius*, in Latin, in 1610, Galileo (who was a Tuscan from Pisa) then wrote all his major works in the vernacular,

47 The *Accademia* was founded in 1582 and Salviati was admitted in 1583. He died in 1589.

48 On the different steps that led to its publication, see Parodi (1974).

49 The two parts were published together for the first time only in 1709–11, in two volumes.

despite strong protests from foreign fellow-scholars. Wanting to break with the Aristotelian and university tradition, he opted for Florentine, convinced as he was of its richness and perfection, and offered in his work a clear and elegantly flowing example of Italian prose, which allowed him to widen his readership to encompass those who had a poor knowledge of Latin or none at all.⁵⁰

In the Italian tradition, many dialects had a long-standing rich literary tradition going back to the Middle Ages. If before the sixteenth century this production cannot properly be designated as 'dialect literature', given that it was not subordinated to a superior standard, with the establishment of a literary language generally accepted across Italy came the development of dialect literature proper, that is, literature deliberately written in dialect in order to contrast with the literary language and its academicism. This lively phenomenon, described by Benedetto Croce as '*letteratura dialettale riflessa*' (Croce 1927; see on this also Jones 1990; Paccagnella 1994:495–98), was 'an insurrection, revenge against the tyranny of the literary language and its literature', even though its real aim was to complement rather than replace literature written in Italian (Croce 1927:223–27). Hence it comprised self-conscious, sophisticated works of considerable merit which are quite distinct from 'popular literature'. The genres attempted in dialect were those that did not aspire to the validation that the polished learned language could give, such as theatre, satires, love and mock-heroic poems (see Cortelazzo 1980:73–91). It is mostly in these fields that dialect authors expressed in full their artistic and linguistic vein, with contributions coming from across the peninsula: Alessandro Tassoni (1556–1635) from Modena, Adriano Banchieri (1568–1634) from Bologna, Giulio Cesare Cortese (c. 1570 to c. 1640) and Giambattista Basile (c. 1572–1632) from Naples,⁵¹ Giuseppe Berneri (1634–1701) from Rome, Carlo Maria Maggi (1630–1699) and later Carlo Goldoni (1707–1793) from Venice, Carlo Porta (1775–1821) from Milan, Giovanni Meli (1740–1815) from Palermo.

In France, the ideas on language use of François de Malherbe (1555–1628), who from 1605 onwards was official poet at the court of Henry IV and later Louis XIII, were in tune with those of the monarchy and were to permeate the works of the normative grammarians of the seventeenth century. Malherbe

50 On Galileo's language, see Spongano (1949), Altieri Biagi (1965) and Marazzini (1994:287–92).

51 The dates of birth of both Cortese and Basile and the year of Basile's death cannot be established with certitude. I am following here the indications provided in Malato (1997:820–22, 844–46). See on this also *DBI*: s.v. and *LIE* 1990: s.v.

was not a systematic grammarian and his ideas can rather be deduced from his *Commentaire* on the poetry of Philippe Desportes (1546–1606) (see Brunot 1891; 1966–79: III (1–2), 1–14), in which he headed the reaction against what he saw as the linguistic excesses, untidiness and confusion of sixteenth-century poets. Seriously engaged in the creation of a more efficient linguistic system and exclusively concerned with the distinctive character of court poetry, Malherbe's purism, essentially restrictive in nature, insisted upon the need for strict form, restraint, clarity of expression and purity of diction, paving the way for French Classicism. Archaisms, dialect words, technical terms, diminutives, neologisms, borrowings, alongside any potential ambiguity, were condemned; clarity and precision (*clarté* and *précision*), as well as ready intelligibility were strongly promoted. Malherbe's search for a more accessible style did not mean, though, that he wanted the varieties of French used by the labouring classes to enter court poetry, because, on the contrary 'he was primarily concerned with the production of a literary style which would distinguish the refined minority from the rest' (Lodge 1993:174). His saying – as his disciple Racan recalled – that whatever was written had to be comprehensible to the 'hay-pitchers in the Haymarket' ('crocheteurs du Port au Foin') did not imply that he valued their coarse talk, but rather that poetry had to be intelligible to all (Hall 1974:174). The truth is that in France, just as in Italy, the spoken usage of the illiterate masses remained largely unaffected by high-level language codification.

In 1635 the founding of the *Académie française*, sponsored by Richelieu and modelled on the Florentine Crusca, was calculated to enhance the status of the French tongue nationally and abroad and was driven by a purist movement that wanted to define and fix the literary language. A crucial contribution in this sense came also with the *Remarques* [...] *sur la langue françoise* (1647) of Claude Favre de Vaugelas (1585–1650).⁵² Like Malherbe, Vaugelas – who considered himself a language observer rather than a law-giver – did not compose a formal grammar, and his work is rather a collection of scattered linguistic observations. At the core of his views on what constituted correct language was usage. In particular, the model for his language was based, he explained, not on the usage of the people at large, but rather on the most sensible section of the court ('la plus saine partie de la Cour'), as well as, when usage was uncertain, on the best authors of the time. His point of view was puristic and clearly aristocratic, rejecting with disdain the low talk of the

52 On the *Remarques*, see Ayres-Bennett (1987) and, in particular, pp. 191–200 for their socio-cultural background. See also Brunot (1966–79: III (1–2), 46–56).

populace ('la lie du peuple'). Vaugelas conceded, though, that, in some instances, some comparatively less-educated people may have a better linguistic instinct than the learned. This was the case with women, who were identified by Vaugelas as the repositories of a purer language and models of good linguistic usage to be consulted as authentic sources of untainted opinions on the French tongue. In polite society – and the *Remarques* were discussed and debated in the *salons* – women's lack of excessive formal education in the ancient languages was seen as a guarantee against pedantry and artificiality, neither of which conformed with the taste of the age.⁵³ These considerations were also expressed in the rich conduct literature of the time and Nicolas Faret, for instance, in his *Honneste homme* (1630), encouraged his readers to seek women's company and conversation, in itself a form of social education (Ayres-Bennett 1994b:45). Paradoxically, other more traditional contemporary attitudes on women and the question of *bon usage* were decidedly more negative and perceived women's language as being incorrect and weaker than men's, with poor pronunciation, poor spelling and grammatical errors.⁵⁴

Perhaps also on account of the more positive views on women and their use of language, we find at this time some of the first instances of the production of grammars, and linguistic reflection, by women in the Romance area. Marie Le Jars de Gournay, editor of Montaigne's *Essais*, and his 'fille d'alliance' (adopted daughter), 'one of the most intelligent of the early seventeenth-century French thinkers on language' (Hall 1974:175f.), became involved in the discussions on poetics of the first half of the century, objecting to Malherbe's ideas and expressing instead her admiration for Ronsard and the *Pléiade* and, clearly, for Montaigne. Marguerite Buffet's *Nouvelles observations* of 1668 was composed specifically to teach ladies the art of proper speaking and good writing, offering brief, familiar and simple rules very much dependent on Vaugelas's style of presentation (Ayres-Bennett 1987:205; 1994b:39–41).

The continuity of seventeenth-century French production of grammars with the past was ensured by the ever-growing publication of grammars of a practical

53 But on the *Précieuses* and their use of language (mocked by Molière in his *Les Précieuses ridicules*), see Lathuillière (1966), Brunot (1966–79: III (1), 66–74) and Timmermans (1993:104–23).

54 Ayres-Bennett (2004:111–80) deals specifically with these issues and views, also offering a detailed analysis of general and specific features of women's language as described by seventeenth-century authors of metalinguistic texts (see in particular pp. 126–76). On women arbiters of linguistic usage, see also Timmermans (1993:135–51).

nature (and often contrastive with other Romance and non-Romance languages), which went hand in hand with the cultural expansion of French abroad. Among the innumerable French grammars of this kind, whose success continued well into the following century, were Charles Oudin's *Grammaire française* (1632) and *Grammaire royale française & allemande* by J.-R. des Pepliers (1689). At the same time, a strong element of innovation in the field of the production of grammars came from Cartesian philosophy and method, and their application to language, with the publication of the *Grammaire générale et raisonnée*, more commonly known as the *Grammaire de Port-Royal* (1660), the work of Claude Lancelot and Antoine Arnauld.⁵⁵ Frequently reprinted and re-edited, and repeatedly translated and adapted into several languages, it began a new and influential direction in grammar at an international level, by combining the 'high' tradition of linguistic thought concerned with general theoretical claims with the 'low' tradition of materials, procedures and criteria for language teaching. A model for a whole series of similar grammars, it was used in half the schools in Europe at least until the first decades of the nineteenth century, in its original version, as well as in a variety of adaptations (Simone 1998:166).

In Romania, between the second half of the seventeenth century and the first half of the eighteenth, Romanian supplanted Slavonic as the language of the Church and of the princely chancelleries. The seventeenth century, in particular, was a golden age for Romanian literature. Chroniclers like Grigore Ureche (c. 1590–1647) and Miron Costin (1633–91) in Moldavia and Constantine Cantacuzino (1640–1716) in Wallachia promoted Romanian as a language of religious, as well as lay literary works.⁵⁶ The language of the common people was no longer deemed inferior to Slavonic: on the contrary, these authors celebrated, in Romanian, the Latinity of the Romanian language linked to the consciousness of the common Roman origins of all Romanians, thus raising the issue of their fundamental unity, whether Moldavians, Wallachians or Transylvanians. Of great significance is the translation of the Bible into Romanian thanks to the extensive collaboration of some famous scholars of the time: the *Biblia de la București* or *Bible of Prince Șerban Cantacuzino*, as it has come to be known, was published in Wallachia in 1688 (see Rosetti *et al.* 1971:167–83; Rosetti 1973:118).⁵⁷ The seventeenth century also saw the publication of the first legal code in Romanian in Iași in 1646 – the *Carte românească de*

55 On Port Royal, see Padley (1976:210–59) and Simone (1998:165–70) and, for a more detailed discussion, Donzé (1967), Pariente (1985) and Tsiapera and Wheeler (1993).

56 For an outline of the literary language between 1640 and 1780, see Gheție (1997b).

57 The New Testament had been translated into Romanian (*Noul Testament*) and printed in Transylvania in 1648 under the direction of Metropolitan Simion Ștefan.

învățătură also known as *Pravila lui Vasile Lupu* – as well as of Metropolitan Varlaam's well-known *Cazania* (1643), one thousand pages containing sermons, fragments of the Gospels and tales of the saints' lives, as well as the first philosophical writing in Romanian culture, the *Divanul sau Gilceava Înteleptului cu lumea sau Giudețul sufletului cu trupul* ('The Divan or the Sage's Quarrel with the World or the Dispute of the Soul with the Body', 1698) of the renowned and cultured Moldavian prince Dimitrie Cantemir (1678–1723). Cantemir, a historian, philosopher and writer, was also the author of the first novel in Romanian literature, *Istoria Ieroglifică* ('The Hieroglyphic History')⁵⁸ (1703–05) – an allegorical novel about the struggle for the throne between factions of boyars in the Romanian lands – and of valuable historical works: among these, in Romanian, he wrote between 1719 and 1722 the *Hronicul vechimei a romano-moldo-vlahilor* (*The Chronicle of the Ancientness of the Moldavian-Wallachian Romans*; first composed in Latin), in which he demonstrated the Roman origins of the Romanian people and their continuity in the territory of former Dacia. Numerous popular works translated or adapted usually from Slavonic or Greek had a large circulation in all Romanian provinces, thus contributing too to the diffusion of the literary language: they included hagiographic, apocalyptic and apocryphal narrations, as well as didactic literature and romances. The *Alexandria* (1713), the story of Alexander the Great, and the *Floarea darurilor* (1700), the flower of grace (the medieval *Fior di virtù*, an account of virtues and vices), enjoyed particular success as indicated by the number of extant manuscripts and printed editions (Rosetti 1973:120).

Throughout the sixteenth century and until the first half of the seventeenth, Spanish continued to enjoy widespread recognition and prestige across Europe, thanks to a degree of maturity and flexibility that expressed itself in the admirable production of the *Siglo de Oro*.⁵⁹ The Golden Age of Spanish culture, backed by the strongest political power existing on the continent, stretched across almost two centuries. With authors such as Garcilaso de la Vega (1501?–36), Cervantes (1547–1616), Góngora (1561–1627), Quevedo (1580–1645), Tirso de Molina (1583–1648) and Calderón de la Barca (1600–81), Spanish

58 In his writings, Cantemir came face to face with the necessity of enriching Romanian vocabulary through the introduction of borrowings from other languages and neologisms, and drew up for his novel an explanatory list of foreign terms (Rosetti *et al.* 1971:374–407; Niculescu 1981:104). On Cantemir's lexicon in the *Istoria* and the *Divanul*, see Miron (1978).

59 On Spanish in the *Siglo de Oro*, see Lapesa (1980:291–366) and, specifically on its linguistic features, pp. 367–417.

literature excelled in all genres. After 1665, though, a series of French victories started to show the cracks in the Spanish Empire and the decrepitude that had taken hold of it. The Catalan revolt of June 1640 and the Spanish counter-insurgency exhausted Spain and ultimately contributed to its decline as a world power. In the same year, from December 1640, the Portuguese revolt began a war of twenty-eight years that eventually saw Portugal regaining its independence despite every effort by Spain to reconquer it. The heyday of the Spanish Empire was over, while France, on the contrary, was becoming ever more powerful.

4. The eighteenth century and the universality of the French language

From the last decades of the seventeenth century, concomitantly with the first years of the reign of Louis XIV ('le Roi Soleil') and the publication of the Port Royal grammar, French began its ascent to the status of internationally recognized language of prestige. Its production of grammars reached its apogee in the century of Enlightenment with the works of Du Marsais, Girard, Beauzée and the *Idéologues* (see Joly and Stéfanini 1977; Désirat *et al.* 1982).⁶⁰ These grammarian-philosophers developed a science of grammar concerned with the explication of linguistic structure by making use of universal principles, thus creating a type of grammar distinct from the practical manuals for language teaching. In between these two types of grammatical approach there were texts such as the *Traité de la grammaire françoise* (1705) of François-Séraphin Régnier-Desmarais and the *Principes généraux et raisonnés de la grammaire françoise* (1730) of Pierre Restaut, in which traditional attention to language description and pedagogical needs are integrated with theoretical reflection. French grew in prestige to embody everything that was noble, polished and reasonable in human speech. Antoine de Rivarol peremptorily stated in his influential *Discours sur l'universalité de la langue française* (1784) that French was superior to all other languages and synthesized his views by claiming that whatever was not clear was not French ('ce qui n'est pas clair n'est pas français') (Hall 1974:177).⁶¹

An interesting, and less known, by-product of the philosophical and universal-principle based approach to grammar, coupled with the practical

60 On French in the eighteenth century, see Brunot (1966–79: VI (1–2)) and Seguin (1972).

61 On Rivarol, see also Brunot (1966–79: VIII (2–3), 848–64).

needs of studying a language, is the so-called ‘grammaires pour les dames’. Between the end of the seventeenth century and the first decades of the nineteenth, a number of grammars, claiming to target also (or more specifically) ladies, were published across Europe.⁶² In France the *Grammaires pour les dames* included not only grammars of foreign languages, but also of the national idiom. We find works such as Prunay’s *Grammaire des dames* (1777), Nicolas Adam’s *Grammaire française universelle à l’usage des dames* (1779), Antoine Tournon’s *Les Promenades de Clarisse et du marquis de Valzé, ou Nouvelle méthode [...] à l’usage des dames* (1784), or abbé Barthélemy’s *Grammaire des dames* (1785). We often find in these works a marked insistence on the importance of gaining a proper knowledge of one’s mother tongue and its fundamental principles (which can then be applied also to the acquisition of foreign and classical languages) as well as a specific focus on the question of correct orthography (see on this Reuillon-Blanquet 1994; 1995). Authors of these works often claimed that, in order to meet women’s specific demands and needs, they adopted special strategies to make them easy and accessible, the underlying implication being that if these grammars were suitable for women, then they were suitable for anyone, and specifically for those with no knowledge of Latin. Ultimately, grammars ‘pour les Dames’ became synonymous with elementary manuals for people with more limited education or intellectual means, determined by gender, lack of study or age.

With the transition from the *Siglo de Oro* to the age of Enlightenment, Spain lost to France the political and cultural hegemony it had enjoyed at an international level until then. But the foundation of the *Real Academia Española* in 1713 accelerated the process of the stabilization of the language and usage prescriptions that are at the basis of modern Spanish. The publication of the *Diccionario de la lengua castellana*, also known as the *Diccionario de Autoridades* (Madrid, 1726–39, 6 vols.), drew its inspiration from some foreign lexicographical works, such as the *Vocabolario della Crusca* and the *Dictionnaire de l’Académie Française*, as well the dictionaries by Richelet, Furetière, Trévoux, and Danet and some Spanish ones, among them the *Tesoro de la lengua castellana, o española* by Sebastián de Covarrubias (Madrid 1611). With their dictionary the academicians aimed to offer a wide and modern lexical inventory that would re-establish the prestige of Spanish, much damaged by the political decadence of the country (see Gili Gaya 1963; Lázaro Carreter

62 On grammars ‘for the ladies’ as a wider European phenomenon, see Ayres-Bennett (1994a) and Finotti and Minerva (2012). On the main features of this production, see Minerva (2000). Specifically on Italy, see Sanson (2011:209–32).

1972; López Morales 2005). Another goal of the *Diccionario* was to withstand and counteract the growing numbers of Gallicisms entering the language and, it was felt, corrupting it.⁶³ Grammar production was dominated by the publication of the *Gramática de la lengua castellana* (Madrid, 1771, with further editions in 1772, 1781 and 1796)⁶⁴ by the *Real Academia*, which was officially imposed by Charles III as the grammar to be used in all the kingdom's schools from 1780 (Lázaro Carreter 1985:189). Following a moderately normative perspective, the *Gramática* offered a complete and structured description of Spanish usage, which was, rather than the written and literary language, the overriding principle at the core of the *Real Academia*'s linguistic positions (Hernández 1992:362). The *Diccionario* and the *Gramática* were complemented by the *Orthographia española*, which the academicians had published thirty years earlier, in 1741.⁶⁵ Clearly, women too were not exempt from having to learn adequate and correct Spanish and in 1790 Josefa Ama y Borbón, in the female education plan she drew up in her *Discurso sobre la educación física y moral de las mujeres*, stressed the importance for well-to-do women who were to marry educated men of a good knowledge of the correct grammar and pronunciation of Castilian, which had to be supported by the reading of the great Spanish classics (Kitts 1995:204).

The early eighteenth century brought some unpromising signs for the cultural life that was left in the Catalan-speaking territories. Politically speaking, the reprisals taken against Catalonia after the War of the Spanish Succession, which ended in 1714, threatened to abolish its national identity altogether. Since 1704 the House of Bourbon had reinforced the organization of the state, adopting a French-style centralizing policy with inevitable linguistic repercussions. With the *Decreto de Nueva Planta* (1716), the suppression of the autonomous government and of the existing Catalan universities was imposed, together with limitations on the use of the Catalan language. From then on Castilian became the only official language, and in 1768 more severe restrictions came into effect when the teaching of Catalan in schools was prohibited and Castilian was imposed in the courts by the *Real Cédula del Conde de Aranda* (Hernández González 1992:411; Terry 2003:52). It is from this

63 Specifically on the influence of French over Spanish, see Brunot (1966–79: VIII (1), 39–84), and in particular pp. 49–68 for the eighteenth century.

64 On eighteenth-century Spanish production of grammars, see Niederehe 2001 (in particular pp. 187–92 for the *Gramática* of 1771).

65 In the seven eighteenth-century editions of the *Orthographia* innovations in spelling were introduced, but it is the eighth edition of 1815 that is particularly important for the modernization of Spanish orthography (see Lapesa 1980:421–24; Schmid 1992: 423–25).

time that a period of persecution of the Catalan language began, spurred by a conscious plan to annihilate it (Ferrer i Gironès 1986:47–60; Joan *et al.* 1994:160). Nevertheless, despite the very real obstacles and opposition that the eighteenth century brought for the Catalan language, interest in regional language issues did not cease (Comas 1972:170–86). In 1743 we find the first grammar of Catalan. It is the work of Joseph Ullastra, the *Grammatica cathalana embellida ab dos orthographias*, which remained in manuscript (printed only in 1980) just like Antoni Febrer i Cardona's *Principis generáls de la llengua menorquina* (1804) and *Principis generáls y particulars de la llengua menorquina* (1821), both practical works (on the Menorcan variety of Catalan) inspired by Restaut's and Wailly's eighteenth-century French grammars (Solà 1991:262). The first printed grammar of Catalan appeared only at the beginning of the nineteenth century in Barcelona, that is, the *Gramática y apologia de la lengua cathalana* ([1813–15]) of Josep Pau Ballot i Torres. He clearly stated that his linguistic model was the language of the seventeenth century, when he believed Catalan had reached perfection, thus rejecting numerous contemporary spoken forms (Solà 1991:265; see also Segarra 1987).⁶⁶ After 1750, the growing concern with Catalan tradition was also backed by the founding of new institutions such as the *Real Academia de Buenas Letras* (1752) and the *Academia de Ciencias* (1764). Even though most activities in these academies were in Castilian, Catalan or an interest for Catalan, its literature and the history of the region was often present (Joan *et al.* 1994:155). The *Real Junta Particular de Comercio* created by Charles III in Barcelona (1758) compensated for the absence of the traditional university, which had been moved to Cervera by Philip V, and promoted the creation of technical schools that, in turn, favoured the economic development of Catalonia. And with the opening up of trade with the South American colonies in 1778, a new sense of confidence in the commercial society of Barcelona developed.

Interest in the Portuguese language found its expression in the eighteenth century in, among other things, the production of works on orthography, such as the *Orthografia ou Arte de escrever e pronunciar com acerto a Lingua Portuguesa* of Father João de Moraes Madureira Feijó (1734), the *Orthographia da Lingua Portuguesa* (1736) of Luís Caetano de Lima, the *Orthographia Portuguesa, ou regras pera escrever certo* (1783) of Francisco Félix Carneiro Souto-Major and the creation of the *Academia Orthográfica Portuguesa* by João Pinheiro Freire da Cunha in 1772

66 For an outline of Catalan production of grammars through the centuries, see Solà (1991) and Swiggers (2001, esp. pp. 495–97). On its lexicographical production, see Rico and Solà (1991).

(Gonçalves 2003:50–54). Nevertheless, it was not until the beginning of the twentieth century that a more definitive standardization of Portuguese spelling took place, when the rich nineteenth-century production eventually culminated in the *Formulário da ortografia nacional* of Aniceto dos Reis Gonçalves Viana in 1911. Latin, although starting to lose ground, was another written medium in use among learned members of the clergy as well as lay people. Contemporary local newspapers testify to its use, for instance, with the advertisements in which we read of teachers offering Latin lessons, or of the opening of new schools or the recent publication of easy grammatical methods (Mendes Drumond Braga 2001:473–75). The third spoken and written tongue in Portugal was Castilian, but, as in the rest of Europe, it had to withstand the competition of French. Even for French, newspapers, such as the *Gazeta de Lisboa*, were full of advertisements by teachers offering their services to an elite public (Mendes Drumond Braga 2001:476–78). Italian, too, was of interest for a certain type of audience and indeed there was still considerable knowledge of Italian among educated people abroad, since it continued to enjoy a role of prestige in elegant conversation, literature (for example with the great success of Metastasio) and music (see Folena 1983). But appreciation of Italian too was eventually ousted by the unstoppable influence of French.⁶⁷

In Italy, the ‘language question’ was far from settled and debates were as alive as ever.⁶⁸ There were two main points of controversy. The first was whether written Italian should continue to be based on fourteenth-century Tuscan, as codified by the *Crusca* in its *Vocabolario*, or whether the principle of imitation should be relinquished because it was out of tune with contemporary taste. The second concerned the attitude to be adopted towards Gallicisms, which were increasingly present in both spoken and written Italian. From the last decades of the seventeenth century, French progressively became the necessary tool to be up-to-date with political, scientific and literary novelties, and to access, by means of French translations, works originally in other languages, such as German and English. It soon became an integral part of the upbringing of the young men and women of the smart set (Dardi 1984; 1992; Morgana 1994).⁶⁹ The Enlightenment and its philosophy,

67 On the influence of French on Portuguese in the eighteenth century, see Brunot (1966–79: VIII (1), 29–38).

68 On linguistic debates in eighteenth-century Italy, see Vitale (1984: 94–156), Puppo (1966; 1975), Migliorini and Griffith (1984: 297–347) and Matarrese (1993).

69 On lexical borrowings between Italian and French, see Hope (1971). For a bibliography of texts used to teach French in Italy between 1625 and 1860, see Minerva and Pellandra (1997). On the influence of French across the different states in the peninsula, see Brunot (1966–79: VIII (1), 85–137).

as well as rationalist philosophers, the sensationalists and the encyclopaedists, not to mention France's political and military power and its distinction and glamour in a variety of fields (e.g., science and technology, literature, fashion, cuisine, transport and etiquette), were all important elements that contributed to the prestige and success of the French language throughout Europe in this period. Arguably, in the eighteenth century, Italian was an uncertain tool of expression and French was often the preferred choice of intellectuals and scientists for their works or their memoirs (for instance, the Neapolitan Ferdinando Galiani used French for his *Dialogues sur le commerce du bleds* of 1770 and so did the Venetian Goldoni for his *Mémoires* of 1787),⁷⁰ as well as for personal correspondence: this was the case with Antonio Conti's letters to Vallisnieri and Scipione Maffei, Lazzaro Spallanzani's to his cousin Laura Bassi, or Vittorio Alfieri's to his sister and brother-in-law (see Bédarida and Hazard 1934:29–32; Dardi 1984:362f.; Migliorini and Griffith 1984:317). The 'Enlightened' and revolutionary Giovanni Ristori in 1788 alleged that 150,000 learned people in Italy had at least a reading knowledge of French.⁷¹

French also played an important role in elegant mundane conversation in north Italian urban centres. In Venice, for instance, speaking the language was very fashionable, a sign of distinction for the upper classes, but, on a more neutral level, using French in everyday conversation stemmed from the need to cover a gap in the Venetian linguistic repertoire, whenever the speaker found the local dialect inappropriate, in either a specific instance or for a specific topic, and Italian itself inadequate because it was too literary and stiff (Cortelazzo and Paccagnella 1992:259). In Piedmont, because of its proximity to France and the bilingual structure of the state, the influence of French was especially strong. The renowned author Vittorio Alfieri reported that the upper classes in Turin used French and Piedmontese almost exclusively (Marazzini 1992:4), indicating, therefore, that knowledge of French was in this area much more than a simple aristocratic whim and rather a matter of everyday communication. Opposition to what was perceived as a real linguistic invasion was voiced by several men of letters throughout the century and, in particular, by the Veronese Father Antonio Cesari (1760–1828) and later by the Neapolitan Basilio Puoti (1782–1847), the most representative exponent of purism between the two centuries (see Vitale 1986).⁷² Women especially were

70 But on the possibility that Goldoni actually wrote his memoirs in Italian first, before rendering them into French, see Wallington (2008).

71 Cited in Dardi (1984:349f.).

72 Polemical movements against the overarching influence of French and its culture developed also in Spain and Portugal, often provoking a purist reaction to counteract

deemed to be particularly keen to adopt and follow this Francophile linguistic trend at the expense of their knowledge of Italian, which, it was felt by men of letters and pedagogues alike, was unacceptably poor (see Sanson 2011:143–72). French had become the culprit for what was perceived to be a linguistic and moral decadence of the peninsula that concerned all speakers alike, but women to an even higher degree. Some scholars adopted a more lenient and pragmatic attitude towards French: the Paduan Melchiorre Cesarotti, for instance, the most authoritative Italian linguist of the time, observed in his *Saggio sopra la lingua italiana* (1785) that the French language had become extremely common throughout Italy and, although he disapproved of those who used French terms needlessly and out of context, he also realistically accepted the positive contribution that French had to offer in enriching and modernizing Italian (Cesarotti 1785:114–15).

As for the medium of communication of the majority of the population, the Piedmontese Giuseppe Baretta in 1768, touching upon the linguistic variety of the peninsula in his *An Account of the Manners and Customs of Italy* (written as a defence of his country against some unfortunate observations made by Dr Samuel Sharp in his *Letters from Italy* of 1766), explained that the dialects were preserved ‘in what may be called their barbarous purity’ and that in their daily intercourse all Italians used the speech of their own restricted area, using Tuscan only if they had to converse with strangers (Baretta 1768, II:183). Using Tuscan outside Tuscany was indeed considered a clear sign of affectation, with the risk of being ridiculed and mocked (Baretta 1768, II:183f.).⁷³

It is in the eighteenth century that the first grammars of some Italian dialects were published, overcoming the grammarians’ prejudice according to which only the grammar of a ‘language’ could be described, but not that of a ‘dialect’ (Benincà 1988:45). In Naples, for instance, the economist and man of letters Ferdinando Galiani published (anonymously) in 1779 his *Del dialetto napoletano*, with the aim of describing the grammatical foundations of an ‘illustrious’ Neapolitan that could aspire to be the national language of the new kingdom of Naples (Benincà 1994:567).⁷⁴ Long-existing ambitions to codify and standardize the regional dialect of Piedmont saw a high point with the publication of the *Grammatica piemontese* (1783) of the physician

the perceived excessive borrowing and diffusion of Gallicisms. See on this Martinell (1984) for Spain; Boisvert (1983–85), Machado (1984) and Teyssier (1994:468–70) for Portugal. In general on inter-Romance borrowings, see Hall (1974:159–63).

⁷³ On Italian and foreign travellers’ remarks on the use of Italian and the dialects in eighteenth- and nineteenth-century Italy, see Cartago (1990) and Serianni (1999).

⁷⁴ Galiani’s was the first Neapolitan grammar to be published, but the first to be written was Francesco Oliva’s *Grammatica della lingua napoletana* (see Oliva 1989).

Maurizio Pipino, who took as his model the language spoken at the Turinese court and who hoped to promote the use of Piedmontese as a written medium. Both works present a strong focus on the issue of orthography, a first crucial element to codify the dialect. We also have the first grammatical description of Sardinian: this is Matteo Madao's *Saggio d'un'opera intitolata il Ripulimento della lingua sarda lavorato sopra la sua analogia colle due matrici lingue, la greca e la latina* (1782), a traditionally structured work in which the author intended to 'illustrate' Sardinian (specifically the Logodurese variety) by adopting a purist stance and pointing out the differences between Sardinian and Italian.⁷⁵ This work was later followed by Vincenzo Raimondo Porru's *Saggio di grammatica del dialetto sardo meridionale* (1811) and by Giovanni Spano's *Ortografia sarda nazionale ossia grammatica della lingua logudorese paragonata all'italiana* (1840, 2 vols) and Giovanni Rossi's *Elementus de gramatica de su dialettu sardu meridionali e de sa lingua italiana* (1842), composed in Campidanese. In the second half of the century several dialect dictionaries were published, amongst them the *Vocabolario bresciano e toscano* (1759), a collective work (sometimes attributed to the abbé Paolo Gagliardi), the anonymous *Raccolta di voci romane e marchiane* (1768), the *Vocabolario veneziano e padovano* by Gaspare Patriarchi (1775, with a second edition in 1796), the *Vocabolario piemontese* (1783) by Pipino, the *Vocabolario etimologico siciliano, italiano e latino* by Michele Pasqualino (1785–95) and the *Vocabolario delle parole del dialetto napoletano che più si scostano dal dialetto toscano* (1789) by Ferdinando Galiani and Francesco Mazzarella Farao (see Cortelazzo 1980:105–8).

The Enlightenment marked a crucial moment in the demise of the role of Latin as a language of culture across Europe. To start with, it lost its battle with French as the international language of diplomacy. Besides, as seen earlier, until then Latin had been the official language of schooling and of the Jesuit order, which, with its firm grip on the educational system of the different states, was a strenuous defender of the status quo. But the claims of the vernacular languages to be taught as subjects in their own right were strengthening. With the expulsion of the Jesuits from France, Spain and Italy in the second half of the century came a strong encouragement to educational plans for reform in which the national, or literary, languages were eventually given priority and took on the role of the medium of teaching. In the Bourbon-ruled Duchy of Parma and Piacenza, for instance, following Guillaume Du

⁷⁵ For a general outline of Sardinian production of grammars and lexicography, see Dettori (1988). For an outline of its external history from the Renaissance onwards, see Blasco Ferrer (1984; 1988).

Tillot's initiatives in 1768 (the year in which the Jesuits were expelled from the Duchy), a series of reforms were launched: in the *Costituzione per i nuovi regi studj*, inspired by Father Paolo Maria Paciaudi, a *Piano per le scuole gratuite dei fanciulli* was drawn up, a first attempt in Italy to create a free, state-run primary school, in which reading, writing and basic numerical skills would be taught, while strictly forbidding Latin grammar and privileging Italian instead (Lucchi 1985:36f.).

In Romania, political unification was still distant. Starting from 1711 in Moldavia and 1716 in Wallachia, the thrones of the two principalities were in the hands of the so-called Phanariot rulers (Florescu 1999:173–85). They were members of prominent Greek families (from the Greek district of Phanar in Constantinople) who acquired wealth and power by occupying high political and administrative posts in the empire of the Sublime Porte. Even though one of the main objectives of the Phanariots' presence in Romanian territory was to weaken its political elite, the regime was still a compromise that allowed for the preservation of the autonomous status of the Romanian principalities without imposing the direct administration of the Ottoman court. Inevitably, though, until 1821, when Phanariot rule was brought to an end, Greek culture had a strong influence on Moldavian and Wallachian society: Greek schools were opened, and literary, philosophical and scientific instruction in Greek was introduced, replacing Slavonic theological education. Greek became the official language of the courts and, between 1770 and 1820 especially, there was also a great influx of Greek neologisms into Romanian vocabulary (see Gáldi 1939; Rosetti *et al.* 1971:426–38; Rosetti 1973:132–34). The Phanariots were also responsible for the introduction of French culture into the Romanian lands: as polyglot officials of the Ottoman Empire, they used French as their diplomatic language and were often accompanied by French secretaries. Soon French language tutors were employed to teach the language at court, as well as to instruct the children of Romanian boyars with more cosmopolitan social and cultural aspirations.⁷⁶

In Transylvania, between the eighteenth and nineteenth centuries the Latin-oriented cultural movement of the *Școala ardeleană* (the 'Transylvanian School') espoused the ideas of the Age of Enlightenment and promoted the development of scientific knowledge and its diffusion among the masses (Rosetti 1973:138–45; Florescu 1999:166–72). Its members, who included Samuil Micu (or Clain, 1745–1806), Gheorghe Șincai (1754–1816),

76 On the influence of French on Romanian, see Heliade-Rădulescu (1898), Rosetti *et al.* (1971:577–84), Brunot (1966–79: VIII (1), 3–8) and Goldiș-Poalelungi (1973).

Petru Maior (1760–1821) and Ion Budai-Deleanu (1760/63–1820), also attempted to educate and gain political rights for the Romanian population of Transylvania who lived under Habsburg rule. In order to do so, they provided historical and philological arguments to support the thesis that the Transylvanian Romanians were the direct descendants of the Roman colonists brought to Dacia after its conquest in the second century AD, and in so doing they used the idea of the Latinity of the Romanian language and people (already expressed by the chroniclers of the seventeenth and eighteenth centuries, as seen earlier) as a political instrument to promote a nationalist spirit.

Relatively late compared to other Romance varieties, in the second half of the eighteenth century, we find the first grammars of Romanian, namely Dimitrie Eustatievici Braşoveanul's *Gramatică rumânească* (1757, in manuscript only; first printed in Ursu 1969), Samuil Micu and Gheorghe Şincai's *Elementa linguae daco-romanae sive valachiae* (1780, revised second edition in 1805) and Radu Tempea's *Gramatică românească* (1797).⁷⁷

5. The nineteenth century: between political centralization and linguistic revival

At the end of the eighteenth century and the beginning of the nineteenth, associated with revolutionary ideas came the belief that language could be a badge of cultural and ethnic identity. With the new ideology, the concept of 'nation' changed and language took on a new role in society: in a state where the citizens were bound together by a 'social contract', all members also had to be united by sharing the same language.

In France the onset of the Revolution marked a direct state intervention, to an unprecedented extent, in the field of language. The democratic ideals according to which all citizens should participate in public life, alongside the desire to improve communicative efficiency and increase the internal cohesion of the state, led revolutionary leaders to impose a policy of linguistic assimilation that left no room for regional varieties (Renzi 1981).⁷⁸ This approach has been condemned as destructive of the local dialects. The state needed one and only one idiom: from this came a strong push for the diffusion of the standard language, which *de facto* complied with the Parisian norm.

⁷⁷ On Romanian production of grammars, see Turculeţ (1989) and Swiggers (2001:479–82).

⁷⁸ On French at the time of the Revolution, see also Brunot (1966–79: IX (1), IX (2)).

Paradoxically, then, the French state since the 1789 Revolution aimed to impose 'in the name of democracy and egalitarianism [...] a standard variety which had been crystallized under the *ancien régime* as a hallmark of class distinction' (Lodge 1993:216). In 1794 the abbé Henri Grégoire presented to the National Convention a report on the linguistic situation of France, tellingly entitled *Sur la nécessité et les moyens d'anéantir les patois et d'universaliser l'usage de la langue française*, from which it emerged that out of an estimated population of twenty-five million, at least six million, mainly in the south, did not know French at all, and another six million only had a smattering of the language. Those who were able to speak it purely, either because of the region they happened to live in or because of their education, numbered only three million (Rickard 1989:121; see also Certeau *et al.* 1975). The intention of the revolutionary government was to provide a primary school, with a teacher paid by the state, in each *commune*, but the shortage of teachers and funds hampered the plan and primary state education was only established by law in 1832, when it was neither free nor compulsory (it became free only in 1881 and secular and compulsory in 1882). In 1832 it was also decided that reading was to be taught from French texts and not Latin ones, and that a standard spelling and grammar had to be adopted for all state examinations (Rickard 1989:121). In France, in the course of the nineteenth century, the use of dialect declined rapidly and linguistic unification was swift. Napoleon had devised a highly centralized governmental and administrative system, later adopted and elaborated by successive governments, which made knowledge of French a must. Conscription introduced by the Revolution, better communications, and national and regional newspapers, journals and magazines, as well as universal male suffrage in 1848 were all elements that contributed to the decline of dialects.

The strong centralism that had dominated western Europe and its cultural life in the seventeenth and eighteenth centuries, promoting the establishment and maintenance of the ideal of a single standard for any given political unit, both in terms of language and of non-linguistic issues (Hall 1974:125, 181–83), meant that, as a consequence, certain Romance varieties that had once enjoyed the status of literary standards lost prestige. With Romanticism, though, came a revived interest in prestigious literary languages of the past, and in the nineteenth century renewed impulses towards local autonomy in culture – and in some instance also in politics – came to the surface and grew in strength. It was indeed at this time that minority languages within the Romance area sought to assert their worth and seek cultural, if not political, autonomy. Two Romance languages in particular which shared a common

destiny of falling from the splendour of their ancient literary production to a period of oblivion, followed in the nineteenth century by a strong desire to revive their prestige, are Occitan and Catalan.

Old Occitan literature, especially the lyric poetry of the *trobadors*,⁷⁹ is acknowledged to be among the finest in Europe. The peaks of its expression were reached in the twelfth and thirteenth centuries. Forms of Occitan had remained in use in southern France until the end of the fifteenth century and beyond, but as a result of the edict of Villers-Cotterêts, mentioned earlier, by the end of the sixteenth century French was virtually the only vernacular written and printed in the south. Occitan remained, nevertheless, the real everyday spoken language in its home territory and French was spoken by only a tiny (yet influential) minority. With the new climate established by the Revolution, French made rapid headway as a spoken medium, above all in the cities and among the upwardly socially mobile. In the second half of the nineteenth century, in the hope of stemming the tide of French, various attempts were made in order to re-establish some form of Occitan as a literary medium:⁸⁰ Rochegude, Fabre d'Olivet and Honnorat are some of the main names linked to the 'restandardization' (Bec 1991:48) of the once prestigious *langue d'oc*. In particular it is worth mentioning the Félibrige movement (see Ripert 1918; 1948; Camproux 1953:156–200; Jouveau 1970–84), founded in 1854 by a group of seven young local poets, and generally associated with the main luminary in the study of the language, the poet Frédéric Mistral (1830–1914), winner of the Nobel Prize for Literature in 1904. Félibrige members cultivated both poetry and prose (basing their language on the contemporary speech used in the lower Rhône valley) and drew up a plan for the rehabilitation of the Occitan language and literature, comprising also a precise programme of grammatical and orthographic codification. Mistral's idyllic poem *Mirèio* ('Mireille') of 1878 was enormously popular and widely read, in the original or in translation, throughout the French-speaking world. In terms of language codification, though, the *Félibréens* obtained only limited results, hindered as they were by the conservative and folkloristic attitudes of the adherents to the movement as well as by the dialectal fragmentation of Occitan

79 On the *trobadors*, see, for instance, Jeanroy (1934), Marrou (1971) and Kay and Gaunt (1999). For a history of Occitan literature, see Camproux (1953).

80 In the nineteenth century, mainly thanks to German linguists, Occitan came to be known as Provençal, this being the name used in the medieval period for the literary *koiné*: today Provençal refers properly only to the local dialect of Provence. For an outline of the external history of Occitan from the sixteenth century to the present time, see Kremnitz (1991). On its production of grammars and lexicography, see Schlieben-Lange (1991). On the revival of Occitan, see Lafont (1974).

(Bec 1991:49–51). The masses spoke one of the Occitan dialects, read only French and did not find it easy to read archaic literary Occitan, which seemed artificial and too distant from their everyday speech (Rickard 1989:122).

In Catalonia, the same Romance variety had been in use for centuries and, as seen earlier, an interest in language and an impulse towards local autonomy had never completely disappeared, and had persisted, more or less underground (Hall 1974:125; Joan *et al.* 1994:163). From the third decade of the nineteenth century,⁸¹ the foundations were laid for the re-establishment of Catalan as a literary language as part of a movement that reached its apogee in the 1870s and was from that moment on called *Renaixença* (see Montoliu 1962; Pi de Cabanyes 1979). The revival of Catalan was certainly aided by the prosperity of its main urban centre, Barcelona, and in 1859 the interrupted tradition of the *Jocs Florals* ('Floral Games'), the poetry contest of medieval origin in which the winners were awarded jewels in the forms of flowers, was restored (see Miracle 1960): the *Jocs* had to be held exclusively in Catalan (either ancient or literary modern) – we read in the founding programme – with the aim of reviving the poetical glory of the country and reinstating and preserving the purity of the Catalan language (Martí i Castell and Moran 1986:354).⁸² However, the period of linguistic oppression that Catalan had to endure was far from over and would reach new peaks in the twentieth century, especially under Franco (Ferrer i Gironès 1986:177–201).

A similar destiny was to befall Galician in the north-west part of the Iberian Peninsula. Galician, or more precisely Galician-Portuguese (because the two regions were united at the time), enjoyed a period of flourishing literary production in the thirteenth and fourteenth centuries (with its apogee between 1250 and 1350) with the lyric poetry of the *trovadores*.⁸³ To a lesser extent, it had a further literary flourishing, in which Portuguese did not share, between 1350 and 1450, only then to have to face a long period of literary neglect after the time of the *Reyes Católicos* (Mattoso Câmara 1972:10f.; Teyssier 1980:48f.).⁸⁴ Politically subjugated to Spain, from the sixteenth century onwards, Galician was no longer cultivated as a literary language but it continued to be used as a medium of everyday communication and, especially

81 The poem *La pàtria* (1833) of Bonaventura Carles Arribau is traditionally seen as symbolically marking the beginning of the movement (Joan *et al.* 1994:168).

82 For a general outline of the codification of Catalan through the centuries, see Solà (1977), Carbonell (1979), Casanova (1991) and Rogge and Beinke (1991), and, with specific reference to the literary language, Calveras (1925).

83 On the Galician-Portuguese of the *trovadores*, see Silva Neto (1979:403–5).

84 For an external history of Galician, see Brea (1994). For the period between the thirteenth and the sixteenth century, see also Maia (1986).

after the seventeenth century, it was mostly restricted to rural areas or to the more popular and lower-class parts of the urban centres. Portuguese and Spanish theatre, from the sixteenth century onwards, reserved Galician for more rustic characters and for certain professions, such as water-vendor or charioteer (Teyssier 1980:48f.; Brea 1994:87). Then, again on a lesser scale, Galician had its own process of rediscovery and renewal of the language in the nineteenth century, especially after the 1840s, alongside claims for political autonomy. It had its own *Xogos Florais*, celebrated from 1861, and Rosalía de Castro, the poet and novelist, initiated Galicia's own *Rexurdimento* ('Renaissance') with the first work completely written in Galician in that period, the *Cantares Gallegos* (1863). The first grammars and dictionaries of the language followed soon after: the *Compendio de gramática gallego-castellana* (1864) of Francisco Mirás, the *Diccionario gallego-castellano* (1865) of Francisco Javier Rodríguez and the *Gramática gallega* (1868) of Juan Antonio Saco y Arce (Brea 1994:90). Ultimately, though, the revival of Galician was less successful than that of Catalan: the area had been for a long time a backward and impoverished part of the country (mostly reliant on agriculture and fishing) and not only had the political and religious authorities suppressed the local language, but it had also been spurned by the middle classes. Last but not least, there was considerable disagreement about the shape its standardization should take (Posner 1996:215).

As for Romansh, its production of grammars had begun at the end of the seventeenth century, with Father Basilius Meyer composing a grammar of Surselvan in 1685, followed by the publication of the *Fundamenti principali della lingua retica, o grigiona* (of the Surselvan and Surmiran varieties) by Father Flaminio da Sale in 1729. But it was after the 1820s that the Romansh production of grammars really flourished, mostly targeting a public of learned readers or German speakers (Swiggers 2001:482f.). In 1864, Josef Anton Vian's *Gröden, der Grödnert und seine Sprache*, published in Bolzano, is the first printed grammar of Ladin, whereas the first autonomous description of Friulian was published in Jacopo Pirona's *Vocabolario friulano* of 1871 (pp. xlv–lxxiii). A Friulian native speaker and a bilingual writer, Caterina Percoto, had composed in 1865 a short (still unpublished) grammar of Friulian (Caira Lumetti 1985:32; Percoto 1988:185). Her literary production and personal correspondence testify to her interest in linguistic issues and in the relationship between the national language and the dialects (see on this Sanson 2010:1034–41).

In the second half of the nineteenth century, Italian and Romanian gained official recognition. In Italy the spread of an officially recognized and

universally adopted standard was delayed until political unification took place. The gulf created in the sixteenth century between the literary language and everyday speech had widened over the ensuing centuries, and the level of command of the Italian language differentiated the social elites from the lower classes. By the time political unification was reached, in 1861, the need to have a common spoken language became urgent. Italian dialects had persisted in everyday use for centuries and, owing to the lack of a unifying political centre, they had never really been threatened by enforced centralization. In post-unification Italy debates regarding the *Questione* took on a different dimension and a clear social import (see De Mauro 1963; Migliorini and Griffith 1984:403–50; Metzeltin 1988:361–79; Serianni 1989; 1990). Now they were not and could not be limited to only literary Italian, but were also extended to the use of the national language in social contexts. Similarly, the ‘language question’ widened in scope and attracted broader interest: it was no longer merely a matter of debate for scholars and men of letters, but concerned the general public too.

In 1861, the estimated percentage of those who, out of the nation’s twenty-five million inhabitants, could have been said to know Italian reached only 2.5 percent (De Mauro 1963). This figure included the better educated as well as the inhabitants of Tuscany and Rome. More optimistically, Castellani (1982a) estimated this figure to be between 9 and 10 percent: either way it was an extremely low figure. Because Italian was above all a written language, only those with a certain level of literacy were likely to be able to acquire a full command of it. Since rates of illiteracy in 1861 were high, with an average of 78 percent and peaks between 80 and 90 percent in the south and the islands (De Mauro 1963:36f., 95), it followed that the majority of the population did not have access to the literary language and used only their native dialect. Even though a certain degree of passive knowledge of Italian was not the exclusive preserve of the literate, it would be fair to say that at the time of Unification Italy was a country of monolingual speakers, who had as their mother tongue one of the many dialects spoken in the peninsula.⁸⁵

The *Questione* debates also dealt with the best means of extending knowledge of Italian to all Italians. Structurally remote from the everyday linguistic practice of most people, the now national language was also poorly equipped for everyday discourse, having remained for centuries above the needs of daily

85 To these we must add the percentage of those who belonged to one of the many linguistic minorities that existed across the country: in addition to speakers of other Romance varieties (French, Franco-Provençal, Occitan, Ladin, Catalan), there were also those who used German dialects, Slovenian, Greek, Albanian, and Serbo-Croat. On linguistic minorities in Italy, see Telmon (1992).

life. The renowned man of letters Alessandro Manzoni (1785–1873), from Milan, intervened in the debates and, rejecting the idea that the common language of Italy ought to be based on the written tradition, he suggested instead that the national tongue of the new state had to be the ordinary language of the educated middle classes in Florence.⁸⁶ In order to make the final version of his novel *I promessi sposi* (1840–42) a practical example of his positions, he carefully and systematically reviewed the language of the entire text of the 1827 edition with the help of Florentine native speakers. Among these were the men of letters Gaetano Cioni and Giambattista Niccolini, as well as Canon Giuseppe Borghi and the mathematician Guglielmo Libri. But alongside these well-educated intellectuals, Manzoni also enlisted the help of a woman, the Florentine Emilia Luti, who was employed in his household for a period of time as governess to his daughters and later kept up an assiduous correspondence with the writer. Of modest social background and limited education in comparison with Manzoni's other language 'advisers', Emilia was able to offer him that much sought-after direct contact with contemporary Florentine on which he had based his linguistic theory (Sanson 2011: 287–99; see also Amoretti 1992).

Manzoni also explicitly set out his recommendations on language, in 1868, in his report *Dell'unità della lingua e dei mezzi di diffonderla* ('On the unity of the language and on the means of disseminating it'), commissioned by the Ministry of Education in the person of Emilio Broglio. Here Manzoni proposed, among other things, that Florentine should be taught in schools – where, in order to spread living Florentine, preference had to be given to Tuscan teachers – and that a dictionary of contemporary usage had to be produced. The *Novo vocabolario della lingua italiana secondo l'uso di Firenze* ('New Dictionary of the Italian Language According to Florentine Usage') saw the light in 1870–97 (4 vols); it is known as 'Giorgini–Broglio', from the names of the two main editors, Emilio Broglio and Giovan Battista Giorgini, Manzoni's son-in-law. It provoked a memorable reaction and critical intervention from the linguist Graziadio Isaia Ascoli who, in his 'Proemio' (dated to September 1872) to the first volume (1873) of his journal *Archivio glottologico italiano*, pointed out the impracticality and non-viability of Manzoni's proposals. Ascoli disapproved of his interventionist attitude and opposed the artificial imposition of a linguistic model, whether archaic or modern,

86 Manzoni's linguistic choice has been traditionally defined as 'il fiorentino delle persone colte' (Florentine of educated people). On Manzoni's own use of this expression, see Dardi (2008; 2009).

believing instead that the traditional literary language should act as the basis of Italian. Its evolution as the national language of all the inhabitants of the country had to be the result of a natural process, which had to imply also an increased intellectual activity on the part of the Italian people. The two main obstacles to linguistic unity were indeed, in his view, the low density of culture and an excessive preoccupation with form (Ascoli 1975:30).

In the end, it was literary Tuscan which was disseminated via the educational system: the Casati Law of 1859 (initially passed for Piedmont and Lombardy, and then from 1861 extended to the entire territory of the newly created state) and the Coppino Law of 1877 were pivotal moments in creating free elementary schooling for everyone (boys and girls alike) and in introducing the principle of compulsory attendance. But organizational problems of various kinds and the widespread use of dialect by both pupils and teachers in the classroom, especially in the countryside, deflated the initial hopes about the impact that schooling could have on the spread of the national language.⁸⁷ The process was in fact a lengthy and complex one with Italian progressively gaining ground in the everyday life of its inhabitants thanks also to other factors: military service, industrialization, bureaucracy and increased geographical and social mobility linked to the urbanization and internal migration processes, as well as the mass media, newspapers and magazines and, in the twentieth century, radio and television (De Mauro 1963).

In the nineteenth century, Italy also saw the birth and development of Italian dialectology as a scientific discipline, as a result of the trend towards studying comparative linguistics, which had originated in Germany and dominated the century (see Benincà 1994). Over the century a number of important dialect dictionaries were published, with the aim of documenting dialect words and supplying those who needed them with the equivalent words in Italian. Among the earlier ones were Francesco Cherubini's *Vocabolario milanese-italiano* (1814, second edition 1839–56), Giuseppe Boerio's *Dizionario del dialetto veneziano* (1829, second edition 1856) and Pietro Monti's *Vocabolario dei dialetti della città e diocesi di Como* (1845). But in the second half of the century, interest in the dialects grew exponentially, with scholars embarking on wider projects aimed at recording dialect examples: Attilio Zuccagni Orlandini, for instance, published in 1864 a collection of translations of a realistic sample dialogue between a master and a servant in

87 On state education in post-unification Italy, see Talamo (1960), Vigo (1971), De Fort (1996) and Serianini (1990:18–26).

different dialects. Giovanni Papanti, in 1875, with his *I parlari italiani in Certaldo*, brought together 700 versions of the first story of the first day of the *Decameron*,⁸⁸ comparing Italian dialects with Romance dialects used outside the borders of Italy as well as with the main European Romance languages. These little-studied idioms also offered the opportunity to apply in practice the comparative method, used to reconstruct proto-Indo-European languages, as a well-documented mother tongue, that is, Latin, lay to hand. With a figure of the intellectual standing of Ascoli, mentioned earlier, and his journal *Archivio glottologico italiano*, Italian dialectology took on a completely different dimension (see Benincà 1994:581–90). Interestingly, the rise of interest in the dialects flourished at a time when Italian was progressively, but unstoppably, beginning its process of becoming a mother tongue, thus leading Italy from being a newly founded state of monolingual dialect speakers to a country of bilingual speakers.

Similarly to the Italian situation, in Romania for centuries there had been no organized teaching of Romanian in schools and only in the first half of the nineteenth century did Romanian become the language of teaching and the subject of study (Rosetti 1973:136). In the early 1820s, when, with the demise of Phanariot rule the Greek hegemony in government and education came to an end, Romania entered the modern age of its history, with a period of dynamic political, social and cultural development (Treptow 1997:239–42). New literary and artistic activities ranged from the expansion of schools to the opening of theatres, to the founding of newspapers (the first Romanian-language newspaper, Bucharest's *Curierul românesc*, was published in 1829) and journals, such as *Dacia literară* (1840), *Arhiva românească* (1840) and *Magazin istoric pentru Dacia* (1845). In 1828, the man of letters Ion Heliade-Rădulescu (1802–72) published the *Gramatica românească*,⁸⁹ the most important nineteenth-century grammar of Romanian that marks a crucial moment in the battle for establishing the modern Romanian literary language. Indeed, a good part of the preface consists of proposals for simplifying the Cyrillic alphabet, arguing that this would facilitate the future development of the language – the Latin alphabet replaced the Cyrillic one only in 1863 (Florescu 1999:171) – accompanied by a spirited defence of the Romanian language, which had to be used also for literary and technical writings.

88 Papanti drew his inspiration from Salviati's small collection of dialect examples, which was published as part of his *Degli avvertimenti della lingua sopra 'l Decamerone* in 1584. There the same 'novella' is recorded in twelve Italian dialects.

89 On Heliade-Rădulescu's *Gramatică*, see Guțu Romalo (1980).

Heliade-Rădulescu was part of a group of intellectuals who belonged to radical literary and political societies in Bucharest. Aiming to popularize advanced political theories, they wanted to encourage a Romanian-based education in Wallachia and to stimulate literary, didactic and technical writing, for which a rich and flexible language was invariably needed. Members who gathered first in the *Societatea literară*, and then in the *Societatea filarmonică*, found the existing Romanian language unable to fulfil this requirement and its vocabulary too restricted for the full expression of their ideas. For this reason they devoted their efforts to translations from classical and modern Greek, as well as from French, Italian and German, but also wrote original works, including poetry, plays, travel-notes, lectures and sermons. As part of his linguistic theories, Heliade-Rădulescu also came to consider the urgent problem of the creation of a modern literary language that could be used for literary and technical works of all kinds. Eventually he came to the conclusion that literary Romanian had to conform as much as possible to literary Italian. By virtue of the many similarities between the two languages, he had indeed convinced himself that they were variants of a single language and that Italian, as the more cultivated of the two, provided a suitable norm for Romanian (see Close 1974:69–74).

The rise of a national consciousness among Romanians led to the outbreak of the revolutions of 1848–49, which advocated liberal reforms in Moldavia, Wallachia and Transylvania. Even though these revolutions failed, the movement for unification of the principalities gained momentum, and in 1859, with the *de facto* union of Moldavia and Wallachia, modern Romania was born. With the independence of the country brought about in 1877, and Romania becoming a kingdom in 1881, the country enjoyed a long period of economic and cultural development (see Treptow 1997:299–314): the ‘Age of the Great Classics’ also saw the appearance of Romania’s national poet (Mihai Eminescu, 1850–89), its classic story-teller (Ion Creangă, 1837–89), its principal dramatist (Ion Luca Caragiale, 1852–1912) and the Romanian Academy (founded as the ‘Romanian Literary Society’ in 1866, then renamed the ‘Romanian Academic Society’ in 1867 and finally the ‘Romanian Academy’ in 1879). Transylvania was annexed by Hungary in 1867 and Romanians in this territory were subjected to an aggressive policy of denationalization by the government in Budapest, which, for instance, with a series of school laws (in 1879, 1883, 1891 and 1907) imposed the teaching of Hungarian in all Romanian confessional schools (under the control of Orthodox and Greek Catholic Churches), and made its use compulsory in nursery schools. Eventually, though, a strong national movement among the majority Romanian

population prepared the ground for the union of Transylvania with the other historical Romanian lands into a single nation in December 1918.⁹⁰

So, in the world of Romance idioms, by the end of the nineteenth century, some varieties had already enjoyed stability and acceptance for centuries, some had succeeded in growing in prestige to become the official language of a state, others had to fight for their recognition after suffering neglect and oppression, and others were slowly starting to transform themselves from essentially literary to national languages. But alongside these varieties that, albeit at different rates, have benefited from codification and standardization, there is a far richer selection of Romance idioms that have survived until the present time above all in their spoken, everyday reality, despite lack of official support. Even though modern technology will help preserve some of them from total oblivion, many are waning and are destined to perish (e.g., Judaeo-Spanish). Some already have: Dalmatian, for instance, became extinct at very the end of the nineteenth century.⁹¹ It was formerly in use along the coast of Dalmatia, from Istria to Ragusa, but by the nineteenth century survived only on the island of Veglia (today Krk), off the coast of Croatia, south of Fiume. It is attested in some medieval documents and in the texts taken down by some scholars in the second half of the century (see Ive 1886 and Bartoli 1906; 2000:12–66). In 1898 the last speaker of Vegliote, Antonio Udina, called Búrbur, died in a dynamite accident (Bartoli 2000:17). His mother tongue was a variety of Venetian, but he had learnt Vegliote as a child from his grandparents and his parents, the very last native speakers of the language. Confined as it was to the everyday life of more humble speakers – fishermen, peasants and artisans – Vegliote had eventually had to give in to Venetian, deemed more prestigious and offering more hope for social promotion.

90 Romanian was one of the last Romance varieties to become the national language of a state. In 1938, Romansh was recognized as a Swiss national language following a referendum held in the same year. Article 116 of the Old Federal Constitution defined German, French, Italian and Romansh as the ‘national languages’ of Switzerland, and German, French and Italian as the ‘official languages’ of the Swiss Confederation. The amendment to this Article, voted for by the Swiss in 1996, specifies that ‘for the purpose of dealing with persons of Romansh tongue, Romansh shall also be an official language of the Confederation’ (see www.liarumantscha.ch). As for Moldovan, the official language of the Republic of Moldova (independent from the Soviet Union since 1991), it is effectively identical to Romanian, with which it shares the same literary standard (nevertheless, the question of the relation between Romanian and Moldovan is a contested political issue and is fraught with controversy). See on this C. King (2000) and Dima (2001); also Andreose and Renzi, this volume, chapter 8, and Varvaro, this volume, chapter 9.

91 For an outline of the external history of Dalmatian, see Doria (1989). On its main features, see Tagliavini (1972:374–77) and Muljačić (2000).

Geography and distribution of the Romance languages in Europe

ALVISE ANDREOSE AND LORENZO RENZI

I. Overview

Over a broad and continuous swathe of territory extending in Europe from the Atlantic (France, Spain, Portugal) to the Adriatic (the eastern coast of Italy), including some major islands of the Mediterranean (the Balearics, Corsica, Sardinia, Sicily), as well as in an isolated area in the Carpathian-Danubian area extending just beyond the Dniester, languages and dialectal varieties are spoken which continue the Latin spoken in the Roman Empire.¹ This geolinguistic area is called the *România*. The national languages of this domain, in official use in their respective countries, are: Portuguese, Spanish (Castilian), French (in France, Walloon Belgium and Swiss Romandy), Italian (in Italy and Italian-speaking parts of Switzerland), and Romanian (in Romania and the Republic of Moldova). There is also Romansh, which, despite very low, and falling, numbers of speakers, is one of the three official languages of the canton of Grisons/Graubünden, and one of the four national languages of the Swiss Confederation. Beneath this layer of official languages there is a continuum of different, more or less distinguishable, linguistic varieties: the so-called 'dialects'. Their distinctive identity, which sets them apart from the corresponding national languages and from each other, is particularly striking in Italy, where they were still in general use 150 years ago. In northern France, besides French, there are many *patois* – rustic dialects of very limited social and geographical scope, which are the heirs of what were at one time widely used regional varieties. Even in the Franco-Provençal area of the south-east, and the Occitan south, the ancient speech varieties are reduced to the status of *patois*, which are everywhere in rapid retreat. Between the two poles of the national languages on the one hand, and the dialects and *patois*, generally

¹ This chapter is the result of close collaboration between the authors. Sections 1–5 are particularly the work of Andreose, and 6–8 of Renzi. The statistical data regarding the numbers of speakers in the various linguistic areas are correct for 2007.

lacking official recognition, on the other, there is a series of varieties which, in Portugal, Spain, France and Italy, have received some degree of recognition, generally acquiring the status of 'co-official' or 'regional' languages.

Romance languages are in immediate contact with a number of other language varieties (see also Varvaro, this volume, chapter 9). There are some enclaves of non-Indo-European languages within the Romance area: Basque (*euskara*), straddling the Pyrenees between France (Pyrénées-Atlantiques) and Spain (*País Vasco* and Navarre); Hungarian, spoken in Romania (Banat, Transylvania); Turkish and Tatar, spoken in Romania and Moldova. As for Indo-European languages, the Celtic language Breton, spoken in Brittany, is the result of medieval colonization from the British Isles, and there are numerous German-speaking outcrops in Italy, due to medieval or modern colonization in the valleys or individual villages of the Italian Alps and pre-Alps. The Walsers of Piedmont (in Valsesia and Ossola) and the Val d'Aosta, the twelve Veronese Communes, centred on Giazza (so-called Cimbrian), and found in numerous areas of decreasing size in the Veneto, Trentino and Friuli, speak archaic German varieties. Finally there are enclaves of other Indo-European languages in southern Italy: Albanian (in Molise, Campania, Basilicata, Puglia, Calabria and Sicily), Croatian (in Molise), as a result of medieval and later migrations; Greek (in Puglia and Calabria), which may go back to medieval colonizations, or perhaps to the ancient Greek of Magna Graecia.

In Romanian territory there are, besides the Romanians, various other kinds of linguistic group: the most numerous (and the largest linguistic minority in Romance-speaking Europe), is formed by the Hungarians, with approximately a million and a half speakers, settled principally in the historical territory of Transylvania (including Banat, Crişana, Maramureş); then Rom (or Roman), German, Ukrainian, Serbian, Tatar, Turkish and other small linguistic groups. There are conspicuous minorities in the Republic of Moldova: principally Russian, Ukrainian and Turkish (Gagauz) (cf. also Varvaro, this volume, chapter 9, §5.3).

French borders Germanic varieties to the north-east and east (Flemish, middle Frankish, Rhine Frankish, Alemannic). Italian, the dialects of northern Italy, and Raeto-Romance varieties border on southern German varieties. The Lombard dialects of Canton Ticino and Romansh are in contact with Swiss German, a dialect of Alemannic stock; further east, Trentino dialect and Ladin are in contact with Tyrolean (a variety of the Bavarian dialect group); in the furthest east, Friulian is in contact with the Carinthian German dialect (Bavarian group) and Slovenian. In southern Italy, across the Ionian Sea, the historical contact between Greek and southern Italian varieties reappears.

Similarly, Sicily and southern Spain are separated from the Arabic- and Berber-speaking Maghreb, but also united with it, by small stretches of water. Romanian is in contact with Hungarian to the north-west and with Slavonic languages (Serbian to the south-west, Bulgarian to the south, Ukrainian to the north and north-east). Such contacts have left their historical stamp on the various Romance varieties (cf. Sala, this volume, chapter 6; Varvaro, this volume, chapter 9).

Until recently the Romance panorama was more varied than today. Some Romance varieties died out in the Middle Ages, such as the Mozarabic of the southern Iberian Peninsula, which yielded to Ibero-Romance varieties imported from the north. Dalmatian, spoken on the eastern shore of the Adriatic and at one stage used in writing, together with Latin and Croatian, in Dubrovnik (Ragusa) was already falling out of use from the fifteenth century under the pressure of Croatian and Venetian, and became completely extinct with the death of its last speaker on the island of Krk (Veglia) in 1898. The same fate befell Istrian, which persisted until the second half of the twentieth century in Dignano and Rovigno (Croatia). The old Friulian of Trieste died out in the first half of the nineteenth century. Another variety which is nearing extinction is Judaeo-Spanish, used among the Jewish communities of the Mediterranean and the Balkans (and by immigrants to Israel, the USA, etc.). Varieties on the brink of extinction are Istro-Romanian in the north-eastern Istrian peninsula, that of Bonifacio in southern Corsica, the Catalan spoken in the area of El Carche, in the Spanish region of Murcia. So-called *tabarchino*, the Ligurian dialect which survives in Sardinia, died out in an outcrop on Spanish territory near Alicante at the start of the twentieth century. The Genoese of Gibraltar is also extinct. The use of Venetan (dialects akin to that of Venice) in Dalmatia (Croatia) is drastically reduced, if not wholly extinct, although it was once widespread in the coastal towns and villages. Until about the nineteenth century the French, Occitan and Franco-Provençal dialects of the north and south of France, Belgium and Switzerland, while reduced to *patois*, were still alive and well, but now have nearly died out. The number of speakers of Italian dialects, Sardinian, Ladin and Friulian, and even the regional languages of Spain, is in decline but they show no sign of dying out within the short to medium term.

2. The origins of Romance Europe

The Romance languages continue *in situ* the Latin spoken in the western part of the Roman Empire. The dividing line between Latin-speaking and

Greek-speaking areas was established by Jireček (1893; 1901; 1911) on the basis of the predominance of Latin or Greek inscriptions, and was subsequently refined by Philippide (1925), Skok (1931; 1934), Gerov (1948–53; 1980) and Mihăescu (1978). The line seems to have started in the south from Lissus (Lezhë, north-western Albania), continuing north-eastward to the north of Stobi (near Gradsko, Macedonia), and to the south of Scupi (Skopje, Macedonia), then going first to the north-east, along the old administrative frontier between *Thracia* and *Moesia Superior*, then eastward, along the frontier between *Thracia* and *Moesia Inferior* (modern northern Bulgaria) to the shores of the Black Sea, where it turned northwards, up to the Danube delta (Gerov 1980). To the north of this line what was prevalently spoken was Latin, with Greek (and the other local ancient languages) to the south. A bilingual zone has been claimed for the western part of Thrace, south of the line. On the south-west shore of the Mediterranean, the Latin-speaking zone included the whole area west of *Cyrenaica* (north-eastern Libya).

Not all the Latin-speaking area developed a Romance language. The ‘lost’ territories – where the evidence of placenames, and loanwords surviving in the languages that replaced them suggest that a Romance language had originally formed – are labelled *Romania Submersa* or ‘the lost România’. Along the north-eastern continental frontier, the penetration of Germanic-speaking populations pushed westwards the ancient *limes* between *Romania* and *Germania*, which around the first century AD basically followed the course of the Rhine from its mouth as far as Mainz (*Magontiacum*), went south-east and joined the Danube near Regensburg (see also Varvaro, this volume, chapter 1, §8). Populations of Franconian stock settled along the lower Meuse and the Rhine (*Germania Inferior*), on the middle Rhine (northern *Germania Superior*) and along the Moselle (eastern *Gallia Belgica*). The course of the upper Rhine and the Danube (central and southern *Germania Superior*, northern and central *Raetia*) was occupied by populations of Alemannic and Swabian stock. In eastern *Raetia* and in *Noricum* (modern Bavaria and Austria south of the Danube) there settled peoples of Bavaric stock. Along the mid course of the Danube, Pannonia was abandoned by the Roman population between the fifth and seventh centuries, because of Gothic, Hunnish, Avar and Slav invasions. All trace of Latinity was already lost with the Hungarian occupation of the territory north of the Drava. Slav and Bulgar populations settled in southern Pannonia, Moesia and northern Macedonia and Thrace between the seventh and eighth centuries. *Moesia Superior* (modern eastern Serbia and northern Macedonia) and *Moesia Inferior* were probably abandoned by most of the indigenous population between the fifth and seventh centuries.

The Slav penetration into *Dalmatia* (Croatia, Bosnia-Herzegovina, western Serbia), from the seventh century to the early Middle Ages, was more gradual. African Latinity – which in the fourth century AD covered the provinces of *Tripolitania* (north-western Libya), *Byzacena* (southern Tunisia), *Proconsularis Zeugitana* (northern Tunisia), *Numidia* (north-eastern Algeria), *Mauretania Sitifensis* (northern Algeria), *Mauretania Caesariensis* (north-western Algeria) and *Mauretania Tingitana* (northern Morocco) – was obliterated by the seventh-century Arab invasion. In some zones the retreat of the *Romània* went on for centuries. In Switzerland, *franc-comtois*, Franco-Provençal and Romansh went on losing ground to Alemannic dialects until the fifteenth century. Romansh is still retreating today. In modern Alto Adige / Südtirol, the immigration of populations speaking Bavarian dialects continued to push back Ladin and Trentino varieties until the eighteenth century. The penetration of Germans and Slovenians in medieval and modern times has led to the southward and westward retreat of Friulian. Centuries of Slav and Venetian expansion led to the complete disappearance in the nineteenth century of Dalmatian, and to the decline of Istrian. Other areas of the Empire where Romance varieties did not develop are central and northern Albania, Britain (see Varvaro, this volume, chapter 1), and the Basque regions of the Pyrenees.

In contrast, extensive ‘neo-Romanized’ areas are found north of the Danube. The area of colonization of Dacia essentially corresponded to modern western Romania (Transylvania, Banat, Oltenia, western Muntenia). The thesis of ‘continuity’, according to which Romanian continues the Latin of Dacia (a Roman province from 107 to 275 AD, when it was abandoned by Aurelian), is not universally accepted. Some scholars hold that Romanian was formed wholly or in part to the south of the Danube, and that the current location of Romanian is the result of internal migrations. The historical, archaeological and linguistic data available do not seem adequate to give a definitive answer; see also Niculescu (1992). The modern extension of Romanian includes numerous areas that were not part of the Empire: Crişana, Maramureş, Moldova, Bucovina, Bessarabia, and central and eastern Muntenia. Among the new acquisitions are the areas south of the Jireček line occupied by Aromanian and Megleno-Romanian and the territories which Ibero-Romance and Gallo-Romance have gained from Basque. Then there are ‘re-Romanized’ territories: the eastern fringes of Breton, conquered by *oïl* varieties; the formerly Germanized territories within the French state and today partly retaken by French (Alsace, north-eastern Lorraine, western Flanders); the originally Flemish and now prevalently French area of Brussels; the (minority) Italian settlements in the German-speaking Alto Adige in the twentieth century;

the centres of the Istrian and Dalmatian coast – today almost totally Slavonicized, where colonial Venetian was widespread from the fourteenth century. It should be mentioned that in the view of some scholars (e.g., Gerhard Rohlfs, Giuliano Bonfante), medieval Sicily was ‘re-Romanized’ after the expulsion of the Arabs (on this issue see Tagliavini 1972:121, 410; Melazzo 1984; Fanciullo 1993).

3. Dynamics of the Romance languages in medieval and modern times

The crucial event in the Iberian Peninsula is the Arab conquest (711–18), whose first consequence was the division of the peninsula into a Muslim south and a Christian north, in which some Christian kingdoms organized resistance, with the long process of reconquest (*Reconquista*) ultimately leading to the complete elimination of Arab domination (1492). Around 1095 the county (later Kingdom) of Portugal broke away from the Christian kingdom of León (continuing the kingdom of Asturias) as had, already in the tenth century, the county (later kingdom) of Castile, which was soon to become the most important political and cultural centre in Christian Spain. In 1230 it was reunited with the kingdom of León, but this time under Castilian hegemony.

Reconquest meant, politically, the progressive occupation, and repopulation, of the central and southern peninsula by the northern Christian states and, linguistically, the southward expansion of the Romance languages spoken in the north, which replaced Arabic and absorbed the local, Mozarabic, Romance dialects. Not all the northern languages enjoyed equal success: in the west, Portuguese (now separate from Galician), in the centre Castilian, in the east Catalan, all achieved profound penetration, but Asturo-Leonese and Aragonese had a much lesser role.

In the part of the southern Iberian Peninsula occupied from 711 by the Arabs (*Al Andalus*), the Christian populace retained its own Romance speech, until the *Reconquista*, namely Mozarabic, a language also spoken by part of the Muslim populace. Given the predominance of Arabic and Hebrew in this area as languages of culture, our attestations of Mozarabic are largely indirect (treatises, placenames and personal names, and often the final verse of Arabic *muwaššahas*). From the end of the eleventh century, political persecution of Christians led to a major northward migration of the Mozarabic population, leaving wide areas uninhabited. The subsequent wholesale repopulation from the north favoured the replacement of the language of the Mozarabs, with traces surviving in placenames and occasional effects on the development of

the varieties (Portuguese, Valencian, Andalusian, Castilian) imported by the *reconquistadores*.

Castilian conquered the most ground in the *Reconquista*, occupying all the central-southern part of the peninsula, increasing its area of diffusion tenfold in three centuries. In 1253, Alfonso X 'the Wise' chose the language of the capital, Toledo, to be used in the royal chancellery, thereby establishing the supremacy of Castilian in the administrative sphere over the other languages of the kingdom, Galician and Asturo-Leonese. The great prestige of Castilian, as the language of the royal court and administration, but also as the vehicle of an already flourishing literary tradition, favoured its diffusion among the aristocracy and the clergy and in local civil and legal institutions. The political union achieved by the Catholic Monarchs (1479) widened its area of influence to Aragon, Catalonia and the area of Valencia. Under the Habsburgs (sixteenth–seventeenth centuries) Castilian spread into the Americas and Asia (the Philippines). Charles V proclaimed it the 'official language of diplomacy and the common language of all Christendom'. The process of centralization and repression of local autonomies which commenced with the Bourbons (1700–1931), had major linguistic repercussions, leading to a further consolidation of Castilian in Basque-, Catalan- and Galician-speaking areas. In 1780, Charles III decreed that Spanish should be the only language allowed in primary education, and in 1813 this disposition was extended to other levels of education. Reforms introduced by the 'ley Morano' of 1857, aimed at reducing the level of illiteracy, certainly further helped spread Castilian into areas where it had not previously been spoken. With the Romantic movement there emerged various movements for the preservation and promotion of local linguistic varieties. After the Primo de Rivera dictatorship (1923–30), which was hostile to local autonomies, a new phase began which led, with the arrival of the Republic and the promulgation of the democratic constitution, to the recognition of Catalan (1931), Galician (1936) and Basque (1936). Castilianization gathered strength under the Franco dictatorship (1939–75), which prohibited the use of regional languages in public and imposed the exclusive use of Spanish in administration, schools and the media. The 1978 democratic constitution established that Spanish is the sole official language of Spain, but also granted minority languages co-official status in their respective areas. From the early 1980s the various policies seeking to widen the social use of Catalan, Galician, Basque and, to a lesser extent, Asturo-Leonese and Aragonese managed to contain the expansion of Castilian, but not to stop it completely. Spanish is today the official language of the Kingdom of Spain, and spoken by all the inhabitants of the Spanish state (about 45 million). With

its dialectal varieties, it may be reckoned to be the mother tongue of over 36 million people in Europe.

Diffusion of Spanish outside Spain is due to two events. The first is the creation of an overseas empire. The second is the expulsion of the Jews from Spain in 1492. Having been accepted into the Ottoman Empire, these Jews were scattered in the Balkans, in what is now Romania, in Anatolia and in north Africa. These Sephardic communities kept until at least the mid twentieth century an archaic form of Castilian lexically enriched with forms from the local languages (but cf. Penny, 2002:27–29, for the admixture of features from other parts of the Iberian Peninsula, especially the west), and called *ladino* or *judezmo* by its speakers, and ‘Judaeo-Spanish’ by scholars. The Balkan communities were almost wholly destroyed during the Holocaust. The north African communities were largely expelled by nationalist Arab governments after the Second World War. Some groups survive in Greece, notably at Thessaloniki, in the Balkans (Bosnia-Herzegovina) and above all in Turkey. Many of the Sephardic Jews who escaped extermination emigrated to Israel and the Americas. There is also a community in France. Recent estimates put their numbers at about 200,000, although most speakers are aged over sixty (Montanyés Gómez 2005).

As for Asturo-Leonese, in the thirteenth century the kingdom of León comprised Asturias, the western third of Cantabria, the provinces of León, Zamora and Salamanca (in León), and the north-eastern quarter of the province of Cáceres (in Extremadura). Southward expansion did not always lead to homogeneous expansion of the language, because some zones (e.g., Salamanca) were repopulated by peoples from other parts of Spain. After the union of León with the kingdom of Castile (1230), Asturo-Leonese gradually lost ground to Castilian, the official language of the royal chancellery and one of great cultural prestige. Already in the Middle Ages Castilian began to spread onto the southern plains and the major urban centres of León, pushing the linguistic boundary westward. The gradual reduction of the domain of Asturo-Leonese has continued in modern times, affecting other areas as well. In western Cantabria and northern Extremadura it has been replaced by fundamentally Castilian dialects. In León it has retreated increasingly to the north and west, in rural and mountainous zones, disappearing completely from the towns. It stood its ground better in Asturias where, at least until the early twentieth century, the expansion of Castilian had been almost entirely limited to the middle and upper urban classes. In subsequent decades the profound social and political changes, and especially the repressive policies of the Franco regime, led to a rapid decline of Asturian. In Asturias, from the 1970s, a cultural movement for the defence and preservation of the local linguistic heritage has

been active. Modern Asturian, or 'Asturo-Leonese', 'Leonese' or simply *bable*, is spoken by about 600,000 people, alongside Castilian, the sole official language, in the Principality of Asturias, in León, and some areas of Cantabria and Extremadura (Fernández Rei 2007). It is spoken throughout Asturias except in those western areas which speak Galician. In Asturias its approximately 350,000 speakers in 2002 (one-third of the population) generally have good competence in the more prestigious Castilian (Llera Ramo and San Martín Antuña 2003; Fernández Rei 2007).

We turn now to Navarro-Aragonese, Aragon attained its modern boundaries at the start of the thirteenth century. After a period of Catalan influence, that of Castilian appears as predominant already in the fourteenth century, and becomes yet stronger after the union of the crowns of Aragon and Castile (1479). Thereafter there is a slow but steady northward retreat of the linguistic frontier so that by the mid twentieth century the indigenous variety is confined to the northern area (in the Pyrenees). The kingdom of Navarre, which at various stages of the Middle Ages was united with Aragon and formed a unitary linguistic territory with it, was Castilianized in the fifteenth century. La Rioja, which partly belonged to the Navarro-Aragonese linguistic domain, was Castilianized in the twelfth to thirteenth centuries.

Modern Aragonese is spoken in the northern part of the autonomous region of Aragon, known as Alto Aragón (Nagore Laín 2001). Most speakers are found in the Pyrenean valleys of the province of Huesca. It has retreated considerably in recent decades. According to Llera Ramo *et al.* (2001), in Alto Aragón, of 130,000 inhabitants, native speakers numbered just over 8 percent of the population, active speakers nearly 19 percent, and those with passive knowledge nearly 27 percent. It does not enjoy official status, but is protected by regional legislation.

In the Middle Ages, Portuguese and Galician were forms of the same linguistic variety, but the different destinies of Portugal and Galicia had major repercussions for their development. The reconquest of Galicia by the kings of León was effectively over by the eighth century. In the tenth, the southern boundary had pushed south as far as the Douro (Sp. Duero), almost reaching the Mondego. In 1093, in the territories south of the Minho, the county of Portugal was created, becoming an independent kingdom from 1143. At the end of the twelfth century there emerged a *koiné*, known to scholars as Galician-Portuguese, which in the thirteenth and fourteenth centuries was to become the vehicle of lyric poetry of *trobador* inspiration in Galicia and Portugal and initially in Castile, as well as in documentary, ecclesiastical and administrative functions, both north and south of the river Minho. In Portugal, the political

and cultural centre of the state had shifted as far south as Lisbon by 1255, the official language of the kingdom being influenced by the Romance varieties of the reconquered territories and assuming new linguistic features which differentiated it from the *koiné*. In 1249, with the conquest of the Algarve, Portuguese reaches its present territorial extent. Portuguese is today the official language of Portugal, which has over 10 million inhabitants.

In Galicia, which after brief periods of independence between the tenth and twelfth centuries was part of the kingdom of León, then Castile (1230) and eventually Spain (1479), Galician continued to be used until the end of the Middle Ages, with the increasing encroachment of Castilian hegemony. In the sixteenth century, Castilian established itself definitively as the language of the local aristocracy, while Galician started a decline which continued into the late nineteenth century. Between the sixteenth and eighteenth centuries, Castilianization extended to the urban bourgeoisie. At this period, Galician was perceived as an unprestigious rural dialectal variety. The political and cultural movements between the end of the nineteenth century and the 1930s, which sought to extend the use of the language, were unable to hold back Castilian, which gathered strength under the Franco dictatorship and the industrialization of the 1960s. From the 1970s there began a process of rediscovery and normalization of Galician, leading to its official recognition and its introduction into the public domain. Galician has been, with Castilian, an official language in the autonomous region of Galicia since 1981. It is used in all aspects of public life. According to a 2004 survey, 93 percent of the 1.5 million residents aged between fifteen and fifty-four claimed to understand it, and some 82 percent to speak it (MSG 2004). Native speakers formed just over 20 percent. The domain of Galician also continues to the east of Galicia proper, in the so-called *Franxa Exterior*, comprising the far west of the Principality of Asturias and the Community of Castile and León. In this area, the number of native speakers, about 35,000, varies from zone to zone. Transmission of the language within the family has declined considerably in the last decades (Fernández Rei 1999; 2003).

After absorbing the Balearic islands (1229–35) and Valencia (1238), Catalan expansion halted just south of Alicante. The only surviving linguistic trace of Catalan domination, between the thirteenth and fifteenth centuries, of wider areas of the central Mediterranean is the city of Alghero, in Sardinia, which was repopulated by Catalans between 1354 and 1372. When the Aragonese crown passed to the Castilian dynasty of the Trastámaras (1412), and especially after the union of the kingdoms of Castile and Aragon as the Kingdom of Spain (1479), Catalan began to be challenged by Castilian. From the end of the fifteenth century Castilian began to penetrate the Catalan aristocracy and

educated classes. But Catalan continued to be widely used until the beginning of the eighteenth century, when the strong centralism of the Bourbons led to the imposition of the exclusive use of Castilian in administration, justice, education, etc. This policy brought about a situation of diglossia, resulting in the partial Castilianization of the commercial and industrial bourgeoisie in the nineteenth and twentieth centuries respectively. The cultural revival movements which arose between the 1830s and 1880s (*Renaixença*) laid the ground for the experiments in self-government of 1914–25 and 1932–39, which sanctioned the equal official status of Castilian and Catalan in Catalonia. The process of Castilianization gathered new momentum during the anti-Catalan repression of the Franco regime (1939–75) and was favoured by the massive immigration of Castilian speakers, which reached its height between 1950 and 1970. This process was particularly intense in the towns of the Valencia region, which for centuries had included Castilian-speaking Aragonese territories to the west. After Franco, Catalan returned to co-official status in Catalonia (1979), where it has considerable social prestige and is used in all walks of public life, in the País Valencià (1982) (where it is mainly used by the middle and lower classes and has lower social prestige than Castilian) and the Balearics (1983), regaining its strength and considerably broadening its social use (see also Varvaro, this volume, chapter 9, §2.1). Catalan now has over seven million speakers in Europe, 3.4 million speaking it as a first language.² It is also spoken in the Principality of Andorra (where it is the sole official language); the French *département* of the Pyrénées-Orientales (where it enjoys some degree of recognition, but is rapidly retreating before French), and Alghero in Sardinia (where it is recognized under Italian law as a minority language).

The Romance dialects spoken in the northern part of France (known as *oïl* varieties after the old French word for ‘yes’; cf. ‘Occitan’ named for the Occitan *oc*, ‘yes’) and in part written during the Middle Ages are today mere *patois*, largely of limited social and geographical diffusion (old people, rural areas), and low social prestige.³

In the centre (Île-de-France, Centre, southern *départements* of Oise and Aisne in Picardy, southern Orne in Basse-Normandie, north-eastern Pays de la Loire, northern Allier in Auvergne, western fringe of Champagne–Ardenne) there is a wide area in which the original languages (*orléanais*, *tourangeau*, *berrichon*, *bourbonnais*) have been replaced by regional varieties of French, with some dialectal

2 See further *EULCat* (2003), *EULBal* (2004), *EULAr* (2004), *EULCatNord* (2004), *EULAL* (2004) and *Libre blanc* (2005).

3 There are, however, some very highly regarded contemporary literary figures who write or wrote in dialect, especially in Normandy (for instance, the poet Côtis-Capel).

phonetic and lexical characteristics (Simoni-Aurembou 2003). The indigenous varieties are also in sharp retreat everywhere else in the domain. The original boundary between *oil* and *oc* varieties ran north of the modern one, embracing much of the Poitou-Charentes region. By the thirteenth century, the *oc* varieties of these areas were completely ousted by northern dialects. Modern *poitevin-saintongeais* is an *oil* *patois* with some *oc* features. French expanded into the central and western areas (Champagne, Orléanais and Pays de la Loire, Brittany, Poitou, Saintonge) in the fourteenth century. The *scriptae* retained dialect features longest in Normandy, Picardy, Wallonia, Lorraine, Burgundy and Franche-Comté. About 1500, the written language was in line with that of the Ile de France throughout the north except Normandy and the north-east. Many of the northern varieties survived reasonably well in speech until the nineteenth century, being the only medium of communication for rural populations and the middle and lower classes. The use of these dialects declined at the end of the nineteenth century and the beginning of the twentieth, and again in the second half of the twentieth. Nowadays they have largely been replaced by regional or dialectal varieties of French.

The proximity of French, itself arising from a variety spoken in the Île de France, makes it difficult for speakers to perceive the distinctness of their own varieties, which often come to be considered degenerate forms of French. The 1999 census showed that adults who had had an *oil* variety as their initial language numbered about 570,000 (2 percent), but little more than 0.7 percent of the population of the *oil* region in question actually used it (Héran *et al.* 2002). Some studies, however, indicate that the varieties spoken in the northern swathe of the *oil* domain enjoy greater vitality. In the Pays de Caux (roughly, Seine-Maritime) 19.5 percent of the population spoke *cauchois* (see Bulot 2005). Speakers of *picard* in Picardy in the early 1990s made up 7 percent of the population of the area, with particular concentration in the *département* of the Somme (Parisot 1998; Parisot *et al.* 1998). The languages of the east (*champanois*, *lorrain*, *bourguignon*) are now on the brink of extinction. In the south-west, *poitevin-saintongeais* survives in peripheral and rural areas (Blanchet 1996; Auzanneau 1998). The only *oil* variety which enjoys a measure of recognition in France is *gallo*, included in 1989 among the languages admitted in public education and declared in 1999 an official language of the region, alongside French and Breton (Hervé 2005).

Oil varieties are also widespread in Switzerland, Belgium and the Channel Islands. *Franc-comtois* is still used in the mountainous areas of the north of Canton Jura (1,599 speakers in 1990; cf. Kristol 1998). In Belgium, various *oil* dialects are spoken: *champanois* and *lorrain* respectively in the west and south

of the province of Luxembourg, *picard* in the west of Hainaut, *wallon* in the provinces of Liège and Namur, in Walloon Brabant, in the east of Hainaut and the north of Luxembourg. In the late 1990s, *wallon* was spoken by under a million of the 4.5 million francophones, *picard* by about 200,000, *lorrain* by under 20,000, *champenois* by about a thousand (Fauconnier 1998). The Channel Islands, directly dependent on the English crown since 1204, traditionally belong to the Franco-Norman linguistic domain. Today, the original varieties, almost completely replaced by English, are still used on the islands of Jersey, Guernsey and Sark by the oldest people (about 8,000 speakers) (Toso 2006). See further Jones (2001; 2008) for the French of these islands.

The political and military successes of the French monarchy favoured the diffusion of the language of the royal court. From the mid twelfth century, the language of Paris and the Île de France acquired particular prestige and began to influence the literary and documentary *scriptae* of the north. After becoming the language of the royal chancellery under Louis IX and Philip the Fair, between the fourteenth and fifteenth centuries it gradually overlay the other *oïl* varieties. Its advance into the Occitan domain was slower, but it was effectively over by the end of the sixteenth century. A decisive factor was François I's Ordinance of Villers-Cotterêts (1539), imposing the use of French in all courts of law throughout the kingdom. The initial purpose was to ban the use of Latin and favour the *langaige maternel francoys* (lit. 'French mother tongue'), in the words of the text, but in time it was interpreted as banning equally the other tongues of the kingdom, particularly Occitan. French, as the 'langue du roi' ('the king's language'), became the only variety which could be used officially. The expansionist policies of the French crown between the sixteenth and seventeenth centuries brought numerous linguistic varieties within the kingdom: Basque (in what is now Pyrénées-Atlantiques); Catalan, in Roussillon; Corsican; Breton; Flemish in French Flanders; Franconian and Alemannic German dialects in Lorraine and Alsace (and in 1861 the Ligurian dialects in the eastern part of the county of Nice). In this period, the creation of a strongly centralized state favoured the diffusion of French in cities, especially among the aristocracy and the bourgeoisie. From the seventeenth century, French began to spread through European cultivated circles, thanks to the importance of France and the prestige of French literature, philosophy and science.

However, it took the Revolution to establish for the first time the ideal of national linguistic unification, to be realized by the promotion of French and the elimination of dialects and languages other than French. In 1794, abbé Henri Grégoire presented a *Rapport*, which, from its very title, asserted the

'need to annihilate the patois and universalize the French language'. Diffusion of French at the expense of dialects and minority languages was to remain an aim of all French governments from the nineteenth century until the mid twentieth. An 1864 inquiry into elementary education and the diffusion of French and the patois in France revealed that in the non-Romance areas and the Occitan-speaking south and Corsica over a quarter of the population did not know French. Linguistic unification gathered pace with industrialization (roughly from 1830), the introduction of obligatory primary schooling in French (1881–82) and the First World War – which brought together citizens from all over France. In 1951, however, the *loi Deixonne* allowed optional teaching of Breton, Catalan, Basque and Occitan in schools: for the first time the French state was officially recognizing the existence of regional languages. The 1960s saw multiple factors disrupting the transmission of regional languages within the family: industrialization, the growth of the cities and the movement of the population into cities, internal migration, the expansion of the railway and road network, and the mass media. The linguistic unification of France was effectively complete by the end of the twentieth century: varieties other than French are substantially in retreat, and all citizens of the Republic speak French, predominantly as their first language.

In modern times, French has almost completely absorbed the original *oïl* dialects and the Franco-Provençal dialects of Swiss Romandy. In Romance-speaking Belgium it has partly replaced the indigenous Picard and Walloon varieties. From the eighteenth century French spread into the Flemish areas, especially among the aristocracy and the *haute bourgeoisie*. The process intensified under French domination (1794–1814) and after independence (1830), but then retreated in the second half of the nineteenth century, when Flemish was allowed in the public domain and became the second official language (1898), and especially with the introduction of regionally based monolingualism (1930–32), subsequently embedded by the institutional reform of 1968–71 and by the federal constitution of 1994.

French is today an official language in France (where it is the only official language), Belgium (where it is the official language of the Wallonia region, and co-official with Flemish in Brussels), the Grand-Duchy of Luxembourg (where it has official status as a legislative, administrative and judiciary language), Switzerland (where it is one of the four official languages, and is mainly found in the western part of the country), Monaco (where it is the only official language) and the Italian region of Val d'Aosta (where it has official status together with Italian). It has about 67 million speakers in Europe, the great majority in France.

The domain of Occitan is roughly coextensive with the southern third of France. The boundary with *oïl* varieties goes from the Gironde estuary, passing north-east above Libourne, Confolens, Guéret and Montluçon, then south-east above Vichy, Ambert, Tain-l'Hermitage and Briançon, taking in some Alpine valleys at the eastern edges of the Italian provinces of Turin and Cuneo, down to the Mediterranean between Nice and Monaco. A transition zone between *oc* and *oïl*, 20–50 kilometres wide (the so-called *Croissant*) stretches from La Rochefoucault to Roanne. To the south-west the boundary of Occitan essentially follows the crest of the Pyrenees, retreating northwards in the Pyrénées-Atlantiques and Pyrénées-Orientales, and extending into northern Catalonia along the Aran valley. A medieval Waldensian colony of Occitan speakers exists at Guardia Piemontese in Calabria (Italy).

The south of France was the first Romance region to develop a literary and documentary *koiné* in the Middle Ages (twelfth century). The mid twelfth to mid thirteenth centuries was the golden age of the courtly lyric in the *langue d'oc*, which endured until the fifteenth century and was soon exported, particularly into northern Italy and Catalonia. The social and political conditions which had favoured the blossoming of this literary tradition were undermined by the anti-Albigensian crusade (1209–29), by the introduction of the Inquisition (1234), and by the northern feudal lords' seizure of the southern landed estates. The 'langue du roi' quickly penetrated the northern parts of the *oc* domain (1350–1400) and the chancelleries of the great feudal seignories (Foix, Armagnac, Polignac). The unification of the southern regions under the French crown was complete by the end of the fifteenth century. Between the mid fifteenth century and the end of the sixteenth French spread throughout the *oc* area, gradually replacing the local varieties first in administrative writings, then in documents of record and then in the private sphere. The only exception was Béarn, where the local *scripta* persisted in some domains until the Revolution. At the start of the seventeenth century, the aristocratic and bourgeois élites were bilingual. In the following centuries French gained further ground, without reaching ordinary people in the towns and countryside. The percentage of Occitan monoglots began to fall from 1881, with the introduction of free primary schooling (which became obligatory in 1882). There emerged a general system of diglossia, with French as the 'high' language and Occitan as the 'low', which gradually moved in the 1960s towards exclusive French monolingualism. There were indeed some high-minded but ill-fated attempts to introduce a written *koiné* for use in the official and public spheres, from the mid nineteenth century. For example the *Félibrige*, founded in 1854 by Frédéric Mistral, and *Institut d'Estudis Occitans* (IEO),

founded in 1945 (although their aims were in many respects very different). From 1951 Occitan has been recognized by the French Republic as a regional language, but the progressive loss of competence by the population and the lack of widespread social usage greatly limit its spread.

According to the 1999 census, about 526,000 adults spoke an *oc* variety; other estimates suggest 2 million speakers. Most speakers are over sixty, belong to the lower sectors of society and are mainly concentrated in rural areas. *Oc* varieties are largely limited to private and informal use. The northern varieties, especially *limousin* and *auvergnat*, are in absolute decline; the language is most actively used in the south (broadly, Gascon, *languedocien* and Provençal). For further details, see Deguillaume and Amrane (2002), MPM (1997; 1998), Dubarry and Dupouts (1995), Moreux (2004), Janik (2004), Blanchet *et al.* (2005) and Blanchet (2002:20–24; 2004:139). The Occitan varieties in Italy are legally recognized and protected, and have about 40,000 speakers (Arneodo 2003). At Guardia Piemontese, in the province of Cosenza, there is spoken an archaic Occitan variety, *guardiolo*, imported between the thirteenth and fourteenth century by Waldensian colonists from the Val Pellice. In the sixteenth century the Waldensian population was largely exterminated for religious reasons (Telmon 1994:941–42). In the early 1990s there were 400 speakers, and this variety is in rapid retreat. In the Val d'Aran in Spain, *aranès* has been recognized since 2006 by the *Generalitat de Catalunya* as a variety of Occitan, and, as a result, Occitan is a co-official language of Catalonia, not just in the Val d'Aran, but in the whole Catalan autonomous region.

Franco-Provençal is the name given (since Ascoli 1878) to the group of related varieties spoken in south-east France (southern Franche-Comté and Rhône-Alpes; see Tuaillon 1983), in Swiss Romandy, in the Val d'Aosta and some valleys in Piedmont in Italy. There are also Franco-Provençal dialects spoken in the villages of Faeto and Celle San Vito, in the province of Foggia (Puglia, southern Italy). These dialects are traditionally held to have been imported in the Middle Ages by settlers from the Isère or the Ain (Morosi 1890–92; Melillo 1959; Schüle 1978), but it has more recently been suggested that they could be the result of interference between indigenous Pugliese dialects and the speech of Occitan speakers from Provence or Piedmont, who settled in the area at some time between the thirteenth and fifteenth centuries (Telmon 1985; 1992).

From the earliest attestations, Franco-Provençal has formed a heterogeneous set of dialects, lacking any literary or documentary *koiné*. The reasons lie in the political fragmentation of the area over which it is spoken, abetted by the mountainous nature of the terrain. In the Swiss area, from the eleventh

to the thirteenth century, and from the fifteenth to the eighteenth century, the eastern boundary shifted westward under pressure from Alemannic dialects. Around Lyon, which in the Middle Ages was the greatest centre for the production of literary and documentary texts in Franco-Provençal, the French *scripta* did not appear until the early fourteenth century, but by the second half of that century it had established itself over the local language, making Lyon one of the major centres for the spread of French. In 1560, the House of Savoy introduced French as the official language of the transalpine dominions, the Val d'Aosta and western Piedmont. In the Protestant cities of Swiss Romandy (Geneva, Lausanne, Neuchâtel), the process of linguistic replacement began towards the end of the eighteenth century, and ended around the middle of the nineteenth. In the towns and the countryside, Franco-Provençal/French diglossia ceased by the 1930s. In agricultural cantons of Catholic tradition (Valais, Fribourg, Jura) the development has been slower, so that Franco-Provençal *patois* (Valais, Fribourg) and *oil patois* (Jura) have still not wholly disappeared. The attempt, from the 1970s, to introduce a unified norm (so-called *arpitan*, *arpetan* or *harpeitan*) has had no concrete result. In France Franco-Provençal is estimated today to be spoken by 60,000, mainly elderly, people. Most speakers are concentrated in the rural areas of Savoy and Bresse (Sibille 2003). In Switzerland, speakers in 2000 constituted about 1 percent of the population of Swiss Romandy (Lüdi and Werlen 2005:39). In Italy, Franco-Provençal has had the status of a minority language since 1999, but it is little used in public life. In Val d'Aosta, according to a 2001 survey, 62 percent of the nearly 92,000 inhabitants declared that they understood it, and just over 45 percent that they spoke it (Fondazione Emile Chanoux 2003). It is in retreat before both Italian and Piedmontese. It is also spoken in some valleys of the north-west of the province of Turin, in Piedmont, where speakers are estimated at about 100,000 (Telmon 1994). The dialects of Faeto and Celle San Vito are now severely threatened by the decline in population and the spread of the local Pugliese dialects and Italian.

Modern Romansh (*rumantsch/romontsch/rumauntsch*) is widespread in the Swiss canton of Grisons/Graubünden. It is spoken, often with considerable differences from place to place, in Surselva (Vorderrhein), Sutselva (Hinterrhein), in the valley of the Albula, in the Engadine and in Val Müstair. In antiquity, Romansh extended beyond Grisons into Val Venosta, where it came, from the sixth century, to be completely absorbed by Tyrolean German dialects. The Swiss Romansh area was orientated towards the Germanic world, when the bishopric of Chur was annexed to the archdiocese of Mainz. The original area of diffusion of Romansh has been constantly shrinking over the centuries, first

because of a strong influx of Germanic elements, then because of the constant assimilatory pressure of German which, until 1794, was the only official language of the region. Germanization increased yet further after the Lega Grigione joined the Swiss Confederation (1803). Since then the percentage of the Grisons population speaking Romansh as its mother tongue has more than halved. Romansh became one of the official languages of the Swiss Confederation in 1938. To overcome the problem of linguistic fragmentation, in 1982 Heinrich Schmid developed a unified written language, *Rumantsch Grischun*, which in 1996 became one of the official languages of the Canton and, since 1999, the official language in relations between the Confederation and Romansh-speaking citizens. None of this seems, however, to have checked the continuing replacement of Romansh by German. Speakers in 2000 were recorded at just over 35,000. In the Romansh linguistic area it was spoken by 69 percent of the population (Lüdi and Werlen 2005).

Ladin was for centuries divided between the county of Tyrol (the Sella group comprising the valleys of Gardena, Badia, Marebbe and Fassa, and Livinallongo; the Val di Non, Val di Sole and Ampezzo) and the Republic of Venice (Cadore and Comelico). The Gardena area was strongly influenced by German until the beginning of the twentieth century, leading to the retreat of Ladin. In the south, Romance varieties of the *trentino* type have pushed back the *fassano* varieties along the valley of the Avisio, and have almost completely absorbed the Ladin varieties of the Val di Non and the Val di Sole. The Ladin of the areas dominated by Venice has lost some of its characteristics under the influence of neighbouring Venetian dialects.

Since 1948 the Ladin communities of Alto Adige (Südtirol) have enjoyed the same kind of official recognition and protection as the German spoken there. Ladin is there used in placenames, is taught in schools, and used at the university, in the press and in administration. Since the mid 1990s the *Servisc de Planificazion y Elaborazion dl Lingaz Ladin* has been working on an official standardized version of Ladin. The result, called *ladin dolomitan* or *ladin standard*, blends linguistic features of the main Ladin varieties. While it was recognized as a minority language under Italian law in 1999, in Trentino and in the Veneto it does not enjoy full official status alongside Italian. In Trentino, its use in administration and teaching was only recently authorized (2006). In the Veneto there are no official measures in place to promote it as a second language. In 1969, Pellegrini (1971) estimated the total number of Ladin speakers in the areas of the Dolomites at 55,000.

Friuli, which had long periods of autonomy between the sixth and fifteenth centuries, largely came under the Republic of Venice in 1420, where it remained

until 1797. The south-western part of the region was penetrated from the west by Venetan varieties of the Trevisan type which, further to the west (the area of Sacile), completely supplanted the original Friulian dialects, while further east they gave rise to a Friulian–Venetan transition zone. Venetian dialect began to spread from the nineteenth century into the middle and upper strata of the population of the major urban centres (Udine, Cividale, Spilimbergo, Palmanova, Cervignano, etc.). Venetian-based varieties then became established in Venezia Giulia, in Trieste and in Muggia, replacing the old Friulian dialects. Original Venetan dialects are spoken on the coast, at Marano Lagunare, Grado and in the Monfalcone area (*bisiacco*): they probably continue indigenous varieties (Marcato 2002). After annexation to the Kingdom of Italy, Venetian influence gave way to Italian, which only began to achieve major penetration after the Second World War. Although use of Friulian has declined, especially since the 1960s, it still appears to be flourishing.

According to a survey in 1998–99, in the Friulian-speaking area of the Friuli–Venezia Giulia region (715,000 inhabitants), 60 percent habitually spoke Friulian (Picco 2001). In 1996 it was recognized as ‘one of the languages of the regional community’, and since 1999 it has enjoyed the status of a minority language. A unified written *koiné*, based on the central variety, has been made available, which should, at least in official usage, avoid local differences. Its diffusion in the public sphere is presently limited to placenames and some aspects of culture and public information.

The Romance varieties of Sardinia form a very distinctive dialect group on the Romance scene. Sardinian varieties proper occupy the central-northern (Logudorese or Logudorese-Nuorese) and southern (Campidanese) parts of the island. In the north-west (around Sassari) and north-east (Gallura), dialects of a Sardinian–Corsican mixed type are spoken.

From the fifth century AD onwards Sardinia developed separately from other regions of the former Roman Empire. From the eleventh or twelfth centuries, legal documents were drawn up in Sardinian. The subsequent economic and political penetration of the Genoese and the Pisans (eleventh to twelfth centuries), and Catalan domination (fourteenth to fifteenth centuries), later passing to Spain, had a major effect on the linguistic physiognomy of the island. The north-western (Sassari) and north-eastern (Gallura) areas, controlled by Pisa and Genoa, were repopulated in medieval and modern times by immigrants of Pisan, Ligurian and particularly Corsican origin. The dialects of these (so-called ‘Sardo-Corsican’ areas) nowadays present a decided affinity with the dialects of southern and central-western Corsica. Under Catalan domination (1323–1479), official use of Sardinian was gradually

replaced by that of Catalan, which in turn, after the union of the crowns of Aragon and Castile (1479), slowly gave way to Castilian. Remnants of Catalan and Castilian domination are, besides the Catalan-speaking linguistic 'island' of Alghero, the numerous Hispanisms, Catalanisms and Castilianisms which have entered some Sardinian varieties. Sardinia's passage to the House of Savoy (1718) brought the introduction of Italian as an official language (1764). With the unification of Italy there began to emerge a situation of diglossia (Italian as 'high' language and Sardinian as 'low' language), which today characterizes the majority of the Sardinian population.

In 2006 (Oppo 2007), 52 percent of the nearly 1,600,000 residents of Sardinia could speak Sardinian, and just over 83 percent could understand it, while just under 26 percent spoke it as a mother tongue. Numbers of speakers increase in the centre and south, in smaller centres, among old people, in the lower social classes, and among the relatively uneducated. Transmission from one generation to the next has dropped sharply since the 1970s. Sardinian is recognized as the second official language of the autonomous region of Sardinia. A written unified variety, the *Limba Sarda Comuna*, was introduced into regional administration from 2006, but Sardinian remains almost completely restricted to day-to-day, family and private interaction.

Istrian has been attested since the nineteenth century in the south-western Istrian peninsula, in Rovigno, Dignano, Valle, Gallesano, Fasana and Sissano. The evidence of placenames suggests that it once covered the central part of the peninsula (Crevatin 1989), gradually shrinking following the expansion of colonial Venetian on the coasts and the penetration of Slav populations into the hinterland. Until recently, it held out in Rovinj (It. Rovigno) and Vodnjan (It. Dignano), but is now virtually extinct (Delton 1999; 2000; 2001).

The dialectal variety within Italy is unparalleled in Romance or in any other linguistic domain in Europe. There are three main areas, each internally differentiated:

- northern Italian dialects: Piedmontese, Ligurian, Lombard, Trentino, Venetan, Emilian, Romagnol;
- Tuscan dialects: the varieties of Tuscany (except Massa, which belongs to the northern type);
- central and southern dialects: Marchigiano-Umbrian-northern Laziale (minus the varieties of the Marche to the north of Senigallia, of Romagnol type, but including the dialect of L'Aquila in Abruzzo), southern Laziale-Abruzzese-Molisan-northern Pugliese-Campanian-Lucanian-northern Calabrian, Salentine-southern Calabrian-Sicilian.

Italo-Romance varieties are spoken not only in Italy but also outside it. A western Ligurian dialect (recognized as a national language – although in reality French is the official language of the state) is spoken by most of the 6,000 indigenous inhabitants of the Principality of Monaco. Ligurian-speaking communities are also found in the French *département* of the Alpes Maritimes along the border with the Italian province of Imperia, mainly in the coastal areas of Roquebrune and Cap Martin and some villages in the upper valley of the Roya (Dalbera 2003; Toso 2006), which passed to France in 1861. The Ligurian dialect of Bonifacio in Corsica, imported from the late twelfth century, is now virtually extinct. In the Swiss cantons of Ticino and Grisons, dialects of the Lombard type are spoken, and just over 100,000 people spoke dialect (albeit not exclusively) in the family at the end of the twentieth century (Lüdi and Werlen 2005). In Slovenia and Croatia, Veneto-Giulian dialects are spoken in some towns on the Istrian and Dalmatian coast, the remains of centuries of expansion of Venetian. After 1992, the Italian population of Istria, already drastically reduced after the exodus of the period 1945–60, was divided between Slovenia and Croatia. In Slovenia, in 2002, just over 3,500 inhabitants were recorded as having ‘Italian’ as their mother tongue, most of them concentrated in the towns of Izola/Isola, Koper/Capodistria and Piran/Pirano (POPIS 2002). In Croatia in 2001, just over 20,000 ‘Italian’ speakers were recorded (CROSTAT 2006). Much of the Italian community of Slovenia and Croatia is triglossic, speaking Slovenian/Croatian in official contexts, learning Italian at school as a second language, and using their dialect in the family and informal contexts. In Croatia, over 70 percent of the population of Italian origin spoke dialect in the 1990s (Ursini 2002).

The Italian dialects are local continuations of Latin, and not varieties which have developed from Tuscan-based Italian. Almost all Italian dialects were being written in the Middle Ages, for administrative, religious and often artistic purposes, subsequently giving way to Tuscan, which was soon called ‘Italian’. The influence of the Tuscan *scripta* is manifest in many areas of Italy by the fourteenth (Veneto, Emilia, Naples) and the fifteenth (Lombardy, Piedmont, Rome) centuries, and is generalized in the sixteenth century, when Tuscan imposes itself on dialects as a written language in all domains – literary, religious (subordinate to Latin), administrative, judicial and scientific. Until the unification of Italy, no political event had permitted the diffusion of Italian beyond restricted circles. Only in Rome did the prevalent use of Tuscan in the papal court, and then the repopulation of the city by Tuscans in the fifteenth and sixteenth centuries, profoundly influence the local variety, originally of a central-southern type (Ernst 1970; Trifone 1990; 1992). After Unification, the

use of dialect retreated everywhere, albeit to different degrees in different areas. Apart from regions speaking varieties structurally akin to Italian (Tuscany, Rome, parts of the Marche, Umbria and northern Lazio), the regions where the retreat of the dialects is strongest are those where, especially after 1960, there has been major industrial development, leading to internal immigration and urbanization: Piedmont, Lombardy and Liguria. At the other extreme are economically less developed (Sicily, Campania, Calabria, Basilicata, Puglia) or recently industrialized (Veneto, Trentino) areas.

The commercial and military expansion of many Italian cities in the Middle Ages led to the diffusion of their linguistic varieties beyond their original boundaries. Pisa, from the eleventh to the thirteenth century, dominated northern Sardinia and northern Corsica, leading to notable Tuscanization of the local varieties. Pisa was replaced in the fourteenth century by Genoa, which left obvious traces in the dialect of Bonifacio (in Corsica) and of the island of La Maddalena (Sardinia). In Sardinia there are also Genoese varieties at Carloforte and Calasetta on the islands of San Pietro and San Antioco (south-west), founded by Ligurian colonists from Pegli who in the sixteenth century had colonized the Tunisian island of Tabarka, and moved to their present location in the eighteenth century: their dialect is called *tabarchino*. The centuries-long domination by the Republic of Venice led to diffusion of Venetan in Friuli, Istria, the gulf of Quarnaro and the Dalmatian coast. When these territories passed to Communist Yugoslavia (1954), with the consequent mass emigration of the Italians (some 200,000), the number of Venetan speakers fell sharply.

Other shifts in the geographical distribution of dialects are due to more or less officially 'planned' colonizations. From the eleventh to the thirteenth century groups originating in north-west Italy were transferred into the hinterland of Sicily to repopulate uninhabited areas. Later (twelfth–thirteenth centuries), Ligurian settlers were introduced into Basilicata and Campania. In the Tremiti islands (province of Foggia) a Neapolitan dialect is spoken, the result of resettlements carried out by the Bourbon authorities from 1792. The colonization policy of the Fascist regime led to the settlement of colonists from north-eastern regions (mainly the Veneto, but also Friuli and the Ferrara region) into the provinces of Grosseto (Tuscany), Latina (Lazio), Carbonia and Oristano and around Alghero (Sardinia). In some areas (Grosseto, Latina) these communities kept their dialect until recently. There have been numerous exports of dialects out of Italy. In Gibraltar, after it passed to the British (1713), the authorities encouraged immigration by a substantial Genoese community. Their dialect survived into the 1980s. Among more recent migrations are those

of Ligurians and Pugliese into the Crimea (1830–70), that of Trentino settlers into Bosnia-Herzegovina (1878–82), that of Friulians, Venetans and people from Trento into Romania (second half of the nineteenth to first half of the twentieth centuries). While in the Italian communities in Crimea the original dialect has long since ceased to be used, the Trentino community in Bosnia and the Friulians and Venetans in Romania have maintained the use of their dialect, at least until recent times (Toso 2006; Scagno *et al.* 2008).

In some areas of the Italian peninsula (the south, Veneto), dialects to this day enjoy a measure of social prestige, even if they have no official legal recognition (they are so recognized in some regions, such as Piedmont). This situation was still general in the first half of the twentieth century. Today, in most regions of Italy, the linguistic situation is of a diglossic type, with Italian as the ‘high’ variety and dialect as the ‘low’ variety. Italian may have become ever more predominant in the private sphere, especially over the last thirty years and particularly in the younger generations, but there is still considerable use of dialect in informal communication. In 2006 (ISTAT 2007), people who spoke dialect in Italy (exclusively or as well as Italian) numbered, out of a population of just over 55 million people over six years old, some 49 percent within the family, and 46 percent in relations with friends, but just 24 percent with strangers. It should be realized that the use of dialect is mainly concentrated in two areas, the south and the north-east. In the south (with Sicily) over 70 percent of the population uses dialect in the family, albeit not exclusively. The figures for the north-east are: 70 percent in the Veneto and 64 percent in the province of Trento. In central Italy, only in the Marche and Umbria is the use of dialect higher than the national average. Dialect is least widespread in Piedmont (35%), Lombardy (36%) and Liguria (26%). In Lazio and Tuscany the percentage of dialect speakers is respectively 35 percent and 12 percent, but for Tuscany and Rome (in Lazio) the distinction between ‘dialect’ and ‘Italian’ is far from clear-cut.

‘Corsican’ is really a cluster of dialects divided into two principal groups, *cismontincu*, spoken in the north-east of the island, and *pumontincu*, in the south-west. Until the mid eighteenth century, Corsica gravitated politically and linguistically towards Italy: it was under Pisan domination from 1077 to 1284, and under Genoese domination from 1284 to 1768. This gave rise to Tuscan influence on Corsican varieties, particularly in the centre and north. Italian was the administrative language until annexation to France (1769) and the cultural language until the mid nineteenth century. The penetration of French became more intense in the post First World War period and then in the 1960s. Active use of the native varieties has been further reduced in the last

few decades, but has still remained strong. In 1974 the local home rule movement succeeded in getting Corsican recognized as a regional language. It is taught in schools and universities, but is little used in public life and in recent decades has been in retreat before French. The 1999 census recorded 133,000 adult speakers in France, 90,000 of them resident on the island (but cf. Fusina 2003).

Italian is today the official language of the Republic of Italy, the great majority of whose inhabitants (59 million) may be considered Italian speakers – although many are bilingual, also speaking an Italian dialect or an officially recognized minority language. Italian is also the only official language of the Republic of San Marino (about 31,000 inhabitants), where it coexists with a Romagnol dialect. Italian is also the language of everyday usage in the Vatican City (about 800 inhabitants), where Latin is the official language. In Switzerland, Italian is one of the four official languages of the Confederation. At cantonal level, it is the only official language of Canton Ticino and one of the three official languages of the canton of Grisons/Graubünden, together with German and Romansh. Italian is the second official language in Croatian Istria and in some communes of Slovenian Istria (about 24,000 speakers), and is recognized and officially protected as a minority language of the Italian communities in Romania (about 3,000), in Bosnia (about 1,000) and in the Republic of Moldova (about 100) (Toso 2006).

The basis of modern Italian is fourteenth-century Florentine, which had already been diffused in written usage – mainly thanks to the works of Dante, Petrarch and Boccaccio, and which in the sixteenth century established itself as the common language of Italy's literati and spread into the chancelleries of various Italian states. It is likely that only narrow élites and the geographically limited group of the Tuscans, plus the inhabitants of Rome and parts of central Italy where varieties close to Tuscan are spoken (the province of Ancona, the areas of Perugia, Orvieto, northern Lazio), spoke Italian in past centuries (Castellani 1982a). The number of Italophones immediately after Unification (1861) was between 2,200,000 and 2,900,000, or between 9.5 percent and 12.55 percent of the population, for Castellani (1982a) – just 2.5 percent for De Mauro (1963). However, De Mauro's estimate may be too conservative, because it excludes from this calculation (illiterate) native Tuscan speakers, and speakers of varieties closely related to Tuscan, even though they might reasonably be regarded as speaking 'Italian'. Elsewhere, Italian was for a long time a predominantly written language, used in culture and science, in chancelleries, in administration and in the law (cf. also Varvaro, this volume, chapter 9, §2.3). In the past 150 years its oral use has spread widely. An initial, limited expansion, of the national language was favoured,

between the end of the nineteenth century and the beginning of the twentieth, by the creation of free primary schools (1859), made obligatory in 1877, by industrial development (in the north-west) and consequent urbanization, by the sending of state employees of different origin throughout Italy, by the introduction of obligatory military service, by the press, and by internal and external immigration (De Mauro 1963). The First World War (in Italy, from 1915 to 1918) brought vast numbers of dialect-speaking male citizens into contact for the first time; evacuations from war-torn zones brought whole families into contact with inhabitants of other areas speaking different dialects. The Fascist regime (1922–1943) instigated a harsh policy of repression and assimilation of German and Slav linguistic minorities, whilst measures against dialects were fairly mild. At the end of the Second World War, Italian speakers were still a minority with respect to dialect speakers (De Mauro 1963). The retreat of the predominant use of dialect was due to the profound economic and social changes of the 1950s and 1960s. The rapid shift to Italian by many dialect speakers in recent years has determined the formation of various kinds of ‘regional Italian’, each closely connected – especially in phonology and lexicon – with the local dialect. Yet use of dialect and regional languages is still very strong in some areas of Italy. According to a survey, in 1987 those who spoke only dialect, i.e., Italians who do not know Italian, still numbered 13 percent of the population.

Friuli, Sardinia, the Slovenian-speaking region in the province of Udine, the Ladin-speaking communities of the Veneto and Trentino, and the numerous linguistic ‘colonies’ speaking other languages scattered across Italy have been gravitating linguistically towards Italian for centuries. Other areas speaking different languages have only recently started to do so: the Val d’Aosta in 1861, Alto Adige and the Slovenian areas of Trieste and Gorizia in 1919. These latter were granted administrative and linguistic autonomy by the Italian Republic from the end of the Second World War. The range of varieties receiving official protection was considerably increased by a law of 15 December 1999, which recognizes as languages of ‘historical’ minorities: Albanian, Greek, Slovenian, Croatian, German; and the Romance languages Sardinian, Franco-Provençal, Occitan, Friulian, Ladin, Catalan and French. These languages may be used in placenames and certain administrative documents (beside Italian), and may also be taught in schools. After long legislative discussion, the same rights were not accorded to the dialects, thereby confirming their officially and socially inferior status with regard to Italian.

In Switzerland, in canton Ticino and Grisons/Graubünden, Italian has been an official language since 1803 and has spread progressively at the expense of

the local dialects of the Lombard type, which nonetheless remain fairly widespread (Moretti and Spiess 2002). There were just under 471,000 Italian speakers in these areas in 2000 (Lüdi and Werlen 2005). Penetration of Italian into the Venetan-speaking areas of Istria (now in Slovenia and Croatia) has been less intense, and the use of dialect remains strong there (Ursini 2002). Italian was an official and cultural language of Malta while it was ruled by the Order of the Knights of St John (1522–1800). Today the official languages are Maltese and English, but Italian is known by the élites and to some, frankly very limited, extent by the general population.⁴

Dalmatian designates the cluster of indigenous Romance varieties of some areas of Dalmatia, which had largely died out by 1500, probably surviving just on the island of Krk (Veglia), where it was attested until the nineteenth century (Muljačić 1995; 2000). Originally it was spoken in south-western Krk/Veglia, the islands of Cres/Cherso and Lošinj/Lussino, the coasts of the Gulf of Quarnero, the towns of Zadar/Zara, Šibenik/Sebenico, Split/Spalato, Dubrovnik/Ragusa, Perast/Perasto and Kotor/Cattaro, the islands of Brač/Brazza, Hvar/Lesina, Vis/Lissa and Korčula/Curzola (Doria 1989). The old varieties of Zara, Spalato and Ragusa are known to us from documents of a practical nature, the earliest of which is a letter of 1280. Even at that time Dalmatian appears subject to the influence of Venetian, by which it was ultimately completely absorbed. Ragusan (the Dalmatian of what is now Dubrovnik, in Croatia), is the variety best represented in ancient documents: in the fifteenth century debates were still held in the city's senate in Ragusan (alongside Venetian and Croatian). Dalmatian was absorbed by Venetian on the coast, while Croatian encroached from the hinterland. The island of Krk/Veglia held out longest, and we have direct transcriptions of the dialect of the town of Krk, gathered particularly from its last speaker, Tuone Udaina (Bartoli 1906).

For centuries the Romanians lived cheek by jowl with various peoples: Slavs (Bulgarians, Serbs, Russians, Ukrainians), Hungarians, Pechenegs, Cumans, Tatars, Mongols, Turks, Greeks, etc. The principalities of Moldova and Muntenia (or Wallachia, Ro. Țara românească), tributaries of the Ottoman Empire from the fourteenth century, united in 1859 and became independent in 1878. The official language of the new state was the, basically Muntenian, literary variety, which gradually gained ground from the second

4 Italian has today completely lost its standing as a Mediterranean lingua franca, but we might mention here that it was for centuries, and even as late as the nineteenth century, extensively used in the Mediterranean area as a language of commerce and diplomacy. See, for example, Cremona (2002; 2003), and now Baglioni (2010).

half of the sixteenth century and became definitively established between 1830 and 1880. Its diffusion was certainly favoured by the relative homogeneity of the Daco-Romanian linguistic area, divided into just two dialectal macroareas: the northern, comprising Moldova, Transylvania, Maramureş, Crişana and Banat, and the southern, comprising Oltenia and Muntenia. After the First World War, 'Greater Romania' was created, comprising also Transylvania, Crişana, Maramureş, eastern Banat, southern Dobrogea, Bucovina and Bessarabia. These regions were home to substantial minorities speaking other languages, particularly Hungarians, Germans, Bulgarians, Serbs and Ukrainians. Transylvania was occupied by the Hungarians in the tenth century, becoming part of the Kingdom of Hungary in the following century. In the thirteenth century, the Székely, a Hungarian-speaking population, but one traditionally distinct from the Hungarians, settled in the east and south-east of the region. In the twelfth and thirteenth centuries the Hungarian sovereigns encouraged major immigration of German settlers into Transylvania (the so-called *Săşi* 'Saxons'). Other German populations reached Banat and Maramureş in the eighteenth century, and later (1841) Dobrogea. Between 1526 and the end of the seventeenth century Transylvania was an autonomous principality, subject to the Ottoman Empire. In 1699 it fell, with Hungary, under Habsburg rule and remained thus until 1 December 1918. In the Communist period, particularly between the 1960s and 1980s, many of the Saxons emigrated. The percentage of Hungarians also declined, albeit to a lesser degree. In 1918 the historical Banat was divided between Romania, Yugoslavia and Hungary: this involved the incorporation of Serbian and Croatian minorities into the Romanian Banat, and of Romanian minorities into the Yugoslav Voivodina. Other Romanian communities were incorporated by Serbia in 1829–33, following the annexation of the Timoc valley (Krajina). The 'Quadrilateral', in southern Dobrogea, was a prevalently Bulgarian area but after its transfer to Romania in 1918 it was intensely Romanianized, thanks to the introduction of Aromanians, Megleno-Romanians, and Timoc Romanians. In 1940 it was definitively ceded to Bulgaria, and the Romanian population was largely transferred to northern Dobrogea. Broad areas of Bucovina and Bessarabia became first part of the USSR (1940, 1944) and then part of independent Ukraine after the end of Soviet domination: northern Bucovina, the region of Hotin and Herţa (the latter with a Romanian majority), along the northern border; Bugeac (or *Basarabia veche* 'old Bessarabia'), along the eastern border.

The modern Republic of Moldova occupies the historical region of Bessarabia, excepting the southern part (Bugeac) which was annexed to the Ukraine. In this region, which from the fourteenth to the nineteenth century had formed the

eastern part of the Principality of Moldova, are spoken Romanian dialects of the Moldovan type. In 1812 it entered the Russian Empire, where it remained until 1918, when it was united with Romania. Those territories beyond the Dniester which remained in the USSR became the Soviet Socialist Republic of Moldova, with a prevalently Ukrainian population and a minority (about 30%) of Romanian speakers. In 1944, Bessarabia was detached from Romania and two-thirds of it was annexed to Soviet Moldova, the other third going to the Ukraine (Bugeac). In the new state the majority of the population spoke Romanian, but there were also substantial Ukrainian and Russian minorities, which further increased between 1944 and 1979. The national language was declared to be 'Moldovan' (largely identical to literary Romanian but written in the Cyrillic alphabet), but the general-purpose and administrative language was Russian. With the collapse of the USSR (1989–91), the Cyrillic alphabet, which had been replaced by the Roman alphabet during annexation to Romania, was definitively abandoned and Moldovan (in official documents of the new state sometimes called 'Romanian' and sometimes 'Moldovan') became the only official language of the Republic. In 1990 the territory beyond the Dniester (so-called Transnistria) unilaterally declared independence, adopting as its official languages Russian, Ukrainian and Moldovan (written in Cyrillic).

Romanian is the sole official language of the modern Republic of Romania (over 22 million inhabitants in 2006), and in 2002 (*INSSE* 2002) was the mother tongue of 91 percent of the population; linguistic minorities are generally bilingual. It is also the official language of the Republic of Moldova,⁵ where it is also known as 'Moldovan' (Ro. *limba moldovenească*). Of 3.4 million residents in 2004 (*BNSRM* 2004), 76.5 percent declared Romanian–Moldovan as their native language (16.5% Romanian, 60% Moldovan). In the Transnistrian Republic (the self-proclaimed 'Pridnestrovian Moldavian Republic'), Moldovan (Romanian) enjoys official status alongside Russian and Ukrainian. In 2004 32 percent of the population (just over 177,000) declared themselves ethnic 'Moldovans'.

There are also numerous Romanian-speaking communities outside Romania and Moldova, in north-east Bulgaria, the Republic of Serbia (Timoc valley and Voivodina), Hungary, particularly along the Romanian border, Ukraine, along the Romanian and Moldovan border (in Transcarpathia, Cernăuți and Odessa regions it is protected as a minority language and taught in schools, but also elsewhere). See further Ionescu (1999), *NSIRB* (2001), *SORS* (2003), *HCSO* (2004) and *SSCU* (2006).

5 Russian, which until independence from the USSR (1991) was the language of administration and the state apparatus, remains widely used in the public sphere.

The separation of Daco-Romanians and Aromanians is generally held nowadays to have occurred between the tenth and eleventh centuries. The Istro-Romanians broke away from the Daco-Romanians not before the thirteenth century. The classification of Megleno-Romanian is unclear, some linking it with the Daco-Romanian–Istro-Romanian group and others with Aromanian.⁶ The four varieties seem to have split, in any case, by the fourteenth century.

The Aromanians – or Macedo-Romanians or ‘Vlachs’ – live in small communities scattered over much of the Balkan Peninsula, especially southern Albania, central and northern Greece and south-western Macedonia.⁷ There are five main linguistic subdivisions: *pindeni* and *grămoșteni*, settled mostly in Greece and the Republic of Macedonia; and *muzăcheari*, *moscopoleni* and *fărșeroți*, settled mainly in Albania. Two small individual groups are Aromanians of Maloviște and Gopeš (west of Bitola) and G. Belica and D. Belica (north-west of Struga) in the Republic of Macedonia. For centuries the Aromanians have lived mainly in mountainous areas. They are first indicated in the Balkan area in 976. Their way of life originally developed in the area of the Pindus mountains. They began to give up their originally nomadic life between the fourteenth and fifteenth centuries, expanding northwards, on to Mount Gramos and – in Albania – into the mountainous area around Korçë and Përmet, later settling on the plains of Thessaly and the plain of Myzeqeja. In the following centuries they founded numerous villages in western and central Macedonia, in the region of Bitola and Lake Ohrid (seventeenth–eighteenth centuries), and later eastern Macedonia and western Rodope (eighteenth–nineteenth centuries) (Kahl 1999; 2006). In the first half of the twentieth century, many communities were moved to Romania, first the ‘Quadrilateral’ and then northern Dobrogea. Movement into cities in the second half led to loss of traditional culture and assimilation. Many Aromanians subsequently emigrated to France, Germany, Canada, the United States and Australia.

Lack of official statistics in Albania and Greece makes total modern numbers hard to assess. In Macedonia – the only country where Aromanian enjoys official recognition – the 2001 census recorded 8,467 *vlasi* (but this figure includes the mainly Megleno-Romanian *vlasi* of Gevgelja). The figures from the 2001 Bulgarian census are less reliable, since Aromanians and Daco-Romanians are

6 See Dahmen (1989) and Ferro (1992:225–32). Also the chapters on Aromanian, Megleno-Romanian and Istro-Romanian in Rusu (1984). Recently, the thesis that Megleno-Romanian derives from Daco-Romanian, already put forward by Ovid Densusianu, has been revived with new arguments by Atanasov (2002).

7 See also Saramandu (1984; 1988), Dahmen and Kramer (1985–94), Ivănescu (2000) and especially Kahl (1999; 2006) and Demirtaş-Coşkun (2001).

lumped together. In the traditionally Aromanian areas of the south-west, only 204 people declared themselves to be ethnically 'vlachs' or 'Romanians' (NSIRB 2001). In Greece, the Aromanians have no official recognition. In the last census recording them as a separate category (1951) they were 39,885 (Angelopoulos 1979). In Albania, the last census record for Aromanian as a minority language was in 1955, with 4,249 speakers. In Romania, the 2002 census shows 25,053 people declaring themselves to be Aromanian. Many scholars cast doubt on the reliability of these data, especially for Greece and Albania. Authoritative recent studies concur on between 200,000 and 300,000 (Ivănescu 2000; Dahmen 2005; Kahl 2006). But there are also lower estimates (100,000 for Winniffrith 2002).

Megleno-Romanian is the name given by linguists to the Romanian variety traditionally spoken in the Meglen, a region to the north-west of Thessaloniki, on the boundary between Greece and the Republic of Macedonia, along the course of the river Axios/Vardar. Speakers call themselves generically *vlaș* 'Vlachs' (sg *vla* or *vlau*) (see further Atanasov 1984). They are also to be found in the many towns of Greek Macedonia and the Republic of Macedonia.⁸ The presence of Megleno-Romanians in the Meglen is documented from the late Middle Ages. They remained there, as farmers, until after the First World War, when they were divided between Greece and Yugoslavia. The villagers of Notia, who had converted to Islam, were transferred to Turkey. A large number of Megleno-Romanians emigrated to Romania in the 1930s, ultimately to Dobrogea. Others emigrated to Voivodina after the Second World War. From the 1950s, many moved to Greek and Macedonian cities. Large groups also emigrated to the countries of western Europe. In Gevgelija (south-east Macedonia), in the 2002 census, 214 people declared themselves *vłasi*; elsewhere the term *vłasi* also includes Aromanians. There are no statistics for the numbers of speakers in Greece. Estimates vary between 4,000 to 5,000 (Atanasov 1984; 2002; Kahl 2006) and 9,000 to 10,000 (Țîrcomnicu 2004).

Istro-Romanian is the name given by scholars to a variety of Romanian now spoken in north-east Istria (Croatia), in some localities around Mount Učka.⁹ The Istro-Romanians are probably a surviving branch of the Morlachs (Grk. μαυρόβλαχοι 'black Vlachs'), who were shepherds of Romanian origin recorded in Bosnia, Serbia and Croatia in the late Middle Ages and in Istria from the fourteenth–fifteenth centuries. They settled in Žejane in 1510 (Pușcariu 1926). Old records and toponyms show that originally they were

8 Atanasov (1984; 2002); Țîrcomnicu (2004). Cf. also Capidan (1925).

9 Orbanić (1995); Filipi (2003). Istro-Romanian is now extinct in some localities cited by Pușcariu (1926) and Kovačec (1971; 1984) – Dolinščina, Trkovci, Perasi, Grobnik, etc.

widely settled in Istria, as well as in the islands of Krk and Rab. According to the most recent statistics, there are some 200–250 speakers left in Istria, mainly elderly and all bilingual in Croatian (Orbanić 1995; Filipi 2003).

4. Diglossia, bilingualism, multilingualism

The establishment and consolidation of national states gave rise to various situations of linguistic contact. The ongoing conflict between a process of substitution in favour of national languages and a process of resistance of minority languages involving their linguistic normalization has been a leitmotif of the history of the Romance languages from the beginning of the modern age, and in some cases from even earlier. The modern sociolinguistic situation of the Romance varieties may be efficiently described in terms of bilingualism and diglossia.¹⁰ The former denotes the fact that members of some community master two distinct linguistic codes. The latter indicates the coexistence in the same community of two functionally differentiated linguistic varieties, one of which – the ‘high’ variety – is used only in formal contexts, while the other – the ‘low’ variety – is used only in informal contexts.

The establishment of Castilian, the *langue d’oïl* of the Île de France and Florentine as the official languages of Spain, France and Italy meant that those parts of these countries in which other varieties were used were in a situation of diglossia. For more or less extensive periods, Catalan, Galician, Asturian and Aragonese, *oïl* varieties, Occitan and Franco-Provençal, Corsican, Sardinian, Friulian, and Italian dialects were excluded from official usage and confined to the realm of the informal. The expansion of the national languages first led to a phase of ‘diglossia without bilingualism’, in which the ‘high’ variety was spoken only by the upper strata of society (who also used the other variety), while the ordinary people knew only the ‘low’ variety. Subsequently, bilingualism reached other sectors of the population, and the ‘high’ variety began to be used by part of the community in informal contexts as well (so-called ‘contaminated diglossia’ or ‘dialia’) (Berruto 1995). In many cases the process went further, so that more or less sizeable portions of society ended up speaking only the ‘high’ variety and had no more than a passive competence in the ‘low’ variety, while the number of bilinguals and of monolinguals acquainted only with the ‘low’ variety became smaller and

10 ‘Diglossia’ in the sense of Fishman (1965), in that the relation between the varieties involved is one between ‘sister’ languages rather than a relation between varieties, one of which is historically descended from the other.

smaller. This situation of 'diglossia in regression' is what is found nowadays in most of the linguistic areas mentioned above. What varies is the percentage of speakers who have been completely absorbed into the national language, which is high in France and Spain, but smaller in Italy. Asturian in Asturias and Aragonese in Aragon are protected by regional legislation, but this has yet to lead to a widening of their sphere of usage with accompanying change in their sociolinguistic status. In France and Switzerland, the Franco-Provençal varieties are now in the final throes of the linguistic conflict, having been almost entirely eliminated from speakers' linguistic repertoire. Things are little better in *oil* and Occitan varieties. The various movements promoting diffusion of these varieties in teaching, culture and the media paradoxically affect only some cultured élites who seek to reclaim once prestigious languages which have for centuries been seriously undermined. But they seem incapable of reversing a trend which has been operating for centuries. In Italy, the commonest situation is *dilalia* (Berruto 1995), whereby Italian (albeit usually in regionally marked forms) is used by part of the community in informal situations as well, at the expense of the dialects. The dividing line between the two codes and the numbers of those who speak only Italian (not dialect), those who are diglossic, and 'Italians who have no Italian' still vary considerably from region to region. In the south and north-east, use of the dialect is still strong, dialect maintaining some social prestige and penetrating into contexts normally typical of the 'high' variety. And it is not uncommon for young males (but not females) whose mother tongue is Italian, learned from their parents, to learn dialect in their teens on the street, from their peers. In most of the linguistic contexts considered, the weakening of local varieties has led, in return, to the rise of varieties which graft features of indigenous speech on to the structures of the national language: e.g., *francitan*, in the *oc* domain, *amestáu* in Asturias, *castrapo* in Galicia and the various kinds of 'regional Italian' (northern, central, Roman, southern, etc.).

Some regions with a strong linguistic identity saw the launch, from the late 1970s, of policies in support of the local varieties aimed at stopping or at any rate slowing down the encroachment of national languages. In Catalonia and Galicia, recognition of the official status of indigenous varieties, their standardization and their increasing public usage, are aimed at shifting speakers out of 'conflictual diglossia' (where one language is dominant and the other is dominated) towards 'bilingualism without diglossia' (where there is widespread competence in both systems, but no longer any functional distinction in their use) (Fernández Rei 2007:479). In Catalonia this process is at a fairly advanced stage, while in Galicia the distinction between 'high' Castilian

and 'low' Galician still appears entrenched. In recent years some linguistic minorities in Italy (Friulians, Ladins, Sardinians and, to a lesser degree, Franco-Provençal speakers) have sought to undo the established relation of subordination of local varieties to Italian (and also to German in the case of Ladin). The process is still in its infancy, and only received legislative support in 1999. Initially, these varieties had to develop a supra-dialectal standard language with which the entire community of speakers felt at home. This was a long and laborious process which sometimes met opposition, with the result that, hitherto, there has been only limited public use of these languages. Things are much the same in Corsica, where the measures adopted over the last thirty years to promote the use of local varieties seem to have made little headway against the dominance of French.

There are of course also diglossic situations where a Romance variety 'dominates' a non-Romance language: Spanish and Basque, French and Basque, French and Alsatian in Alsace, French and Lotharingian in Lorraine, French and Flemish in Nord-Pas de Calais, French and Breton in Brittany; Italian and Greek, Albanian or Croatian in the various linguistic enclaves of southern Italy; Romanian and Hungarian in the Banat and Transylvania, etc.

The reverse also occurs, typically in areas where a Romance language is subordinate to a non-Romance one: the Italo-Romance communities of Slovenia and Croatia, the Aromanians in Albania, Greece, Macedonia and Bulgaria, the Megleno-Romanians in Greece and Macedonia, the Romanians in Serbia, Hungary, Ukraine and Bulgaria. The same is true of the Ladin community of the Val Gardena, where the 'high' variety is German, although the population has a high level of competence in Italian. In the Romansh communities of Grisons, despite the official status of Romansh, *Schwytzerdütsch* (or in some cases German) is often used as a language of general communication and even as a 'high' variety. In Belgium, Switzerland, Alto Adige and the Republic of Moldova a monolingual Romance-speaking community lives beside a monolingual community speaking another language. Here, only a small part of the population is bilingual or trilingual, and in general the different languages are used without significant functional differences. In Belgium the francophone Walloons often do not know Flemish, except in Brussels, where bilingualism is favoured. A similar situation occurs in Italy in Alto Adige, where most people speak (Tyrolean) German or Italian (but not dialect), but are not bilingual. The German speakers have long had some knowledge of Italian, while the Italian speakers have begun to learn German better after bilingualism was made obligatory in public employment. In Moldova there was a situation of substantial diglossia during annexation to the USSR, with Russian as the 'high' language,

and Moldovan as the 'low' variety. After independence the use of Moldovan spread to all linguistic domains, but Russian remained the main language of the large Slav (Russian, Ukrainian and Bulgarian) and Turkish minorities.

5. Languages and migrations

Migrations become a major phenomenon by the late nineteenth century. Two factors may contribute to migrants' holding on to their original language: the group's size and their geographical and social isolation. The Sephardic Jews, who emigrated from Spain and Portugal from the late fifteenth century, until recently kept, and to some extent still do keep, their original language, a consequence of the closed nature of their communities. There are also cases of long-term conservation of the original languages in recent migrations, such as the emigrants from Trento (northern Italy) in Štivor, in Bosnia; the Friulian and Bellunese immigrants in Greci, in Romania; or the Venetians who have settled in Brazil. In every case what is involved are fairly homogeneous linguistic communities, living in rural areas. Elsewhere, in similar social and locational conditions, linguistic assimilation has been more rapid. In general, with migrations which took place in the nineteenth and twentieth centuries the original language is not preserved beyond the third generation (Corrà and Ursini 1998). So the main linguistic effect has been a decline in the numbers of speakers from among whom the migration originated. As for the host communities, the introduction of elements speaking other languages has often shifted the balance between national language and dialect in favour of the former, thereby contributing to linguistic homogenization in the region of immigration. In France, the arrival of Italians, Corsicans and people from the Maghreb in Provence, and particularly Marseille, has done much for the spread of French and the weakening of the local Occitan variety. Immigration from southern Spain to Barcelona, Valencia and Alicante favoured expansion of Castilian at the expense of Catalan. In Italy, immigration from the centre and south, Sardinia, the Veneto and Friuli, has been one of the main causes of the decline in the use of dialect in the towns (and in the countryside) of the north-west.

Romance-speaking Europe played the major part in nineteenth- and twentieth-century migrations. Initially, emigration was across the Atlantic. From the 1920s, emigration within Europe begins to prevail. The country most clearly affected was Italy, which saw nearly 26 million people leave between 1876 and 1976, particularly to France, Switzerland, Germany, Benelux and the Americas. Most Italian emigrants were dialect speakers with very limited

knowledge of Italian. They came mainly from southern Italy, the Veneto and Friuli. On the whole they rapidly lost their native dialect. Portugal lost over four million people between 1801 and 1996, until the early nineteenth century mainly to Brazil and Africa (Angola, Mozambique), then from the mid century towards the United States and Spanish-speaking Latin America, and particularly from the 1950s towards Europe and especially France (about 1.5 million). Over four million Spaniards emigrated between 1880 and 1970, initially to Latin America. From the 1920s, and even more so after the Spanish Civil War, Spanish migration was mainly to France. Internal migration was very important. In France it was from south to north. In Italy, until the 1960s, migrants moved from the south, the Veneto and Friuli to Lombardy, Piedmont and Liguria. Only in relatively recent years has there begun to be migration into the regions of the north-east (particularly the Veneto). In Spain, particularly in the 1960s and 1970s, there was a major movement from the south to the region of Madrid and to Catalonia.

In Italy, Portugal and Spain this type of emigration, driven by economic necessity, has now come to a halt, and these countries have in turn started to receive immigrants. From the 1990s, major migratory movements from central and eastern Europe have begun. Of these, the only Romance-speaking migrants are the Romanians and the Moldovans.

6. Linguistic geography

In the major historical grammars (from Raynouard to Diez, Meyer-Lübke, Bourciez, Lausberg, etc.), the number of Romance languages systematically treated is never above six, seven, eight or nine, although there are references to other varieties, especially those with literary traditions. The need was soon felt to take other Romance varieties into account, including those which were not written and not prestigious. These were what were called 'dialects', pejoratively known in French as *patois*, or 'socially degraded' varieties (Dauzat 1946). Groundbreaking work on these new linguistic varieties was done by Graziadio Isaia Ascoli on the 'Ladin' area, and adjoining territories, and 'Franco-Provençal' (the terms were invented by Ascoli 1873; 1878). Theodor Gartner (1879; 1883; 1910; 1923) investigated what Ascoli had called the 'Ladin' area, but for which he used the infelicitous term 'Raeto-Romance'.

The idea of representing dialect forms cartographically was still to come. It appears first in Germany, with Georg Wenker, but is first put into effect in France with Jules Gilliéron, in particular with the *Atlas linguistique de la France*, which appeared between 1902 and 1912 (the work also covered Swiss

Romandy and Francophone Belgium). This brought from the outset an extraordinary quantity and variety of data, especially phonetic and lexical. A linguistic atlas consists of maps, devoted to some 'concept' (which can be an object but also an event indicated by a verb or a noun, the name of a period of time, such as the days of the week, months, holidays, etc.), each of which has a number of points representing localities within a given territory. Alongside these points there appear, in phonetic transcription, the terms for the relevant concept in each locality. The data are gathered by an investigator who has travelled to the various places and questioned a native speaker using a previously drawn-up questionnaire. As a rule, supplementary material, bibliographical references, comments, etc., appear at the side of the map or on a separate page. This is obviously an extremely complex operation, which can take years if not decades to produce. It is also very expensive. No surprise then if many works of this type never reach completion.

The results made visible by the *ALF*, and confirmed by all the subsequent atlases for other Romance varieties, showed that the word for a concept changed from place to place, and that neighbouring villages almost always showed the same words in a different form, or even totally different lexical types. Where grammars, both traditional and historical, gave an ordered presentation of languages, the linguistic atlases showed that language, viewed geographically, is so extraordinarily varied as to appear at first sight chaotic.

But from the outset work by Gilliéron and his school (Gilliéron and Mongin 1905; Gilliéron and Roques 1912; Gilliéron 1918, etc.), Karl Jaberg (1908) on French territory, and later, among others, that of Benvenuto Terracini (especially on Sardinian, 1964), sought a hidden order within this variety, an approach which continues to this day. Gilliéron's school developed a 'stratigraphic' mode of interpretation of the linguistic maps, aiming to modify the then fashionable linguistic model of the Neogrammarians, based on sound laws. From his observation of the frequent replacements of lexical types in the Gallo-Romance domain, Gilliéron developed the concepts of 'phonetic erosion', 'homonymic clash' and 'polysemy', against which speakers react 'therapeutically', showing the 'reactivity' which is an expression of the 'vitality' of the dialects (cf. Grassi 2001).

Having claimed that these mechanisms were what drove language change, and using maps to reconstruct how innovations radiated from some points of innovation to be accepted by others (with the result that older forms get relegated to the remote periphery), Gilliéron, Jaberg, and later Terracini, proclaimed that sound laws had 'failed'. Jaberg (1908) offers an early synthesis of this point of view. For Jaberg, linguistic geography allows us clearly to

observe the advance of prestige forms and their territorial encroachment on less prestigious forms. So each dialect contains new and old forms, sometimes home-grown, sometimes taken from more prestigious centres. Thus, for Jaberg, every dialect is mixed, rather as Hugo Schuchardt (1884b:6) had said that every language is mixed. These claims strike us today as suggestive but obscure, for, coming as they did before Saussure's distinction between diachrony and synchrony, we are unsure to which dimension they refer. At any rate, for these scholars the 'genealogical' conception of languages becomes secondary, whilst every dialect (or language) is subject to innovatory impulses from without and to internal forces of direct, or often indirect, resistance. Where there is 'indirect' resistance, a dialect reacts to an external innovation neither by accepting nor rejecting it, but by modifying its existing form or creating a new one ('linguistic creativity'). In this light, sound laws with their alleged regularity seemed to Jaberg a mere abstraction. In fact, he wrote, every word 'has its own particular history [...] the sound law is an abstraction' (Jaberg 1959:20; quoted also in Varvaro 1968:210).

This approach was favoured in the early nineteenth century by contemporary neo-idealistic theories, mainly inspired by the philosophy of Benedetto Croce and represented in Germany by the Romanist Karl Voßler. These theories aimed to move linguistics away from positivism and its narrow naturalistic methods and into the realm of the psychological sciences. Corrado Grassi (2001) still credits this discipline with having put speakers, rather than the abstraction of language (by now labelled in Saussure's sense as 'langue'), at the centre of observation.

The idea of the 'failure of phonetic laws' led some linguists to go so far as to believe that they could best represent the characteristics of the Romance languages through their lexicon, which emerged so clearly from the linguistic maps, rather than through sound laws. Some individual maps served this end particularly well: for example, the concept 'head' as represented by continuants of Latin *CAPUT* and *TESTA* (a metaphor from *TESTUM*, *TESTA* 'pot'), whence, e.g., Ro. *cap* but Fr. *tête*, or continuants of *AUUNCULUS* 'uncle' (whence, e.g., Fr. *oncle*, Ro. *unchi*) and the Greek loan *θειος* 'uncle' (whence, e.g., Pt. *tio*, Sp. *tio*, It. *zio*, Srd. *tio*, *zio*) and others like *BARBA* 'beard' > Nlt. *barba*, etc. (see further the hundreds of maps in Rohlf's 1954; 1971; 1986). This type of representation looked at first like a static statement of the facts, but actually lent itself to a more interesting, dynamic, interpretation, such that the oldest form was identified and the advance of rival forms could then be observed or hypothesized. Each map gave a different result, but certain patterns occurred more often, indicating specific directions of lexical (and cultural) expansion, and

their limits. This type of lexical observation (and the application of the same approach to phonetic and morphological facts) gave rise to the so-called 'areal norms' of Matteo Bartoli discussed briefly below.

This polemical stance with regard to phonetic laws, and in fact the whole of historical linguistics, certainly provoked reactions. Dauzat (1946:76) felt that the elimination of sound laws was a 'grave error'. Indeed, it has often been observed that the very linguists who condemned sound laws often made use of them, and correctly. At bottom they did not want to appear not to have moved on from positivism, but they could not simply abandon the only available tools of the trade. Those forged by Gilliéron have proved of much more limited use. These involved such criteria as: monosyllabism as the possible origin of the disappearance of certain words, which were already phonetically 'mutilated' (as Gilliéron said) in Latin, and further subject in Romance languages to the corrosive effects of sound laws; homonymic clash as the source of the elimination of lexemes; the 'therapies' resorted to in order to renew the lexicon and compensate for such eliminations – the use of suffixes, especially diminutives (such as *SOLICULUM* for *SOL* 'sun' in Fr. *soleil* and Frl. *sorèli*), loans (which were also diminutives: e.g., Fr. *abeille* from Occ. *abelho* < *APICULA* for *APIS*, 'bee'), metonymic substitutions (Fr. *essaim* 'swarm' for 'bee'; in Italy the type *vespa* also 'bee'), metaphoric substitutions (Fr. *vicaire* 'vicar, priest' for 'cockerel'), periphrases (Fr. *mouche à miel* lit. 'honey fly', for 'bee'); folk etymology as the source of irregularities in sound change (Fr. *fumier* for *fimier* < *FIMARIUM* 'dung heap' + *fumer* 'smoke'). A number of changes conditioned by homonymy and phonetic reduction for the entire Romance domain are given in Rohlfs (1971: ch. 19 and 20).

Thus linguistic geography seems today to have come to the end of its adventure as the bearer of the message we have tried briefly to describe above. No one speaks any longer of idealism in linguistics. Sound laws, which this movement overshadowed, or simply denied, have been wholly rehabilitated and accepted in subsequent directions taken by linguistics, e.g., in structuralism (especially in its very interesting engagements with diachronic phonology: Martinet, Haudricourt and Juilland, Lausberg, Weinrich, etc.), in Labov's socio-linguistics, and, more implicitly, also in generative phonology, etc.

This is not to say that linguistic geography never had and never will have any use. Closely linked to linguistic geography was the development of onomasiology. Moreover, from the beginning, linguistic maps brought to light interesting phenomena which have become accepted wisdom in linguistics and dialectology. One of these concerns the fact that boundaries can be drawn between different varieties. Thanks to geographical observation it was possible

to establish to a higher degree of complexity, and of precision, than before how many linguistic forms there are in a given territory and what their boundaries are (see discussion of isoglosses below in §6.3).

6.1 *Crises and developments*

In the early 1960s, in the light of Saussurean and post-Saussurean linguistics (structuralism, generative grammar, etc.), the methods, techniques and achievements of linguistic geography seemed to many to be outdated. This was the ‘crisis’ of linguistic geography. The neighbouring discipline of dialectology was also in crisis. And so, more generally, and healthily, was linguistics itself. Indeed, linguistics was revising its entire methodology. Language, considered synchronically, is now viewed as a system which only accepts new elements by reorganizing itself. Linguistic geography did not grasp this, focusing on one phenomenon at a time, and thereby failing to see the whole. A phonological study such as André Martinet’s (1945; 1956) on the Franco-Provençal dialect of Hauteville showed that, despite external influences, the phonology of a dialect can be represented as a system. Uriel Weinreich (1954), an American pupil of Martinet, proposed a structuralist reform of traditional dialectology, doubly linked to linguistic geography. Weinreich’s question: ‘Is a structural dialectology possible?’, supposed the answer ‘yes’, and he was already beginning, with characteristic brilliance, to show why.

A different crisis came from outside. Linguistic geography, often together with ethnography, had been concerned, sometimes exclusively, with the world of dialects and rural life and work. After the Second World War, and ever faster as the decades went by, the rural world was disappearing in western Europe as industrialization, and then the service sector, spread. This was accompanied by a sometimes spectacular retreat in the use of dialects (while agricultural tools ended up rusting in the corners of barns and museums of peasant life). Many felt that the very object of study of linguistic geography was in need of a new focus. North American and British sociolinguistics, which mainly studied urban language, appeared both a dangerous competitor and a potentially fertile model for traditional linguistic geography to follow, perhaps the answer to the crisis.

This crisis triggered numerous debates, in which sometimes radically different positive solutions were proposed (see Thun and Radtke 1996; the studies in Ruffino 1992 and 1998; García Mouton 1994; Englebert *et al.* 1998). New technology, especially computers, offered one type of innovative solution. In the midst of the crisis, paradoxically, many projects continued, and others started up, still on traditional lines, while others looked for new ways forward.

Users of linguistic atlases find that they can hear, as well as read, the dialect forms noted. This is the case, for example, with the *Atlant linguistisch dl ladin dolomitich y di dialec vejins* [Linguistic Atlas of Dolomitic Ladin and neighbouring dialects], the work of a team led by Hans Goebel (1998), which has not only maps but also an 'atlante linguistico sonoro' ['linguistic sound atlas'] (now as a DVD, see Bauer *et al.* 2005). A recent project, in the Catalan domain, is that of Maria-Pilar Perea and Germà Colón, who have produced a 'chronostratigraphic' dialect atlas of Catalonia which, exploiting the possibilities offered by computers, shows not only variation through space but also through time, or at any rate their earliest attestations from the time of the earliest texts (*Cronoestratigrafia dialectal* in CD-ROM).

There are even more radical changes to the traditional atlas. It may become a database, which transcends the one-dimensional perspective of maps and gathers dialectological and sociolinguistic data for individual points and speakers; it may contain texts, sound recordings and so forth. This means that it has the advantage (if advantage it be) of being something open-ended that can be built up by additions and updates over the years and decades.

6.2 *Atlases and linguistic maps of the România*

All the major Romance linguistic areas of Europe are today covered by major linguistic atlases. Some have long been completed, others are still in progress, and are often accompanied by monographic studies, collections of texts, etc., while many (regional) atlases give finely detailed coverage of more limited areas, major islands, etc. Only the briefest account of some of these can be given here; see further Thun and Radtke (1996), Ruffino (1992), Winkelmann (1993), García Mouton (1994) and Cugno and Massobrio (2010).

Another approach is to have a, normally large, map which brings together, for a particular territory, the findings of other studies. Thus there are maps for the Iberian Peninsula, or the Gallo-Romance, or Italo-Romance, or Romanian areas, etc., subdivided into the main varieties.

Giovan Battista Pellegrini's *Carta dei dialetti d'Italia* (1977), comprises a large-size (119 × 95 cm) map representing the dialects and languages of the state of Italy (thus including Franco-Provençal, Occitan, German, Albanian, etc.), distinguished by colour. The boundaries between the different varieties are represented by changes in colour, but also marked by isoglosses (see below). The various linguistic areas are precisely identified. Unlike most linguistic maps, which are often merely convenient didactic simplifications, the *Carta* offers a wealth of information. Thus, if one takes the name of one of the localities represented, one can immediately find which variety (or indeed

subvariety) it belongs to, or whether it is in one of the few mixed areas, if it falls within one of the isoglosses traced by the author, and so forth. This system of representation is ideal for representing the complex linguistic system of Italy.

6.3 *Isoglosses and linguistic boundaries*

As we have already suggested, among the problems tackled by linguistic geography from its inception is that of linguistic boundaries. To take an example, as is evident from Pellegrini's *Carta dei dialetti d'Italia*, the boundaries, in the Alps, between Ladin and Tyrolean, and between Lombard and Tyrolean – in other words, between Romance and Germanic varieties – are sharply defined. But this is equally true of the boundary between two genetically related varieties. A case in point, again from the Alps, is the frontier between Ladin and Friulian, and also, in a small contact area, between Ladin and Lombard. Such cases may be favoured, or even created, by the mountainous terrain (villages clustered in valleys, impenetrable and uninhabited swathes of territory, etc.), which make contact difficult. Ladin, Friulian and Lombard are local continuations of Latin. There are no islets of Romance varieties one within another as there are, throughout this Alpine area, islands of German, the result, precisely, of colonization. But there are very different cases in the Romance world as well: over very wide strips of territory villages inhabited by speakers of one language alternate with villages speaking another variety over and over again. This is what happens, for example, in Romania (Transylvania) between Hungarian and Romanian, and between Ukrainian and Romanian in the north-east of the country.

But while adjacent varieties may have the same origin and show linguistic similarities, the transition is more usually gradual. In such cases the boundary cannot be represented as one line, but as many lines. These are 'isoglosses', lines joining geographical points which show the same linguistic phenomena. Given the phenomena (usually phonetic, morphological, or lexical) *a, b, c, d, e* and *a', b', c', d', e'* we have not only varieties which possess all the phenomena of the former group, or all those of the latter, but also varieties that have *a, b, c, d* but *e'*, or *a, b, c*, but *d'* and *e'*, and so forth. If we join the geographical points showing *a*, those showing *b*, etc. and then those that have *a', b'* etc. we get not lines, but series of lines, sometimes (almost) parallel, sometimes intersecting, and these are 'isoglosses'. The phenomena selected must of course be suitable to characterize two neighbouring varieties, as we shall see shortly. Romance linguists have long established that the transition between related varieties is mainly a 'bundle of isoglosses'.

This principle is confirmed by relatively recent detailed studies, such as Francescato's (1966) treatment of the Friulian–Venetan boundary, dealing with subdivisions within Friulian, or Contini's (1987) study of the internal divisions of Sardinian, or Cano González's (1992) study on the boundary between Asturian and neighbouring Ibero-Romance varieties, or the internal divisions of Asturian.

Francescato describes Friulian with forty-four phonetic and morphological features and some lexical examples. These features often show variation within all Friulian subvarieties. In the western area, at the boundary with Venetan, some of these features may give way to characteristically Venetan features. Thus:

- the definite article *el, al* of the Venetan type replaces Frl. *il*;
- the subject pronouns *mi, ti, lu* of the Venetan type may replace Frl. *jo, tu, al*;
- the distinction between the so-called 'strong position' and 'weak position' of stressed vowels¹¹ may be lost;
- final *-t* may be deleted, following the Venetan model: thus *mercà* for Frl. *mercat* 'market', *ferì* for Frl. *ferit* 'wounded';
- the epithetic consonant after a nasal characteristic of Friulian may be absent: e.g., Frl. *omp* 'man' may be replaced by Vto. *oj*;
- original palatals before front vowels (Frl. *θent* 'too') may become *s* or *θ* (*sent/θent*), as in Venetan;
- intervocalic *d* may weaken to *ð*, as in Venetan: e.g., *ruoða* for *ruode* 'wheel';
- the feminine plural may change from *-is* or *-es* to *-i* or to *-e*, losing the *-s* ending typical of Friulian: *džambi, sarièzi* 'legs', 'cherries' (Gruaro); *vorèle, òngole* 'ears', 'nails' (Porcia), etc.¹²

No Friulian dialect loses other typically Friulian traits such as palatalization of /k/ and /g/ before central vowels (*k'ase* or *t'ase* vs. Vto. *kaza*, or the plural in *-s* of masculines like *t'an : t'ans* 'dog, dogs', *ouf, oufs* 'egg, eggs', etc. There

¹¹ A stressed vowel is in 'strong position' if in Latin it was in an open syllable and in Friulian is in a closed final syllable, as a consequence of the regular deletion of final unstressed vowels other than /a/. In this position the Friulian vowel is lengthened: *GOLŌSUM* > *golo:s* 'glutton' (in western varieties *golōus*); *CANTĀTUM* > *canta:t* 'sung', *ŌUUM* > *u:f* 'egg' (in western varieties *ouf*). In all other cases the stressed vowel is in 'weak position', for example, as a continuation of a vowel in a closed syllable in Latin (*FRŪCTUM* > *frut* 'fruit', *ŌSSUM* > *wes* 'bone') or of a vowel in a Friulian open syllable (*golóze, k'antade, kusi, dula*, etc.).

¹² On western Friulian and on the Friulian–Venetan boundary; see also Lüdtke (1957) and Vanelli (2005). Compare also the *ASLEF* (1972–86) directed by Giovan Battista Pellegrini.

seem, then, to be features that Friulian cannot lose, even in the most heavily Venetized varieties. So, the passage between Friulian and Venetan is represented by a swathe of Friulian dialects with varying degrees of Venetization. This includes centres such as Aviano, Polcenigo, Fontanafredda, Fiume Veneto and Azzano, which have some of the Venetan features listed above, and others we have not mentioned. These dialects are thus touched by some of the isoglosses representing Venetan phenomena. This bundle of isoglosses defines a transition area between Friulian and Venetan.

It needs to be added that there are no Venetan dialects in contact with Friulian showing traces of Friulianization. The transition between Venetan and Friulian is due to a bundle of Venetan isoglosses encroaching on Friulian territory, while the territory of Venetan remains compact. This is a result of the greater prestige enjoyed for centuries by Venetan over Friulian. Venetan was spoken by members of the Venetian aristocracy and bourgeoisie, including those settled in Friuli, and managed to impose itself in various towns in Friuli (beginning with Udine) alongside Friulian. In some places it even replaced Friulian completely (as happened to the west in Pordenone, within the above-mentioned bundle of isoglosses, and at the eastern edge in Trieste, which was originally Friulian). As mentioned above, this hegemony of Venetan over Friulian seems to have gone unchallenged until recent times. Today, however, as Italian spreads, Venetan is decidedly in retreat, and may actually disappear in some localities, giving way to a diglossic pair comprising Italian and Friulian. In the historical picture drawn above, however, the western boundary of Friulian has gradually been shifting ever eastwards over the centuries with the result that, as noted, a town like Pordenone, situated approximately in the centre of the transition zone, has long gone over from being Friulian to Venetan.

We see, then, that the boundary between Friulian and Venetan cannot be represented just by the bundle of isoglosses described, but also comprises a series of localities into which Venetan has, as it were, been ‘parachuted’ – introduced without direct adjacency. This situation is perfectly common in other situations as well, and numerous other examples can be given. In the eastern Alpine area, again, in territory in which Ladin, Venetan and Trentino are compactly present (as also further west in Piedmont and Lombardy), there are German (Tyrolean) linguistic islands. These are small medieval colonies in Romance territory, although many more Germanic colonies are more recent, such as the so-called Saxons and Swabians in Transylvania. Rome is a kind of Tuscan-speaking island in Lazio, due to the ancient Tuscanization of Rome, from an influx of inhabitants from southern Tuscany in the sixteenth century

(see Ernst 1970). These are 'linguistic islands', which may consist of a large or small nucleus of inhabitants, as in the cases discussed above, or of more or less extensive bands of territory.

Let us now take an example from the Iberian domain. In north-west Spain, in Cantabria and the Principality of Asturias, we find five linguistic boundaries in quick succession. The westernmost marks the transition between Galician-Asturian (or *extremeiro*) and Asturian (or Asturo-Leonese) and substantially coincides with the course of the river Navia, in western Asturias. The easternmost marks the transition between Asturian and Castilian varieties, and includes the western third of Cantabria, extending north-eastward along the coast, as far as Santander. Asturian is in turn divided into three. Since Menéndez Pidal (1962), a western, a central and an eastern variety have been distinguished. The boundaries continue southwards into the provinces of León, Zamora and Salamanca, in Castilla y León. For simplicity's sake, we consider just the northernmost part of the Asturo-Leonese domain. In tracing the five linguistic boundaries, dialectologists have mainly considered the following phenomena (Menéndez Pidal 1962; Zamora Vicente 1967; Cano González 1992):

- (a) diphthongization of Latin *ĕ* and *ō* in open and closed syllables (*FĔRRUM* > Ast. *fierru* 'iron', *IŌCUM* > *xuegu* 'game') (cf. Loporcaro, volume I, chapter 3, §1.2.1);
- (b) reduction of the falling diphthongs /ej/ and /ow/ from *a* + yod (*CORDARIUM* > *cordeiro* > *cordero* 'lamb') and from /aw/ (*CAUSAM* > *cousa* > *cosa* 'thing'); (cf. Loporcaro, volume I, chapter 3, §1.2.1)
- (c) development of initial /f/ to a velar or glottal fricative (/x/ or /h/): *FACERE* > *hacer* 'to do';
- (d) palatalization of initial /l/ (e.g., *LUNAM* > Ast. *lluna* 'moon', *LARGUM* > *llargo* 'long', etc.);
- (e) passage of the inflectional endings *-as/-an* to *-es/-en*: *CASAS* > C.Ast. *cases* 'houses', *CANTAN(T)* > *canten* 'they sing'.

The boundary between Galician and Asturian is identified with isogloss (a). Phenomenon (b) sets Galician and western Asturian off from the other Asturian varieties, and from Spanish varieties. Isogloss (c) basically follows the river Sella and separates Galician and western and central Asturian from eastern (and Cantabrian) Asturian and Castilian. In Castilian, the velar or laryngeal fricative (/x/ or /h/) was present in the Middle Ages but subsequently fell. It survives today in southern Castilian and Latin American dialects. Phenomenon (d) is found throughout the Asturo-Leonese domain and is one of the features

used to distinguish the Asturian type from the Castilian. To the west it also reaches as far as the eastern fringes of western Asturian, so that the boundary does not exactly match isogloss (a). Development (e) isolates central Asturian from the eastern and western varieties, as well as from Galician-Portuguese and Castilian. Its western and eastern limits do not exactly coincide with isoglosses (b) and (c): in the south-west of the central area we have a zone which retains *-as/-an*, and in the western part of the eastern Asturian domain is an area where *-as/-an* evolves into *-es/-en*. We synthesize the distribution of phenomena (a)–(e) as follows:

phenomenon	(a)	(b)	(c)	(d)	(e)
absent	Glc.	Glc. W.Ast.	Glc. W.Ast. C.Ast.	Glc. (-E) Cst.	Glc. W.Ast. E.Ast. (-W)/Ctb. Cst.
present	W.Ast. C.Ast. E.Ast./Ctb. Cst.	C.Ast. E.Ast./Ctb. Cst.	E.Ast./Ctb. Cst.	W.Ast. C.Ast. E.Ast./Ctb.	C.Ast. (-SW)

None of the isoglosses in question, which are a fraction of what one could mention, coincides with the others. Here too we have a bundle of isoglosses. All the varieties considered (themselves the result of abstraction and simplification) show areas of overlap or transition with regard to specific phenomena. In addition, the spatial distribution of (a)–(c) seems different from that of (d)–(e). While the isoglosses of (a)–(c) mark the gradual passage from the Galician to the Castilian type and may be viewed as reproducing the spread of innovations from the east (Castile) to the west (Asturias), phenomena (d) and (e) have their seat in central Asturias, and seem to have radiated thence both westwards and eastwards.

All our observations until now have been, so to speak, one-dimensional, in that they are apt for representing just the ‘original’ linguistic variety of the localities in question. It is not possible, however, at least with traditional cartographical techniques, to represent diglossic, triglossic or bilingual situations, or code-switching. These are things in which modern sociolinguistics has, however, been extremely interested. Regarding the Venetan–Friulian boundary, though, Francescato (1966) was already observing that diglossia is frequent in the Friulian–Venetan transition zone: in addition to their Venetized Friulian (*furlán*), speakers speak Venetan (*dialecto*). Thus many

speakers are diglossic, or indeed triglossic (Friulian, Venetan and Italian) in Udine and other towns of Friuli. In the Spanish example given above, the overwhelming majority of speakers of non-Castilian varieties also speak Spanish as their 'high' variety.

In the case of French *patois*, presentations of the data in map form subsequent to Gilliéron tell us about the transformation of the speech varieties of France from that time on, but they do not normally give information about numbers of speakers; they do not tell us whether they are diglossic, nor in what contexts they use *patois*.

This fact does not diminish the importance of the study of geographical linguistic boundaries, nor that of the use of isoglosses as, for example, Avolio (2001) rightly stresses. Isoglosses remain the only efficient way of delimiting, for example, the three fundamental dialectal areas of Italy: the north, Tuscan (to the south-west) and central-southern (to the south-east). This is what the so-called 'La Spezia–Rimini' (or 'Massa–Senigallia') line does. Actually this is not *one* line but a bundle of isoglosses, principally phonetic, which diverge mainly at the ends, giving rise to three fairly large intermediate areas (Rohlf 1937, map 2;¹³ Pellegrini 1971). To the north-west, for example, there is an area where northern weakening of intervocalic stops is absent, but in which the Tuscan phenomenon of 'gorgia' (cf. Loporcaro, volume I, chapter 3, §2.2) does not appear, the relevant isogloss appearing rather further south. In the northern Alpine area of Italy, the isogloss indicating the presence of /y/ identifies mixed Trentino–Lombard varieties.

A variety is not necessarily mixed by virtue of lying within a bundle of isoglosses: in the *Carta dei dialetti d'Italia* there are, for example, no Emiliano–Tuscan or Romagnolo–Marchigiano varieties, despite the isoglosses of which we have just spoken. This means, for example, that a variety like that of Fermo is 'northern', even though it does not show lenition of intervocalic stops (a central-southern isogloss which extends further to the north).

Let us now consider as our third and final case Sardinia. Michel Contini (1987) studied Sardinian dialect varieties through sixty-one phonetic characteristics, tracing a corresponding number of isoglosses on maps of Sardinia. Given the extremely fragmentary nature of Sardinian varieties, many of these phonetic

13 The same map appears in Tagliavini (1972:§68). The most traditional and established denomination takes its name from La Spezia, the last clearly northern town to the north-west before the Tuscan area, and from Rimini, the last Romagnol city of the north to the south-east. A rival name takes Massa as the first point to present most of the Tuscan phenomena to the north-west, and Senigallia as the first clearly central-southern point to the south-east.

isoglosses more or less go their own way, but others form bundles which Contini takes as internal boundaries (often reflecting the major mountain chains of the island). These he uses to group Sardinian dialects into five areas, which partly follow and partly diverge from the traditional divisions of Sardinian into dialects:

- 1) Sassarese-Gallurese
- 2) Central-eastern
- 3) Marghine
- 4) Logudorese
- 5) Campidanese

In this subdivision, the use of isoglosses and their grouping into bundles allow fine and extremely detailed distinctions to be made between local varieties, yet still offer considerable scope for synthesis of the results.

A further frequent objection to isoglosses is that the choice of the phenomena they represent is arbitrary. In fact, although bad choices may sometimes be made, any expert dialectologist will certainly be equal to the task of assessing the relative importance of the phenomena examined (see also Contini 1987, I:516f.). Linguistic geography aside, some morphological and phonological phenomena have long been considered by linguists (and sometimes by speakers themselves) as essential criteria for characterizing one Romance variety with regard to another. This is not to say that one cannot descry new features, which may lead one to set up new isoglosses, or that any particular feature cannot be assessed in a new light. Problems of this kind are common to other domains of scientific enquiry, and do not justify doubts about the usefulness of the geographical representation of linguistic variety or the value of isoglosses.

7. Areal linguistics

We have seen that the major calling of linguistic geography in its initial and most successful period was to immerse itself in the great variety of language, to seek its direction in that variety, but without staking its ground exclusively in rigid laws and generalizations. In fact the most daring and advanced work was carried out under the banner of fighting against generalizations, especially of the kind most prevalent at that time – phonetic laws.

This does not mean that there were no attempts to find regularities amid the sea of varieties. For Terracini (1957), crucial to the interpretation of continual change was the dialectic between centres of prestige and socially subordinate areas: the former radiate changes, the latter sometimes accept prestige forms,

sometimes react in creating forms of their own. This means that varieties lacking prestige are not always passive.

The 'areal norms' of Matteo Bartoli seek to explain the distribution of forms in geographical space (Bartoli and Vidossi 1923; Bartoli 1945b). The term 'norm' was intended to suggest that what was involved were tendencies, not hard-and-fast rules, but this did not save Bartoli from much criticism on the grounds of the easily found exceptions to his norms. Gilliéron (1918) in his atlas had noted an apparently paradoxical fact: similar forms expressing the same concept are not always found in geographically neighbouring points, but may be far apart, and separated by whole zones representing other lexical types. An example is the forms *éf, é, és* (< Lat. *APEM*) 'bee', encountered in points far removed from each other in northern France and Switzerland, while the form for 'bee' in the vast intervening area came from Latin *APICULAM* lit. 'little bee' (Occ. *abeio*, Fr. *abeille*) or was *avette* (with diminutive endings) or was a lexical innovation, such as *mouche à miel* lit. 'honey fly', etc. The scarcity of reflexes of *APEM* was due to erosion caused by sound laws, which had reduced the form to a minimal phonetic content (a vowel and consonant, or just a vowel) and to homonymic clash with the continuator of *AUEM* 'bird'. The answer to the question of why these residual forms were so far apart was illuminating: the points showing reflexes of *APEM* were simply the residue of a once compact area which had escaped replacement by innovatory forms. These conservative forms survived at the periphery of French territory precisely because they had not been reached by replacements which had originated in dynamic centres of innovation. A similar explanation is available for the survival of Latin *ka-* in northern (Picard and Norman) and southern (Occitan) Gallo-Romance, while the whole intermediate area develops an innovatory palatal (and subsequent affricate in Franco-Provençal): Lat. *CAMPUM* 'field' > Pic. *kā*, Prv. *camp*, vs. Fr. *champ*, FPr. *tsā*.

On the basis of this type of observation, Bartoli elaborated one of his four areal norms (the norm of lateral areas), according to which, if geographically distant peripheral areas present one form and the central area presents another, the former is the older. Another such example, given by Bartoli, involves the presence of the lexical type *ROGARE* 'ask' in lateral areas of the *România* (Sp., Pt. *rogar*, Ro. *a ruga*), while the type *PRECARI*/*PRECARE* occurs in the central area (Fr. *prier*, It. *pregare*). Now the type *ROGARE* is the older, and original, form in Latin, and its replacement with *PRECARI*/*E* (originally, 'to pray') is later.

The value of this 'norm', as of the others (those of the 'isolated area', the 'greater area' and the 'later area'), has been much debated. Bartoli's norms

have been criticized and rejected by many scholars. Their predictive value is admittedly relative: it would be rash to state that a form found in lateral areas is perforce older than one found in central areas. But neither does this mean that this norm is wholly without value.

The same merits and limitations are present in the norm of the isolated area, which according to Bartoli generally preserves more archaic forms. An 'isolated area' is one less exposed to communication. This is the case with Sardinian, which preserves, for example, *MAGNUM* 'big' (Srd. *mannu*), against the innovative Romance form from *GRANDEM*. But Varvaro (1983) rightly stresses that Bartoli's norms presuppose a static notion of isolation, while an area can be isolated for centuries and then cease to be so, or the reverse. Varvaro shows with rich historical documentation that Lucania, held to be an isolated area *par excellence* (whence its special, 'Sardinian', vowel system; see Loporcaro, volume I, chapter 3, §1.1), was very accessible in the early and late Middle Ages, and became less and less so until the nineteenth century.

As we observed earlier, with sound laws thus pushed to the sidelines, if not wholly removed from the scene, the lexicon is at the centre of attention in geolinguistics. Another example is the work of Gerhard Rohlfs, who had gathered data for the AIS for southern Italy. The large number of mainly lexical maps he presents in Rohlfs (1971) are extremely valuable and the conclusion he reaches in his chapter 23 still seems valid. Consider the twenty-three lexical cases where Rohlfs gives three successive strata of Latin such as *cras*, *mane*, *de mane* 'tomorrow', or *malum*, *melum*, *pomum* 'apple', given in order of chronological progression from the most archaic to most innovative (Rohlfs 1971:206). Sometimes the first form is conserved only in some varieties, as with *CRAS*, continued only in Sardinian (*cras*) and a small area of southern Italy (*crai*), while Romanian continues the more recent *MANE* (*mâine*), and elsewhere one encounters the more recent *DE MANE* (Fr. *demain*, It. *domani*) and (*HORA*) **MANEANA* (Pt. *amanhã*, Sp. *mañana*). At other times the oldest form has no Romance continuants, such as *EDĒRE*, while the subsequent reinforced form *CUM-EDĒRE* is found in Iberia (Sp., Pt. *comer*) and the most recent form *MANDUCĀRE* in the Gallo-Romance, Italian and Romanian areas (Fr. *manger*, OIt. *manicare*, Ro. *a mânca*) (see also Stefenelli, volume I, chapter 11).

Some interesting generalizations emerge. Languages which conserve the first, most archaic, Latin stratum (or the second, if the first is not continued) are Sardinian, central and southern Italo-Romance (over a variable domain, sometimes just the 'Lausberg Zone'; see Loporcaro, volume I, chapter 3) and Ibero-Romance (but not Catalan, which definitely belongs with Occitan). If there is a conservative pair, it is almost always Ibero-Romance and Romanian,

the two lateral areas identified by Bartoli (indeed the relevant examples are those proposed by Bartoli himself). The first and second strata are never represented by Tuscan Italian or French, which are innovative. Rohlfs's conclusion (1971:206) that the data confirm the dictum that 'each word has its own history' is unjustifiably defeatist. The innovatory nature of the Gallo- and Italo-Romance areas, and the relative conservatism of Ibero-Romance and Romanian are data in harmony with the two major observations that Rohlfs himself had set out earlier: (i) that the source of diffusion of innovations was first Italy (Rome) and then France, western centres of political, cultural and linguistic influence; and (ii) that the power of diffusion gradually waned in an empire and a Europe undergoing political, social and economic fragmentation. Rohlfs notes that it is around the sixth century that France supplants Italy as a centre of diffusion of innovations into neighbouring countries (including Italy itself, and to a slightly lesser extent the Iberian Peninsula, while the future Romanian area is now out of reach; Rohlfs 1971: ch. 14). These data from Rohlfs do not divide the Romània into clear-cut areas, but nor are they negligible.

Lorenzo Renzi (1976; Renzi and Andreose 2006) reaches partly similar conclusions to those which emerge from Rohlfs. Renzi attempts to group Romance languages together in terms of the presence or absence of certain syntactic or morphological phenomena and their conservative or innovative nature. The unprecedented decision to observe concordances and discordances among Romance languages at the syntactic and morphological, rather than the phonetic and lexical, level was inspired by the predominant place of syntax in the newly ascendant generative grammar, but also by the notion that there is a lesser degree of variability at these levels, and that the problems of attempting to draw conclusions from an excessively large and intractable mass of material would be avoided. The lexicon comprises an enormous number of units, indeed it is virtually infinite, and phonology is characterized by very great variability. Without arriving at a clear-cut subdivision, this new approach identified Romanian, Sardinian and some areas of central and southern Italy as the most conservative Romance area, while the innovatory area *par excellence* is French (sometimes together with Occitan and northern Italo-Romance). These generalizations match fairly well with Rohlfs's (cf. also Cremona 1970; Green 2006). The correspondence is interesting because it is based on different criteria. It must be noted, however, that Romanian, as well as cases of conservation, presents at every level numerous examples of innovation which are original and unique among Romance languages, and fully justifies the frequent allusions to the originality or the 'individuality' of

Romanian in the panorama of Romance languages (after Bartoli 1945b, see particularly Niculescu 1965; 1978; 1999; 2003; 2007).

Of the methodologies of the latter half of the twentieth century, besides generative grammar, word-order typology (Greenberg 1966) also inspired some interesting observations on Romance, yielding two important conclusions: (i) all Romance languages resemble each other syntactically, offering a more or less rigid SVO type, which they reached via an 'XV' phase (Benincà 1983–84; Adams 1987). In this respect they are differentiated *en bloc* from classical Latin, which was still of the SOV type, as in common Indo-European. This might suggest that, ultimately, it is not entirely necessary to subdivide the Romance languages into groups. One might say that, at least viewing things in very general typological terms, there is only one Romance linguistic type, with some internal parametric differentiations.

8. Classification

Before concluding, we may ask ourselves what is left of the once widely held thesis of Walther von Wartburg (1950), that the Romània is divided into two parts, according to whether the substantival plural is sigmatic or vocalic. The former occurs in the west, from Portuguese across the Gallo-Romance area, to Sardinian, Romansh, Ladin and Friulian. The latter occurs in the central and eastern area (Italo-Romance, Dalmatian and Romanian) (cf. Maiden, volume I, chapter 5, p.164f.). But the question becomes complicated if we bear in mind that in origin at least a part of the vocalic plurals probably have a sigmatic origin (see Maiden 1996). One is bound to ask whether the distinction between plural in -s and the vocalic plural should apply in diachrony or synchrony. In diachrony, going back to the origins, sigmatic and non-sigmatic forms appear to merge, so that the distinction loses importance. In synchrony, the problem is whether, for example, the French plural is still to be considered sigmatic, given that -s is largely only graphic. It only signals plural before a vowel, for example *l'auvent : les auvents* /lovã/ /lez ovã/ 'awning/-s' (and then virtually always only in the article, rather than the noun), unless one allows /s/ to be underlyingly present in the deep phonological structure (Shane 1968). Short of this solution, one has to say that French now has a system (indeed systems) of plural formation different from that of other Romance languages. Clearly, what seemed to be the simple criterion adopted by Wartburg now presupposes a particular theoretical option.

But even if we take Wartburg's proposal back in time, it is unclear why plural formation, rather than other phonological or morphological factors,

should be the chief criterion. Geographically the plural types do not coincide exactly with other phenomena. Only two historical phonetic developments offer a partial correspondence: lenition of voiceless intervocalic consonants, and the loss of final unstressed vowels other than *-a* (cf. Loporcaro, volume I, chapter 2, pp. 65–69, and chapter 3, pp. 150–54). Lenition affects all Romance varieties from Portuguese to northern Italo-Romance (although the latter lacks the sigmatic plural). But Sardinian and Corsican lenition, which has different characteristics, should not be included here (Rohlf s 1966:§209; 1971:44), while Sardinian plural *-s* is truly a matter of general conservation. As for loss of final vowels other than *-a*, this phenomenon only reaches from Catalonia to Friuli, embracing most but not all dialects of northern Italy (not Ligurian or Venetan). In the east, the fall of final vowels other than *-a* in Romanian are of a different date and follow different principles. Thus we see that the geographical coincidence between the three phenomena is only partial.

Wartburg considered the ‘La Spezia–Rimini Line’ – separating dialects of northern Italy from those of the centre-south – to be essential to the division of the *Romània*. In fact, this line only divides the *Romània* in respect of consonantal lenition. It may also delimit the area of fall of final unstressed vowels other than *-a*, but we have seen that the relevant area is smaller than that for sigmatic plurals. Finally, Wartburg included northern Italian dialects in the area of the sigmatic plural on the evidence of an older presence of 2SG *-s* in the verb in some varieties, a move which is quite inadmissible given that the 2SG is something quite different from the sigmatic plural. If this hypothesis might be sustained as a matter of historical reconstruction, as we mentioned, then the original geographical extension needs to be shifted far to the east, covering the whole of the *Romània* and depriving this criterion of any value.

The sociology of the Romance languages

ALBERTO VARVARO

I. The origin of the Romance languages as a sociolinguistic problem

1.1. The process of formation of the Romance languages from the end of the ancient world to the age of Charlemagne has fundamental sociolinguistic implications, regardless of our preferred line of theoretical explanation. In essence the problem is as follows: at the beginning of the fifth century AD the Latin linguistic area stretched from the shores of the Atlantic in modern Portugal to Hadrian's Wall in Britain, to the course of the Rhine and the Danube, in continental Europe, and included the coastal strip of northern Africa from the Straits of Gibraltar to Sirte. The boundary with Greek ran from the Black Sea to the Adriatic, across modern Bulgaria, Macedonia and Albania. In this vast area there were certainly outcrops of other languages (compact remnants of pre-Roman languages or areas of bilingualism), but it should be said that Latin was in use, at least in the army and administration, in the eastern part of the Empire as well.

Now all the information we have, explicit or implicit, suggests that regardless of diatopic, diastratic and diaphasic variations, the Latin system retained a clearly structured compactness, in the sense that: 'Ces variétés se relient entre elles, sur le plan de la langue, par un diasystème commun, et, au niveau des utilisateurs de la langue, par la capacité d'intercompréhension et la conscience d'appartenir à une communauté linguistique identique'¹ (Herman 1996a:44). Witness the fact that, at least until the early fifth century, we cannot make out the area of provenance of a text by its linguistic characteristics, nobody states that they are unable to understand the speech of other Latin speakers and nobody believes that they are speaking anything other than just Latin.

1 'These varieties are interlinked, at the level of language, through a common diasystem and, at the level of language users, by mutual intelligibility and awareness of belonging to a single linguistic community.'

Four centuries later the situation has changed radically. Texts from this area (of course only written texts have come down to us) have unmistakable diatopically differentiated characteristics both when they are assumed to be set down in Latin and, much more so, when they are in the vernacular language; the degree of mutual intelligibility is hard to quantify, but Latin itself seems to be in crisis; leaving aside Latin texts, the popular varieties acquire denominations which are generic (*romanz* or *romance*), or specifically geographic (*español*, *lemozi*, *proensal*, *normand*, *picard*, etc.) (cf. Müller 1996).

What has happened between these two dates is more than a highly complex process of linguistic change, for the basic facts of the sociolinguistic situation, 'who speaks what language to whom' (cf. Fishman 1965), have themselves changed. That this is not simply a matter of explaining how forms and structures of Latin have changed to give rise to Romance forms and structures emerges from two fundamental considerations: Latin is not ousted by Romance varieties but survives alongside them for centuries; Latin does not evolve into a single Romance system, but gives rise to a multiplicity of systems.

The aspect of this process which is most easily understandable is the survival of Latin. The process is characterized throughout by an unbridgeable disparity of prestige between Latin and its alternatives (other languages, or vernacular varieties, with the exception of Greek). This means that Latin retains what sociolinguists call the 'high' functions: culture, religion, administration (or as much of it as is left after the end of the Empire). All other functions are taken over by 'low' varieties. The Latin which remains in use for 'high' functions tends to maintain, within the limits of the capacities of those using it, its normal, standard, form, such as it was at the end of the period when it was still generally flourishing. Deviations from such a norm, due to ignorance or a desire not to lose contact entirely with the vernacular varieties, may be repaired by resorting to ancient models: this is what happens in the so-called Carolingian reform (cf. also Banniard, this volume, chapter 2; Wright, this volume, chapter 4).

What I have said implies a fundamental change. There has arisen a sociolinguistic situation quite different from the ancient one, and of a diglossic type, with functional specialization of various varieties, one of which is reserved for 'high' functions. The shrinking in the range of use of Latin entails a corresponding reduction in the number of those using it and makes it unlikely, if not impossible, that Latin was anybody's first language. In what varieties, then, are 'low' functions carried out?

1.2. It must be said at the outset that the available documentation forces us to speculate, because the texts and the information which have come down to

us involve, almost without exception, the 'high' functions. Hypothetically, since after 400 AD there were indeed major upheavals but not a real demographic substitution, it seems obvious that speakers continued the 'low' varieties of the Latin system for 'low' functions. If the process gave rise to an entirely new linguistic family, as indeed happened, it cannot be explained away by a widening of the divergence between 'high' and 'low' varieties of late imperial Latin. Beyond explaining why, when and how the divergence widened to the point of preventing mutual intelligibility, we have to explain why the diatopic fractures emerged. Here we simply have to argue in socio-linguistic terms.

In the west the collapse of the late fifth-century Empire gave rise to an unstable situation of political fragmentation. The Germanic kingdoms were on the whole ethnically based: they each belonged to one of the federations of tribes, formed in the previous few centuries, which had invaded the Empire. In these relatively unstructured new political entities the legal status of the Romans was formally unchanged, but the rapid Romanization of the Germans weakened the perceived link between language and legal and social status. The unity of the Empire had owed much to the social, cultural and linguistic cohesion of the upper class, first and foremost the senators, whose properties lay in areas far distant from Rome and from each other. The leading circles of the Church, and particularly the bishops, were, on the whole, of the same kind. The linguistic model was the written language of the classics and the spoken language of the senatorial class. The collapse of the Empire led to regionalization of these groups, and of their ideal models. Even the papacy struggled to hold on to property far away from Rome and eventually it fell into the hands of a small group of local families. Throughout the Latin area the new upper classes, with the exception of the Church, were made up of speakers of other languages who were more or less Romanized but estranged from the classical models and from Romans immersed in those models. The collapse of trading links and especially the shrinking of the socio-cultural and political horizon meant that everything was regionalized. In the time of Ammianus Marcellinus (end of the fourth century AD), his readers in Rome were interested in Britain, Arabia and the Sassanid Empire, while the Roman armies and the imperial functionaries, not to mention merchants, travelled the length and breadth of the Empire; two centuries later the horizons of the subjects of the Merovingians had shrunk to little more than their own locality, and did not extend far beyond the boundaries of their region.

The new orientation and the shrinkage of social networks effectively modifies linguistic behaviour. The ancient standard stops being both written and spoken and becomes (save in certain limited domains such as the Church,

schools and in part the law) merely written, and errors are apparent only to those who have studied, but are no longer regulated by normal linguistic intercourse. It is adequate for 'high' functions, but has lost its function of orientating and regulating spoken usage, even at a high social level. Spoken language, deprived of a guiding norm, now turns towards much more locally delimited alternative norms: the usage of people or groups of people whose prestige is much more locally restricted. The linguistic material of this spoken usage is derived from the middle- and low-level spoken language of late antiquity, but the choice of variants (of which there must already have been a good number) had at one time been guided by the prestige model following a scale of acceptability which went all the way down to taboo usages. Now the variants were distributed along mainly diatopic scales which were quite different, and what had once been taboo could become the norm in a particular area. The upshot was the formation of diatopic varieties, all diastatically alternative to Latin but each one distinct from the varieties used in neighbouring areas.

1.3. Here I can give some examples, intended merely to be indicative of extremely numerous and complicated processes. To begin with phonetics, Latin final consonants had a long and intricate history. At the end of the Empire -m must already have long disappeared in almost all positions, while the disappearance of -s remained endemic despite long opposition, a nice example of variation kept under control by the standard norm. Both -m and -s had major morphological functions in the noun and the verb; their definitive disappearance was to favour morphosyntactic processes which the system, as long as it had held out, had rendered superfluous or marginal. Once this collapse has occurred, the -m disappears in all Latin areas while the situation with -s changes from area to area: the eastern part of the Latin area rapidly opts for reduction of -s to -i, or its elimination, and finds other solutions for the problem of the distinction between singular and plural in nouns, and between some forms of the verb; the western part opts to eliminate the loss of -s, which thereto had been endemic. Here -s seems to be restored everywhere, which allows a different, more conservative, solution to the attendant morphological problems. But in fact the isogloss marking retention of -s beats a fairly rapid retreat from the Po Plain; a few centuries later this consonant survives only in liaison (cf. Ledgeway, volume I, chapter 8, p. 399f.) in the northern Gallo-Romance area, notwithstanding it was the essential element in the operation of the residual system of nominal declension; later still, -s was to have yet more different outcomes in Andalusia, the Canary Islands and part of Spanish America.

Yet another example. The late Latin consonant system undergoes a profound change following the transformation of *ɛ* and *i* in hiatus to yod ([j]) and the consequent palatalization of the consonants immediately adjacent to this yod. In the various Romance varieties the outcomes are almost always highly heterogeneous (see Loporcaro, volume I, chapter 3, pp.143–49; Maiden, volume I, chapter 5, pp. 223–25). Take *-li-*, as in *FOLIA* ‘leaves’. The most widespread outcome is the palatal lateral [ʎ], which may be long, as in It. *fogliā* ‘leaf’ [ˈfɔʎʎa] or short as in Occ. and Pt. *folha*, Cat. *fulla*, OFr. *feuille*. But modern French, many Spanish varieties and Romanian have *-j-*: Ro. *foaie*. Many southern Italian dialects have [ʎʎ] (or later reflexes thereof: e.g., *foggħia*); Castilian went from medieval *hoja* pronounced with [ʒ] or [ʎ] to the modern forms with [x]; some varieties of southern Italy, Sardinian and Corsican wholly lack palatalization (*folla*). The results of other palatalizations present areal distributions each of which is different. Clearly in low-level late Latin pronunciation there were always multiple alternatives and in the different areas the vernacular norm has selected different outcomes.

A brief word about the definite article will show that the phenomenon of selection of substandard alternatives also affects categories which were apparently extraneous to Latin of the imperial period. Thus Latin has no article, while all Romance languages do have one and it always originates in the demonstrative adjective, as is the case in other Indo-European languages too. But it is not always the *same* demonstrative: generally it is *ILLE*, but also *IPSE* (in Sardinia, the Balearic islands, in some parts of the Principality of Catalonia, and in the early Middle Ages perhaps in areas of Italy, especially the centre–south; see also Ledgeway, volume I, chapter 8, §3.3.1.2); these articles are generally placed in front of the noun (It. *il lupo* ‘the wolf’, Fr. *le loup*, Sp. *el lobo*, Bal.Cat. *es llop*), but also after it (Ro. *lupul*). There is nothing comparable in Latin texts before the early Middle Ages but, on the one hand, the formation of the article was not unknown to the Indo-European family and, on the other, the fact that the article appears in all Romance varieties suggests that it originates in spoken late Latin at a level of vulgarism such that it was excluded from any kind of documentation. On the one hand, then, the Romance languages adopted and developed this possibility, which was in line with the history of the Indo-European languages, yet on the other they made different choices which, magnified by diverse phonetic developments, gave rise to systems different in both form and function.

As a final example I cite word order. Suffice to say that Latin had a very free order (see Ledgeway, volume I, pp. 387–96) and allowed, at least in high style, the separation of connected elements (noun and attribute, for example,

whether the attribute followed or preceded the noun; for discussion, see Ledgeway, volume 1, chapter 8, §3). All Romance languages have abandoned both SOV order (e.g., *Arma virumque cano* lit. 'arms and the man I sing'), observed only by extremely literary texts as a purely learned calque (see Salvi, volume I, chapter 7, §3.4.9), and the separation of connected elements (e.g., *tacita tecum loquitur patria* lit. 'silent(FSG) with you speaks the country(FSG)'), but the norms that governed, and still govern, the application of SVO order (or the use of clitics, which did not exist in Latin) were and remain different.

1.4. An equally difficult question is the timescale of the process, which we can only guess at. There is no reason to think that it was so slow as to take up most of the approximately four centuries we can allow for, nor that it was everywhere uniform. Far from it: since we are dealing with the linguistic counterpart of a general process of restructuring of identity, society, culture and politics over a very large area, where from an original complex and structured unity various different entities emerge at different stages, it makes sense to assume that the formation of Romance varieties with their own linguistic and social identity emerged at different times and in different ways. That the main Italian areas may have lagged behind, as has often been said and seems to be confirmed by the documentary evidence, gains credibility not only, nor so much, by the greater linguistic closeness between some (but not all) Italian varieties and Latin, but most of all by the fact that a specific socio-political identity, distinct from that inherited from Rome, was slower in taking shape in the Italian peninsula. Conversely, the early awareness of specific Frankishness presumably favoured the early awareness of the difference in the area corresponding to what we call *langue d'oïl*; and the same effect was presumably produced in the Iberian Peninsula first by the strong association of Hispanic identity with the Goths, and then the fundamental contrast with Islam, not in the name of Latinity but in that of religion. In any case the restructuring of a social identity after a period of crisis can come about very rapidly, as can the recovery of equilibrium by a linguistic system after a longer or shorter phase of vacillation.

2. How the Romance languages acquired status

2.1. This issue is generally considered with particular reference to the Middle Ages, in connection with the appearance of the first Romance texts (see also Wright, this volume, chapter 4). I deal here with the acquisition of status in more recent times (see also Andreose and Renzi, this volume, chapter 8, §3).

In modern times the assumption that status is attributable to ideological motives has become increasingly common. The clearest, and perhaps

therefore the most successful, case is Catalan. From the seventeenth century it had lost its status as a language of culture and was partly excluded from written usage. Its recovery of status from the mid nineteenth century and the subsequent achievement of a normalized standard form, which reclaimed ground from written usage, and from elevated usage in general, is a phenomenon of great sociolinguistic interest. But we should not forget that it came hand in hand with the social and economic success of the Catalan bourgeoisie and with their strong sense of identity. The success was complete in Catalonia proper (the so-called Principality) and in the Balearic islands, and here it held out victoriously even against the long oppression of Franco's regime. This was not so in Valencia and Roussillon, for different reasons and with different results. Similar attempts in the Occitan area have to be deemed essentially a failure, for they did not march in step with a parallel social development.

More recently, amid the climate of the reclaiming of ethnic identities characteristic of Europe in the second half of the twentieth century, there has been a burgeoning of initiatives by groups or individuals, often serious scholars, proposing normalized standards for minority languages or for dialects aiming to acquire the status of languages. Suffice it to mention the Asturian of the *Academia de la Llingua Asturiana* in Oviedo (1978–89), the *Rumantsch Grischun* of H. Schmid (1982), the Sardinian of E. Blasco Ferrer (1985), the *ladin dolomitan* of D. Kattenbusch (1991). Such initiatives are neither good nor bad: only the future will tell whether they meet the communicative needs of these groups, and their need for a distinct identity.

2.2. In the last century the Soviet claim that the Moldovan of the left bank of the Prut, spoken in the Soviet Socialist Republic of Moldavia, should be considered a different language both from Romanian and from the Moldovan dialect of the right bank, caused quite a stir (see also Andreose and Renzi, this volume, chapter 8, §3). It was wryly observed that two communities that insisted on their differentness could nonetheless each boast as their own classic writers of the past, such as Mihai Eminescu. Today it seems fair to say that this politically motivated identification of Moldovan as a language in its own right has failed. One of the first measures taken by the Republic of Moldova after the collapse of the USSR was the abandonment of the Cyrillic alphabet, revived (in a new, Russian-based, form) by the Soviets to distance themselves from Romanian, which for a century had gone over to the Roman alphabet. The two-language theory has been given up. But the Republic of Moldova has not united with Romania, something which many had considered merely a matter of time, and the problem could present itself again, albeit in different ways (cf. Heitman 1989; Dyer 1996; 1999).

We would be wrong to imagine that it takes a Stalin to undertake such initiatives. The split between Galician and Portuguese is undoubtedly the consequence of the fact that from 1143 Galicia stayed under the Castilian crown, and did not enter the Kingdom of Portugal. Consciousness of a specific Galician identity, distinct from the Portuguese, goes a long way back – at least to 1576, when Duarte Nunes de Leão (cf. Berschin 1994:368) writes that Galician and Portuguese were once virtually the same tongue, but ascribes the linguistic pre-eminence of Portuguese over Galician to the fact that Portugal had a royal court which provided a kind of workshop for the production and refinement of the language, the like of which never existed in Galicia. Here too there is mutual comprehension, just as there is a shared series of written texts, starting from the medieval poets. While dialects north of the Minho (the river separating Galicia from Portugal) are no different from those to the south, the extensive autonomy enjoyed by Galicia with regard to Spain has not led to any decline in the distinctive identity of Galician with regard to Portuguese, an identity reinforced by the choice of different orthographic norms (cf. Dahmen and Kramer 2001:505).

Typologically analogous is the process at work in Corsica. Here an unquestionably Italo-Romance dialectal system, which moreover is in fact very close to the Tuscan dialect from which Italian emerged, tends to be conceived of as an autonomous language (cf. Marcellesi and Thiers 1988; Goebel 1988). Not wanting to be French – which is what the Corsicans have actually been since the second half of the eighteenth century – has led not to identification with Italian but to the promotion of local identity.

The historical and cultural context of these processes is generally pretty clear. Minor differences are enough to modify the outcome. The dialect of Canton Ticino (Switzerland), a variety of Lombard dialect, never became an autonomous variety within Italo-Romance, even though Ticino was detached from Lombardy from the first half of the sixteenth century and has been an autonomous canton from the beginning of the nineteenth. Nor is there a standard variety of Québécois in opposition to French, even though in this case the actual linguistic differences with standard French are not inconsiderable.

But in this regard the most prominent problem involves the American varieties of Spanish and Portuguese. After the Latin American states reached independence during the nineteenth century (the last was Cuba in 1898), there were indeed sporadic aspirations to create specific national varieties. These were, of course, tendencies widespread in narrow cultivated circles, but lacking a popular basis. In no case was there a complete break, possibly because of the skilful policy of the *Real Academia Española* in striving to

embrace the American academies and in keenly accepting specific Americanisms. This does not mean that standard Spanish in America was the same as that of the Spanish of the peninsula or that it was the same in Buenos Aires and Mexico City, nor that there are not recognizable differences between Brazilian and European Portuguese. But the relation between these varieties is viewed as internal to Spanish (or Portuguese) as an organic system, as is also the case in the anglophone world. In both cases what has saved the unity of the system is not imposed homogeneity but recognition of difference.

2.3. As a footnote to the issue of status-acquisition we need to consider a particularly interesting phenomenon which, once again, is not peculiar to Romance linguistics: the link between the development of the written standard and that of the spoken standard. The latter is generally held to presuppose the former, but the reverse can also occur. The prime example, and one that has been overstated, is Italian. From the first decades of the sixteenth century a barely supra-regional variety of Tuscan used by fourteenth- and fifteenth-century writers became set down as a norm and swiftly acquired the function of a literary standard. Subsequently, Tuscan dialect underwent changes which did not penetrate the standard, so that a perceptible gap between the written language and the spoken variety emerged even in Tuscany (see also Andreose and Renzi, this volume, chapter 8, §3). Talk of generalized diglossia in most Italian regions, and the view that only a tiny proportion of the population understood and spoke the standard are, I fear, exaggerated. After all, from the sixteenth century the standard is widespread, not just in the written form but, very likely, in the spoken form as well, outside the Italo-Romance area: in Malta, in the ports of north Africa, the Adriatic and the Levant (cf. Cremona 2003; Baglioni 2010), and later in Habsburg court circles. It would be odd indeed that a spoken Italian should have existed outside Italy but been virtually non-existent within Italy. There is no doubt, however, that the spoken use of Italian generalized only over the last hundred years, thanks to the progressive spread of the media, and television in particular, as well as the mingling of population brought about by two world wars and later by the internal migrations of the 1950s and 1960s.

The Italian case is not, of course, unique: common Romanian, too, initially emerges as a written variety and only later becomes a spoken variety, setting in motion a process of dialect levelling and homogenization of varieties analogous to what happens with Italian, albeit featuring greater permissiveness of the standard (cf. Ciolac 1989). The same perspective probably applies to the processes of norm-creation from above, which I mentioned earlier, in that these standards initially created on paper exist before anyone uses them.

3. Different sociolinguistic situations in the history of the Romance languages

3.1. A unitary description of the sociolinguistic situation of the Romance languages could not be justified. The situations are different and continually shifting. We would be better off taking the opposite view: one can only give an idea of Romance sociolinguistics by the description of numerous individual situations, without any claim that one can draw generalizing conclusions from them. I shall try to follow a middle path, but I need to say first that we are much better informed on diatopic variation than on diastratic, so we are often obliged, especially where the past is concerned, to resort to indirect information and hypothetical inferences (see also Banniard, this volume, chapter 2).

In the abstract, one might assume that the collapse of the organic linguistic system of the late Empire left the field open to a large number of diatopic varieties, at the same time as the range of diastratic variation grew narrower, or vanished altogether, leaving just a simple Latin vs. vernacular diglossia. It would take centuries for the formation of new, far-ranging systems – the national Romance languages. In effect, the theoretical model of a fragmentation of Latin unity (*ex uno plures*) has widely dominated studies of the subject, and seems obvious and self-apparent.

It seems to me, though, that the process was not so linear, although I must point out that any argument on this subject is bound to be somewhat hypothetical. First, it seems unlikely that the passage to diglossia was instant or even rapid. There must have been an intermediate phase characterized by a continuum comprising different levels of acceptance and prestige. This very structured situation certainly assured more or less general mutual comprehension and the preservation of an awareness of a common linguistic identity. In the different areas of the Romania this will have had different durations, and the break will have come at different moments. It is likely that in Italy this situation endured until quite late, whilst in Gaul such a phase was already over, as seems to be confirmed by Herman's acute study (2002) of the elimination of the Latin synthetic passive – which as we know has no Romance continuants – when texts of Gregory of Tours (end of the sixth century) are reused in the *Liber Historiae Francorum*, little more than a century later (727 AD).

Indeed, the continuum tended to polarize to the point of fragmentation: the acrolect remains as medieval Latin, while the basilect becomes the Romance varieties of each area. Only now do we really have a diglossic situation (cf. Andreose and Renzi, this volume, chapter 8, §4), characterized by the fact that both its poles also have their own sociolinguistic structure. That this happens at

the very end of medieval Latin arises from the differences of level encountered from one writer to another and from formulae like ‘*latinum circa romançum*’ (‘Latin which comes close to Romance’) (cf. Menéndez Pidal 1956:459n). The Romance vernacular side is less easy to investigate, and here it is evident that diatopic variation, leaving aside the continuation of mutual comprehension, was accompanied by typically sociolinguistic differences of status. One of the best known demonstrations of this is what Conon de Béthune (late twelfth century) says about the wry reaction of the Queen of France and her son, the future Philip II Augustus, towards Conon’s Picard, which in the Parisian court was evidently deemed coarser than the local variety, although it presented no problems of comprehension (cf. Wallensköld 1968:5):

La Roïne n’a pas fait ke cortoise,	The Queen has not behaved courteously,
Ki me reprist, ele et ses fieus, li Rois.	In reproaching me, both she herself and her son the King.
Encor ne soit ma parole franchoise,	Although my speech may not be French,
Si la puet on bien entendre en franchois;	One can still understand it well in French;
Ne chil ne sont bien apris ne cortois,	Nor are those people well- mannered and courteous
S’il m’ont repris se j’ai dit mos d’Artois,	who reproached me if I used words from Artois,
Car je ne fui pas norris a Pontoise.	for I was not brought up in Pontoise.

3.2. I will not attempt to catalogue all we know about such disparities of status over the centuries. I may mention how numerous Italo-Romance writers between the thirteenth and fourteenth centuries (from Brunetto Latini to Marco Polo) proclaimed the virtues of the *langue d’oïl*, in justification of their electing to use this language rather than an Italian variety. Between the eighteenth and nineteenth centuries French is once again reckoned as a higher-prestige variety in educated writing and court and diplomatic conversations well beyond the Romance world, for example in Prussia and Russia. A brief mention may be made of the exploitation of similar imbalances of status in the theatre, for comic effect. Just as in Spanish theatre peasant characters are made to express themselves in *sayagués* (a stylized Leonese dialect), in Italian, stage characters in works from the sixteenth century already use a mongrel Italian

mixed with Adriatic varieties of Slavonic and Greek, and the masks in *commedia dell'arte* are portrayed, in opposition to serious characters, as dialect-speaking (Harlequin and Brighella speak Bergamasque and then Venetian with Bergamasque features; Pantalone speaks Bolognese).

The theatre exploited for comic effect the kind of sociolinguistic stratification which comes about in areas subject to immigration of servants: Venice was a prime example of this between the sixteenth and seventeenth centuries. In modern times the same thing has happened with the southern Italian varieties spoken by immigrant workers in Turin or Milan, or with Andalusian or Galician immigrant workers in Barcelona, and so forth. Their speech varieties were considered 'low', even if they were, as with Andalusian in Barcelona, closer to the national standard than was the 'high' local variety. Likewise the speech of the Latinos is bound to seem 'low' in New York City or Los Angeles.

Throughout the Romance-speaking world, it has always been true that major urban centres rated their own speech as superior to that of the surrounding countryside. Representative of countless similar cases is the Neapolitan 'cafone', the country yokel, whose speech is despised not only by users of the standard but also by those who know and use only the urban dialect. This evaluative asymmetry explains why the Romance urban varieties establish themselves in the surrounding territory, but never the reverse. Linguistic geography has very clearly illustrated this contrast. Since Jules Gilliéron, author of the *Atlas linguistique de la France* (ALF 1902–10), the geographical projection of dialect data on to maps has always shown that the countryside imitates the towns.

3.3. Note that we do not mean to say that the contrast between high and low varieties is always bound to lead to the triumph of the former over the latter. The linguistic history of urban centres shows that the processes involved are more complex. From a linguistic point of view we know less than we would like about Paris, long the most heavily populated city in Europe. But it is clear that anything can happen there. The fate of [r] is a good example. The most ancient and widespread feature of its evolution, at least in its popular pronunciation in Paris, is deletion before a consonant: rhymes such as *sage* : *marge* are attested from the first decades of the twelfth century, and Villon rhymes *courges* and *Bourges* with *rouges* and *bouges*; before [l] the pronunciations *Challes* for *Charles* and *pallez* for *parlez* were less stable (Fouché 1961:863f.), as was the assibilation of [r] between vowels: already in Villon *chaire* rhymes with *aise* and then we have *Pazis* for *Paris* (p.603f.). Yet these changes were not ultimately successful: assibilation retreats in the first decades of the seventeenth century and few traces of it are left (*chaise* itself is an example), as with

deletion before a consonant (e.g., *faubourg* for *forsbourg*). But the uvular pronunciation, originally (c. 1600) limited to word-initial position and -rr-, is then extended to all positions, not only in Paris but throughout northern France and thence to much of central Europe, well beyond the frontiers of Romance. This wave-like diffusion is certainly linked with the success of the Parisian Enlightenment model in royal courts and hence confirms that the uvular pronunciation probably imposed itself from on high, from the beginning. Equally complex is the history of the pronunciation of the diphthong *oi*, which in late medieval Paris had at least three realizations: [we], [e] and [wa]. The pronunciation [e] manages, from the sixteenth century, to infiltrate court usage and becomes general in some cases, such as the imperfect and conditional tenses (e.g., *parlais*, *parlerais*) as well as some ethnic names (e.g., *anglais*); until 1700, the distribution seems to be much as follows: [wɛ] in sustained speech, [ɛ] in the normal speech of people of high social class, [wa] in the speech of the *petit bourgeoisie* and the common people. The social shock of the Revolution modified this distribution: the first pronunciation has gone out of use, the second has remained normal only in the categories where it had taken root (hence *Français*, but the personal name *François*, and so forth), the pronunciation [wa] became general (Fouché 1958:272f.; Straka 1990:29; see also Smith, volume I, chapter 6, p. 308).

A special case is the replacement, between 1450 and 1550, of the medieval dialect of Rome, of central-southern Italo-Romance type, with modern Roman, which is of a Tuscanizing type (cf. Ernst 1970; also Andreose and Renzi, this volume, chapter 8, §3). The abandoned features were certainly part of the ancient urban usage, but they also coincided with those of the rural hinterland, particularly dialects of the Sabine area; the Tuscan model was indeed imported by immigrants but these were the popes, the papal curia and their attendant families.

Generalizing these observations, it seems reasonable to state that all the changes which have happened over the centuries in the Romance-speaking world are due not so much to demographic movements as to disparities of sociolinguistic prestige between competing varieties. The centuries-old erosion of Asturian, Leonese and Aragonese by Castilian, the weakening in France not only of Occitan varieties but also of Norman, Picard and particularly central varieties in favour of the Parisian model, the process which has pushed back towards the Alps the dialects of the Po Plain and that which has crushed southern varieties under the weight of Neapolitan, are all examples of this same principle. Sometimes the substitution does not go all the way, sometimes it does. In Trieste the replacement of its originally Friulian type of speech with a Venetian variety is brought to completion only in the

nineteenth century. In Corsica (see also Andreose and Renzi, this volume, chapter 8, §3) the medieval orientation towards Pisa and Tuscany explains the profound change of varieties that were at one time similar to Sardinian.

4. Between diglossia and variation

4.1. It may be possible to attempt a description of the sociolinguistic situations of the Romance world on a scale ranging from the clearly and stable diglossic to the homogeneously variable. But we should be wary of illusions. Friulian was often claimed to be a compact variety, for so long as it was little studied (cf. Francescato 1989:602). The traditional affirmation of the unity of the Romanian language is largely ideologically motivated, and a peasant from Muntenia has a hard time understanding one from Maramureş (cf. Windisch 1989). Portuguese has usually been claimed to exhibit a 'relatively high degree of homogeneity' (cf. Holtus 1994:625), but we may doubt this. Besides, many of these claims really refer to diatopic rather than diastratic differences, although the distinction is not always clear. In the Romance world there are few remaining cases of a degree of socio-cultural homogeneity so great that there is no need for some diastratic variation: an often cited example is the case of Dolomitic Ladin villages where all speakers are peasants, farmers and artisans; but even here diastratic variation has entered the repertory, thanks to the success of tourism in these valleys (cf. Mair 1989:698). Given that field research in these areas has only recently begun and is far from extensive, we may be dealing with affirmations of principle rather than the fruits of observation.

It is often forgotten, as observed above, that evaluation of a sociolinguistic situation must take into account all the factors in play. If we say that diastratic variation is weak in Galicia or Corsica but forget that, regardless of the legal position in these two areas, the high variety in Galicia is Castilian and in Corsica French, we are bound to end up making erroneous generalizations. It is virtually always forgotten that in contemporary Romance societies there are sometimes conspicuously substantial groups of immigrants, often speaking other languages and not always being rapidly assimilated; in recent decades these groups have re-created, in cities and often even in the countryside, sociolinguistic conditions which are far removed from the homogeneity dreamed of by scholars who even now do not take such groups into account (no entry in the *Lexikon der romanistischen Linguistik* alludes to them; but see Andreose and Renzi, this volume, chapter 8, §5). If all Romance-speaking Europe and part of the Romance-speaking world outside Europe, albeit to

different degrees, belong to the most developed and wealthiest sector of the contemporary world, then they are *a priori* unlikely to contain situations which, as in pre-modern societies, lack fine social differentiations entailing complex sociolinguistic architecture.

4.2. All this leads us to expect that it will be very difficult, if not impossible, to place the Romance sociolinguistic situations on a typologically based scale. We shall proceed less systematically, beginning with a fairly well-studied zone, such as Friuli (cf. Francescato 1989). Here a simple stratification along the lines Italian vs. Friulian would do, even though the Italian side would need a finer analysis reflecting the relative weight of tendencies towards the standard or the regional Italian variety of the Veneto, and whether in marginal areas or in the capital (Udine), Venetian dialect needs also to be considered. Today all Friulians are practically bilingual, but this is a recent situation and the number of those speaking Italian as L1 must be fairly limited. Choice of Friulian is favoured by distance from towns, by the isolation of individual areas, by advanced age or by a low socio-cultural level. The more these factors are reversed, the more Friulian becomes restricted to specific situations or domains.

This is so close to a textbook model that we may be convinced that it can easily be applied to other cases. Let us take Galicia (cf. González González 1994). Here too, the use of Galician in competition with Castilian is correlated with features such as 'rural' vs. 'urban', low vs. high social status, low level of education vs. high, advanced age, type of addressee (domestic or local vs. group outsider). Clearly, these are social features everywhere correlated with the use of the low variety. What differentiates the Galician and Friulian cases is that in the former, besides about 85 percent bilinguals there is a group not only of monolinguals in Galician but also of monolinguals in Castilian, and that it has been calculated that only in the villages does the percentage of those who speak Galician well surpass that of good Castilian speakers. This latter group reaches about 70 percent in small towns and is actually greater than 90 percent in the cities. In effect, only in the countryside is a question in Castilian likely to get a reply in Galician.

Both the Friulian and the Galician situations have been defined as bilingualism with diglossia. Note, however, that functional diglossia, where the low variety is used by everyone for ordinary oral communication and the high variety is reserved for formal communication, whether oral or written, is not stable and general in such cases, but restricted to specific places and situations. In general terms, in both cases there is movement towards less rigidly structured systems.

Things are different if, as often occurs, languages of quite different origin are also present in the mix. Consider the Raeto-Romance area of the canton of Grigioni (Grisons, Graubünden, Grischun) (cf. Kristol 1989). Here the varieties in question are the speech of the village, the regional Romance variety, Swiss German and standard German. In the distant past there were probably no social or sociolinguistic stratifications in village communities, except (where relevant) the Catholic vs. Reformed opposition, which had some slight impact on language. In the nineteenth century, the distinction between use of local dialect for internal purposes vs. recourse to standard German (or something approaching it) in relations with the outside world (the capital of the canton, Chur, has been German-speaking since the fifteenth century) is already in place. Later, the generalized spread of schooling led not only to everyone learning standard German but also to their learning the regional variety of Romansh. Demographic shifts gave rise to emigration to German-speaking cities, such as Zurich, and immigration by minor functionaries and office workers or rich holidaymakers, either case reinforcing the diffusion of Swiss German, a phenomenon abetted by frequent marriage of local people with outsiders. The result was the more or less general spread of once limited bilingualism, and the creation of a double diglossia, both on the Romance and the German side, with some intertwining of the two. The outcome today is an extremely complex distribution of linguistic usage and in any case a marked reduction in the use of the village dialect to the advantage of Swiss German which, unlike standard German, is not necessarily identified as 'high' and has become the normal choice for oral communication if the interlocutor is not unequivocally identified as a fellow villager or as a non-Swiss tourist.

Such situations of multiple diglossia are not after all exceptional. Again in Switzerland, this time in Fribourg, there actually existed in the fifteenth century a complex 'pentaglossia' (i.e., involving five languages), between Franco-Provençal, French, Latin, Alemannic dialect and German, while in Alsace there operates an Alsatian–German–French triglossia, albeit undergoing simplification (cf. Lüdi 1990:322f.). In Spanish America, too, one can find cases in which a native local language is overlain by another native language with scope beyond the local, and both are overlain by Spanish (cf. Zimmermann 1992:347). In many cases this situation can be traced back to Spanish identification of certain indigenous varieties as *linguas generales* 'general languages'.

4.3. For situations of simple diglossia (cf. Andreose and Renzi, this volume, chapter 8, §4), analogous to those described by Ferguson, we need rather to look towards some, but not all, of the American colonial territories and some, but not all, of the areas where creole languages are spoken. In

creole-speaking areas, including Haiti, too rigidly described by Ferguson, when the high variety is also the lexifier of the creole, there tends to emerge a continuum that ranges from the lowest form, the basilect, to the highest, the acrolect. This gradation has also been described (cf. DeCamp 1971c) using implicational scales such that the presence, in the speech of a given individual, of a particular vernacular feature implies the presence in that person's speech of certain other vernacular features, some of which imply a larger number of other vernacular features than others.

Although the relation between dialect and standard in Spain and Italy has often been defined in terms of diglossia, in these countries there is generally just a continuum between low usage, wholly dialectal, and high usage, involving the standard, apart from cases in which low usage is attributed to the presence of linguistic varieties of quite different origin. In this direction the most advanced situation is undoubtedly that of French. This has come about, it is usually held, because of the antiquity and stability of the French monarchy, which allowed the consolidation of a French identity earlier than in Spain and Italy. But perhaps even more effective were the consequences of the French Revolution of 1789 on the ideological level, and the practice of subsequent republican governments. For the revolutionaries of the end of the eighteenth century, the fundamental concept of *égalité* implied total commitment to linguistic equality between all citizens. As Barrère (1794) put it: 'chez un peuple libre la langue doit être une et la même pour tous'² (cited in *LRL* (V.1, p. 330; cf. also p. 224)). No wonder that abbé Grégoire was able to prescribe in the same year 'l'usage unique et invariable de la langue de la liberté' (cited in *LRL* (V.1, p. 224)),³ i.e., the use of the educated French of Paris. In reality, however, it was the schools, especially after 1870, with the Third Republic, which translated this ideology into an actual diffusion of French sufficiently powerful to make local dialects give way even before the advent of the modern media.

Thus France has long presented a generalized use of *français commun*, which is geographically differentiated into *français régional* and diastatically into *français familier* and *français populaire*, not to mention variation according to sex and age, or indeed professional and group jargons. The system appears overall multifaceted but relatively simple and particularly devoid of fractures, so that it functions fairly organically and can be analysed in variational terms: a kind of exemplary diasystem which shows how, in a modern society, a unitary

² 'For a free people the language must be one and the same for all.'

³ 'The unique and invariable use of the language of liberty.'

and 'supple' linguistic instrument can operate, however socially articulated it may be, however many speakers it may have and however widespread geographically it may be (see the detailed description by Müller 1985, and the overview by Prüßmann-Zemper 1990). Countries like Italy and Spain seem to be heading the same way, although they lag considerably behind France chronologically.

Nevertheless it needs to be said that this triumphalist account not only overlooks the presence of minority groups speaking other languages (and often not well integrated), but ignores a point of extraordinary theoretical interest with potentially catastrophic consequences for the future. In fact, and this is not a recent development, the distinction between written and spoken language has in France assumed dimensions which far exceed the limits, however elastic, of the merely variational. It is no exaggeration to say that these constitute two different Romance languages, with two different grammars, which are profoundly differentiated not only on the level of writing (the spelling of written French is, like that of English, remote from pronunciation) and the lexicon, but on the far more important level of morphosyntax.

To cite but some of the innovative traits which have become intrinsic to French, we need only say (after Müller 1985) that the wholesale reduction of redundancy in spoken language, whereby a morphological marker is expressed less often than in writing (for example, [ɛlfɛlomonoɔovʁ], which fails to make number explicit, corresponds to two written utterances in which number is actually marked twice: *elle fait l'aumône au pauvre* 'she gives alms to the poor man' and *elle fait l'aumône aux pauvres* 'she gives alms to poor people'), bears highly innovative consequences. In the verb system of the spoken language, person is indicated only by the subject in some tense-forms, in that only the second person plural is distinguished: [tymãʒ] *tu manges* 'you eat' as opposed to [ilmãʒ] *il mange* 'he eats' or *ils mangent* 'they eat', but we have [vumãʒe] *vous mangez* plural 'you eat' (for the first person plural the spoken language uses *on mange*, literally 'one eats', again without a distinctive ending). It follows that spoken French has in many respects abandoned inflectional desinences, moving the indication of person from the right to the left of the verb.

Other equally significant characteristics of the verb system are abandonment of the *passé simple* (*je chantai*, etc. 'I sang') in favour of the *passé composé* (*j'ai chanté*, etc.), of the past anterior (*j'eus chanté*) in favour of the pluperfect (*j'avais chanté*), or the appearance alongside the simple future (*je chanterai*, etc.) of the periphrastic *je vais chanter* ('I'm going to sing'), of the imperfect or pluperfect subjunctive and especially the passive, replaced with forms with *on* 'one' or reflexive structures. All in all, the verb system of the spoken language

is as far removed from that of the written language as that of the Latin system is from that of the Romance languages.

But even more significant is the fact that even in the nominal system, with a few exceptions, number marking has passed from right to left, from the endings to a prefix, generally the article: the opposition between *la ville* 'the town' and plural *les villes* becomes that between [lavil] and [levil]. The slippage between the two languages is so great as to involve in a fairly systematic way typological characteristics: spoken French has moved away from the type common to other Romance languages.

No fracture of such a kind is to be found elsewhere in the Romance-speaking world. If the fabric of French society ever collapsed it could obviously have grave consequences for the sociolinguistic situation: a complete overturning of the norm, a crisis of the diasystem, crystallization of the written standard language as a kind of dead language, estrangement from the other languages of the Romance family. Since the spoken language nowadays allows considerable latitude for variation, such a catastrophe could also bring about the polarization of diatopic varieties distinguished on the basis of different choices made among the available alternatives.

5. Relations between Romance and non-Romance varieties

The interplay of sociolinguistic relations between Romance and non-Romance varieties, which we have occasionally alluded to, is no less varied and wide-ranging than what happens within Romance. I can do no more here than briefly mention two aspects: the historical side, which implies shifts in the external linguistic boundaries of the Romance linguistic family (and hence the process of Romanization of new territories) and the coexistence, particularly noticeable in modern times, of different types of language in the same area.

5.1. On the historical side, our first observation has to be that today, and for a long time, the Romance area does not cover that area which in antiquity had been Latin-speaking. Why are the former Britannia, the regions on the right bank of the Rhine and those once extending from the right bank of the Danube to the Alps, Pannonia, Moesia, Dalmatia and the whole of north-west Africa, not Romance speaking? The answer is far from straightforward and anyway there may well be more than one: in fact it is likely that the causes, process and chronology involved were different. In very general terms, appeal has always been made to demographic explanations: these areas abandoned Latin (or what had become Romance) as a consequence of the Germanic, Slav and Arab

invasions, conceived as actual replacements of people or at any rate as a major demographic addition which swamped the indigenous Latin- (or Romance-) speaking peoples. Things become much more difficult if we approach the question in closer detail, case by case. We know next to nothing of the Danubian areas downstream of Noricum and very little about Africa. Yet the Danubian regions were not just any outlying part of the Empire, and the emperor had long resided at Sirmium. And there is no doubt that Africa was, at least in part, profoundly Romanized, and it is unlikely that the Libyans of the interior, the progenitors of the Berbers, were overwhelmingly more numerous than the Romanized provincials; there can be no reckoning on an invasion from outside before the eighth century. Our sources do speak of the Roman populations fleeing before the barbarians for Noricum, and this is perfectly credible (and may have caused, as some Romanists have argued, greater Romanization of the Alpine areas to which they fled), but it is implausible that such a flight was general throughout the provinces we have mentioned.

Recently the English historian Edward James has rightly observed that the principal difference between the consequences of the Germanic invasions in Britain and on the Continent is linguistic: 'the bulk of the inhabitants of the former Roman Empire on the Continent still speak a language derived from Latin . . . Yet in Britain Latin was wiped out within a matter of decades, surviving only as the language of the Church' (James 2001:61). Why did things turn out in this way? The usual response was that Latin was not deeply rooted in the island, but archaeology has shown the existence of areas of very intense, and refined, Romanization. Other explanations are even weaker. The study of the whole process leads James to a different conclusion: the differentiating factor is that on the Continent the Germanic newcomers were favourably received at least by part of the population of the Empire, while in Britannia the conquest involved a struggle that lasted centuries between the newcomers and the Celts, who had lost their veneer of Latinity. If that is what happened, the conclusion turns out to be typically sociolinguistic: 'The difference between Britain and the rest of the Western Empire may be that the Saxons took over without any local good will. The transfer of power to barbarians on the Continent resulted in little disruption to local power structures or to the basics of Roman life; in Britain that disruption appears total' (James 2001:99).

This may be accurate for Britannia, but that does not mean that the same dynamic can be generalized to the other provinces. The Slav penetration into Pannonia and Dalmatia seems to have involved an ongoing and peaceful infiltration of peasants, who move on to unoccupied or abandoned land and

gradually cut off and surround the towns, obliging the existing non-urban populations to take up pastoral activities and transhumance. The coastal towns may, then, be lost each in different and even opposite ways. The disappearance of Dalmatian in Ragusa and then in Veglia, leaving aside the difference of several centuries, may look similar, but in Ragusa what happened was that the use of Dalmatian remained characteristic of the urban aristocracy, so that when in the late Middle Ages they disappeared or thought it preferable to adopt Venetian, the local speech was deprived of the social group that supported it; in contrast, in Veglia, Dalmatian had remained as an informal, domestic variety spoken by the common people and fell victim, towards the end of the nineteenth century, not to pressure from the Croatian spoken by the lower classes of neighbouring localities, but to diffusion from above of Venetian (cf. Doria 1989).

In the strip of territory along the Rhine and the upper Danube where Romance gave way to Frankish, Alemannic and Bavarian, it was probably partly a matter of violent conquest and actually submersion of the Latinized population by the invaders, but a substantial part was also played by infiltration by the peasantry. The gradual occupation of no-man's-lands or the more or less peaceful sharing out of areas already under cultivation seem to have left a record in the evolution of placenames (cf. Pitz 1997). The principal towns were cut off from their hinterlands, which gradually shrank; Trier, the old imperial capital, Strasbourg and Salzburg long remained Romance. Military power passed to the Germanic peoples or those who came to agreements with them and often shifted into the countryside; religious power remained Latin and urban, but monasteries, centres for considerable prestige and wealth, were sometimes situated outside the towns. For what may have been centuries the linguistic map, if we were able to draw it, would not show a single frontier, with Romance to the west and south, and Germanic to the east and north, but a dappled effect with intermingled patches of greater or smaller size which, after some time, are resolved into more compact areas on either side. The future linguistic frontier, which since the Middle Ages has varied but little, despite the fact that the area was long disputed, is the result of a process of reciprocal assimilation, not of violent substitution.

The process of reabsorption of non-Romance speakers in regions gradually reconquered by Christians in the Iberian Peninsula and Sicily seems to have been different yet again. It is of course possible that some, perhaps substantial numbers, of the Arabic speakers emigrated, by choice or by force, to areas that were still Islamic. But that is not what made Arabic disappear, even though it certainly weakened it, with an effect that ultimately led to a drop in the birth

rate of a kind normally encountered in social groups who experience loss of prestige, power and wealth. It is equally obvious that at the collective level bilingualism must have made great advances, especially insofar as the lower classes acquired a knowledge of their rulers' language. There is no doubt that individuals and entire families abandoned their original faith and language to secure a better future. But the final stage in the complete elimination of Arabic was in general the effect of traumatic circumstances: violent conquest or subsequent religious persecution or, in the case of Sicily, a long civil war followed by forced removal.

5.2. Only for more recent periods can analytically conducted studies be available to us, such as the exemplary investigation in the village of Bonaduz, in central Grisons, a locality which passed from Raeto-Romance to German between 1910 and 1920 (cf. Cavigelli 1969; see also Andreose and Renzi, this volume, chapter 8, §3). The study brings to light the fact that cracks had gradually been appearing in the dominance of the Romance language for at least half a century due to the frequency of marriage to women, mainly, who came from neighbouring German-speaking countries, and to the development of bureaucracy in the canton (recall that Chur, the local capital, is German-speaking), so that the local administrators and religious officials of Bonaduz had become German-speaking. But the collapse of Romance, set underway by this long process, was ultimately very rapid because of a catastrophic event, a fire which completely destroyed the town in 1464. No lives were lost, but the living space of the inhabitants was fundamentally modified as were their habitual interactions with each other. There is every reason to believe that similar processes have been repeated several times in history, especially where there is gradual infiltration. At some point there must have occurred a traumatic crisis which upset the stability of the social group or made it change its old linguistic habits.

While in Europe and Africa the Romance languages were, from the end of the Roman Empire onwards, in overall retreat, except for the Iberian and Sicilian reconquests and the eastward expansion of Romanian, from the fifteenth century they are expanding beyond the Pillars of Hercules. This process has been considered similar to that of the Latinization of the Empire, and in any case it is not simply a matter of the processes summarized above operating in reverse. There is a fundamental difference with respect to ancient Romanization: the Romans won over the upper classes of the provinces and used them as channels of cultural adaptation; in America the upper classes were usually destroyed by violence or deceit and in no way contributed to the diffusion of linguistic change. The absence of the mediating influence of the

old ruling classes may have prolonged the duration of the process of Hispanization, but the change weakened the impetus towards formation of regional Romance varieties and favoured the constant orientation, for centuries, of American speakers towards prestigious peninsular varieties.

In Spanish America, the history of Hispanization seems to continue the practices of the peninsular Reconquest, as also happens in other respects. It is clear that it is not primarily a matter of demographic replacement. However cruel the conquistadors may have been, and without overlooking the complete disappearance of some peoples of the Caribbean area, it is clear that a large proportion of modern Hispano-Americans are biologically of indigenous origin. This proportion is in the majority in almost all the Spanish-speaking countries, except Argentina (and also Portuguese-speaking Brazil, for specific reasons and because of the extreme cultural isolation of the Amazonian natives). Here we have to allow that there has been replacement of one language by another: a vast number of individuals whose ancestors spoke one of the many native languages decided, or were forced, to go over to Spanish and bring up their own children in Spanish. The reason for this is that among natives, such as domestic servants for example, who were constantly in contact with the Spanish, what probably counted was long coexistence and familiarity: in a way, taking their masters' language made them feel members of their masters' families, albeit subordinate ones. It is less likely that they were absorbed into a majority which was of European origin and European language, because colonists from Europe were always, and everywhere, in an (often tiny) minority, and the social distance, plainly advertised by physical distinctness, was virtually always insurmountable. Only women could be accepted by the Whites, but mostly as concubines, not as wives. Religious conversion was not the decisive factor it has often been in Europe. The missionaries soon saw it as their duty to learn the indigenous languages and to convert the pagans through their ability to express themselves as they did. Anyway, extremely large masses of natives gradually turned into Spanish speakers.

The process of Romanization of Latin America in any case lasted for centuries and is far from complete today, and indeed we cannot be certain that in the future one of the very numerous indigenous languages may escape submersion (the best placed candidates seem today to be Guaraní in Paraguay, in particular, and Quechua in Peru).

For slaves of African origin things went rather differently. The linguistic solidarities between groups of slaves were generally already broken in the holding centres set up along the coasts of West Africa and during the long crossings, for reasons of security at least. Among the *indios* nothing of this kind

occurred: group solidarities persisted intact for centuries. Some of the native American languages were elevated by the colonizers to the by no means ignoble status of *linguas generales* ('general languages'), something which never happened for the languages spoken by those of African descent. The *linguas generales* were the most widespread and prestigious indigenous varieties at the time of the conquest, which the conquistadors acknowledged as languages for communication with all the natives of the relevant areas. There was good reason for the belief, around 1600, that to be able to communicate without difficulty with native Americans it was sufficient to know Nahuatl, Guaraní, with perhaps also Aymará, Puquina, and Mochica or Yunga (cf. Pottier 1983).

It looks as if the Church was more concerned with the *indios* than the slaves, even if, as we have said, conversion to Christianity did not involve abandoning their mother tongue for the *indios*, while the mother tongue of the slaves was considered something of a disadvantage. In any case, the acquisition of literacy, which is not rare among domestic slaves, was always in Spanish or Portuguese, and there was no escape route to modernization, no hope of social mobility, without a mastery of the Hispanic languages, except in the case of Paraguay.

5.3. In modern times the internal linguistic minorities are relatively weak throughout the European Romance-speaking countries save Romania. In Italy, such inroads are limited to the Germans of the Alto Adige / Südtirol and the Slovenians of Trieste, both situations going back no further than 1918, and where the minority groups do have some legal protection for their languages. Only small inroads have been made in Val d'Aosta, whose dialect is Franco-Provençal and whose 'high' language is French alongside Italian. Here, membership of the bilingual complex of Savoyard states goes back many centuries. The other linguistic minorities, numerous but widely varied, date from the Middle Ages and have long been integrated. Their languages stand in the same relation to standard Italian as the Italian dialects, from a sociolinguistic point of view: i.e., they are low varieties generally limited to the domestic environment. I doubt that the recent granting of rights to these languages in the school and in administration will make much difference to this situation. The position of Ladin speakers is more delicate, in that they almost always identify more with their German-speaking neighbours than with their Italian neighbours.

In France, Bretons, Flemings, Alsatians or Basques do not seem too badly affected by the minority status of their own languages and diglossia with French. In Spain, the Basque problem has long had serious political aspects, but mercifully does not turn on the issue of language.

As for Romania, within its present boundaries about 12 percent speak other languages, including 8 percent Hungarians, prominently concentrated in some regions of north-eastern and eastern Transylvania which have been colonized by Hungarians since 900 AD. The German-speaking minority, also of medieval origin, which remained prestigious and numerically buoyant until 1944, has dropped to 2 percent. The other minorities, Roma ('gypsies'), Serbs, Bulgarians, Turks, etc., are few in number and wholly lacking in prestige. Romanian is universally recognized as the national 'high' variety, but its status is challenged by the Hungarian group in the areas where they are prevalent, even if bilingualism can now be said to be general among the minority groups (cf. Krefeld and Schmitt 1989).

The situation is less clear in the Republic of Moldova, where the non-Romanian speaking (Ukrainian-, Russian-, Turkish-speaking) population is very strong and in some places even in a majority. The statistics for 1989 (cf. Deletant 1996:68) attest 2,660,000 Moldovan speakers versus 554,000 Russian speakers, 396,000 for Ukrainian, 140,000 for the Turkic language Gagauz, 67,000 for Bulgarian, 17,000 for Yiddish, 9,500 for Romani and 82,000 for Belorussian. The majority of these speakers appear to be able to use and understand Moldovan (i.e., Romanian), but to what extent is unclear (Ciobanu 2002).

It should not be forgotten that the last few decades have seen immigration from outside Europe assume a pan-European dimension, where until half a century ago it was limited to France and Britain. A huge number of Turks and Kurds settled in Germany after the Second World War. From 1989 the ranks of immigrants from outside Europe were swollen by a no less conspicuous influx from within Europe but outside the then boundaries of the European Union (particularly people from the Baltic states, Poles, Ukrainians, Romanians or Albanians). When the number of immigrants was a minority, and European societies were more self-confident, it was taken for granted that the process would lead to rapid integration, both social and linguistic. But the formation of increasingly numerous and compact groups, particularly in major cities or outlying parts of those cities, together with the rekindling of Islamic awareness in many of them, make the possibility of assimilation, or of speedy assimilation, look doubtful. All the young people may be bilinguals, but it is far from a foregone conclusion that the Romance language is bound to become L1 in all the immigrant groups in just a couple of generations, as happened with internal immigrants in Italy, France or Spain, whose second generation always abandoned their original dialect in favour of the standard or the local Romance dialect (apart from the Castilians in Catalonia). There might even emerge enclaves making stable use of the other language, like

the medieval (e.g., Albanian) enclaves of Italy, but with the additional factor of distinct cultural and religious identity.

5.4. A word is in order about a situation which is bound to have consequences for the Romance-speaking world of the future. It is no surprise that the language of the colonizers should have been adopted in Latin American countries when they gained independence after centuries of domination and settlement. But it was less foreseeable that the same thing should have happened in countries decolonized after 1960, which had been European colonies for under a century (apart from a few coastal outposts) and contained extremely low proportions whose L1 was the European language. There are two reasons for this. On the one hand, the frontiers of these countries were laid down along totally artificial lines and were not coextensive with ethnic boundaries: it is almost always the case that peoples speaking different languages live together within the same country, while populations speaking the same language are split between different countries. In such a situation, to elevate an indigenous language to the status of official language would be a major advantage for the ethnic group speaking it, at the expense of all other groups, but the choice of the language of the old colonial motherland would be to nobody's advantage (apart from small upper echelons of society who already know that language, and therefore stand to gain from such a choice). On the other hand, this option lends the new state a veneer of international accessibility and paves the way for the development of the educational system, often with the support of the relevant European nation.

In this way, the last fifty years or so have seen the emergence of a kind of *Romania Novissima*. Much of Africa has officially adopted French, Portuguese and Spanish (likewise English in the former British colonies), which are, however, the L1 of a vanishingly small minority and are spoken as L2, at least to begin with, by limited groups. Schools use the European language almost without exception, and in that language pupils learn to read and write. One cannot tell today whether eventually the Romance language will become the L1 of substantial groups of the indigenous population, and therefore when the countries in question will be able to be considered for all intents and purposes Romance-speaking (at present they are Romance-speaking only for a certain number of 'high' uses). And it is hard to predict whether genuine local varieties (say Congolese French or Mozambique Portuguese) will ever emerge. At present there are certainly differences from area to area, but the fact that schools are the spearhead of this new diffusion of the language, and the incidence of radio and television, are factors which might sustain the unity of these Afro-Romance varieties.

Romance outside the Romània

MARI C. JONES AND CHRISTOPHER J. POUNTAIN

1. Introduction

A number of the Romance languages, most notably Spanish, Portuguese and French, have in the course of the last five centuries been carried beyond Europe, both as the spoken languages of individuals and as the official languages of empire. Indeed, although Romance linguistics as a discipline has traditionally been Eurocentric, with consideration of what Elcock (1975) termed 'Greater Romania' often somewhat marginalized, it is worth remembering that today the vast majority of Romance speakers inhabit areas outside the original Romània: in particular, Mexico and Brazil are respectively, by some considerable margin, the countries with the largest Spanish- and Portuguese-speaking populations. The Romance diaspora resembles the dissemination of Latin over an area far exceeding its original territory of Latium, and we may expect to find many of the same issues raised by this process (for which see Varvaro, this volume, chapter 1): the extent and nature of dialectalization, the rate of linguistic change, and the extent of the influence of the indigenous and other languages with which Romance came into contact. This chapter will explore some of these themes.

2. Figures

Outside Europe (for which see Andreose and Renzi, this volume, chapter 8, and Sanson, this volume, chapter 7), Spanish is a *de jure* or *de facto* official language of Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, the Dominican Republic, Ecuador, El Salvador, Equatorial Guinea, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Uruguay and Venezuela. The Spanish-speaking population of the United States is variously estimated, but is likely to be in excess of 30 million. There is a small minority of speakers in the Philippines, and Spanish is also still used

in the Sahrawi Arab Democratic Republic. Of the 329 million speakers worldwide, some 300 million are from outside Europe.¹

Portuguese is an official language of Angola, Brazil, Cape Verde, East Timor, Equatorial Guinea, Guinea-Bissau, Macau, Mozambique and São Tomé and Príncipe. There is still a significant minority presence of Portuguese speakers in Goa, Daman and Japan. Of the 178 million speakers worldwide, some 168 million are from outside Europe.

French is the official language of the *départements d'Outre-Mer* (Guyane, Guadeloupe, Martinique, La Réunion, and Mayotte), the *collectivités d'Outre-Mer* (French Polynesia, the French Southern and Antarctic Lands, New Caledonia, Saint Martin and Saint Barthélemy, Saint Pierre and Miquelon, and Wallis and Futuna), Quebec province in Canada, Bénin, Burkina Faso, the Central African Republic, the Democratic Republic of Congo, the Republic of Congo, Côte d'Ivoire, Gabon, Guinea, Mali, Niger, Senegal and Togo. It is one of the official languages of Burundi, Cameroon, Canada, Chad, Comoros, Djibouti, Equatorial Guinea, Haiti, Madagascar, New Brunswick province in Canada, Rwanda, Seychelles and Vanuatu. To a much greater extent than any other Romance language, it is also widely spoken as a second language. Of the 68 million speakers worldwide (though the inclusion of second-language speakers would raise this figure to around 100 million), 15 million are from outside Europe.

Italian is spoken by sizeable immigrant communities in the United States, Canada, Brazil, Argentina, Uruguay and Australia; some 5 million of its 62 million speakers are from outside Europe.

3. The historical background and diasporic scenarios

3.1 *Romance as an official and native language*

Portugal and Spain may be considered together. Portugal was initially at the forefront of Atlantic exploration and conquest, partly because of its strategic position, and partly because of dynastic disputes within the other Iberian Peninsular powers following the unification of Castile and Aragon through the marriage of Ferdinand of Aragon and Isabella of Castile (the Catholic Monarchs) and the conquest of the Moorish kingdom of Granada. The Treaty of Alcáçovas / Alcazovas-Toledo of 1479 fixed the Canary Islands as a Castilian

¹ Based on Ethnologue (www.ethnologue.com; last accessed 10 April 2012), which reflects census figures from around ten to fifteen years previously.

possession and Madeira, the Azores and Cape Verde as Portuguese, with Portugal also being given the right to territories further south along the African coast. Columbus's Castilian-sponsored expedition of 1492 was critical for the expansion beyond the Romània of Castilian, which also from around this time can be more properly termed 'Spanish', as it became increasingly adopted at all social levels, not only within Castile but also in Aragon. The Treaty of Tordesillas of 1494 created a demarcation line between territories which could be claimed by Castile/Aragon and Portugal to the west of the Azores; the precise location of the line changed as a result of different expert opinions and it was never rigorously enforced. However, it allowed Spain to claim massive areas of the Americas to the west of the line, and Portugal to claim Brazil when Cabral discovered American land to the south of the Equator in 1500 whilst en route to India. These territories form the major part of the present-day Portuguese- and the Spanish-speaking worlds. In 1529 the Treaty of Zaragoza similarly established a longitudinal demarcation in the Far East, permitting Portuguese control of the Moluccas, although it awarded to Spain the Philippines, claimed for Spain in 1521 by Magellan, whose expedition had proceeded via the Pacific. Portugal was thus able to form a network of outposts and early colonies in various localities along the African Atlantic coast, around the Indian Ocean and in the China Sea. French colonial expansion began in a significant way only in the early seventeenth century with the establishment of Acadie (Acadia, ranging over parts of present-day New Brunswick, Nova Scotia and Prince Edward Island), Quebec and Louisiana.

Generally speaking, in 'settlement colonies' (those which experienced substantial colonial settlement and a long period of rule from the mother country, a scenario that was typical of the American mainland), the colonial language was firmly established and a series of new regional varieties were developed, some of which were eventually significantly different from the European standards. While indigenous languages also survived quite strongly in some of these areas, the colonial language tends to be the majority language (a notable exception to this pattern is Paraguay; see §5.1.1). In 'exploitation colonies', where colonial settlement was slighter, the colonial language was a minority language, as in a number of the Portuguese African and Indian possessions. Where the colonial language formed the basis of a pidgin for a population who spoke mutually unintelligible languages, a scenario which particularly applied in the context of the African slave trade on both sides of the Atlantic, creole languages developed (examples are the French-based creoles of Haiti, Guadeloupe, Réunion and Mauritius and the

Portuguese-based creoles of Cape Verde, Guinea-Bissau and São Tomé and Príncipe) and speakers of the colonial language are still in a minority. Yet another linguistic scenario obtained in a number of American localities where imported black African slaves developed a 'restructured' Romance language (sometimes referred to as 'creolized' because while it has some creole-like features, such as absence of gender and number agreement and simplification of the inflectional verbal system, it has remained closer to the Romance model than a full creole). Such are the Afro-Hispanic languages generically referred to as *bozal*, of which there is ample testimony during the colonial period and subsequently, though today very few survive and these are terminally threatened (see Lipski 2005). Creole languages will not be further examined in this chapter, since they receive special treatment in Bachmann, this volume, chapter II.

3.2 *Romance co-official with another diasporic language*

The rival ambitions of the colonial powers led to territorial annexations and exchanges with resultant changes to language status. In the course of the eighteenth century, France ceded Acadia, Quebec and most of Louisiana to Britain (the area around New Orleans being ceded to Spain: this was transferred back to France in 1800 and then sold to the United States in 1803). In Quebec, French was strongly maintained as a result of the virtual separation of the French- and English-speaking populations, and today it continues to be spoken by some six million people and is the sole official language. In Acadia, French is usually found in a diglossic relationship with English (Arrighi 2007:48) and is considered a minority variety of relatively low status. In the present-day state of Louisiana, French and Spanish continue to be spoken by small minorities of the population as 'home' languages and have no official status.

The Spanish empire in North America had spread to Florida, New Mexico and California (and, as we have seen, temporarily part of Louisiana) between the sixteenth and eighteenth centuries, but this situation changed dramatically in the course of the nineteenth century, with large parts of northern Mexico being lost, ceded or sold to the United States; Spanish has been extensively maintained as a 'home' language in these areas. In more recent times there has been massive immigration of Spanish speakers, especially from Mexico, Cuba and Puerto Rico, who have been atypically slow to assimilate into US society and have maintained close-knit social networks (Bodvarsson and Van den Berg 2009). The consequence is that while Spanish is not an official language of the United States (although it enjoys official special status in

New Mexico), in practice it is widely used and many official documents are routinely published in Spanish.

*3.3 Bilingual contact scenarios: conquest with a
Romance language established at an official level but not
as a native language*

Portuguese contact with Africa has been continuous since the late fifteenth century, but the presence of Portuguese speakers was until the late nineteenth century limited to coastal areas, and so this is in essence an ‘exploitation’ colony scenario. Today, in the African countries in which Portuguese is an official language (often referred to as ‘the Five’), two strongly contrasting situations may be identified. In Angola and Mozambique, Portuguese is firmly established. In the colonial period, learning Portuguese was the only way to economic and social advancement, since it was not only the official language, but the only language of education; in Angola, admission of Africans to the more favoured status of ‘*assimilados*’ (lit. ‘assimilated ones’) was contingent upon their demonstrating proficiency in Portuguese. During the revolutionary struggles which preceded independence, Portuguese was the obvious candidate, as it was a vehicular language which allowed speakers of many different indigenous languages to communicate, and since independence in 1975 Portuguese has been strongly promoted as the language of national unity. The number of speakers has increased substantially since independence as governments inimical towards the Portuguese colonizers nevertheless acknowledged the practical advantages of maintaining Portuguese, which not only had an existing base as the language of administration and education, but was also important as a major world language.² The impact of such a language policy has been dramatic: Mingas (2000:48) reports, for example, that since independence a shift has taken place in Luanda from a population which was predominantly monolingual Kimbundu-speaking to a situation in which bilinguals are now in the majority, with a monolingual Portuguese-speaking group in second place. Angola and Mozambique do not really have any Portuguese creole languages, although what is sometimes referred to disparagingly as ‘*pretoguês*’ (‘Black Portuguese’) is identified in Angola, and the ‘*linguagem dos musseques*’ is an urban slang of Luanda which shows some

2 The Guinéa-Bissauan nationalist leader Amílcar Cabral famously said that the Portuguese language was the most beautiful gift the Portuguese left to Guineans (this view has been strongly challenged by Paulo Freire; see Freire and Macedo 1987:109).

creole-like features, such as the merger of polite and familiar pronouns of address, merger of case and gender in third person pronouns and some reassignment of functions amongst prepositions, with *em* 'in' used extensively (Perl 1994:98–106).³ These scenarios contrast with the other three countries of the 'Five', Cape Verde, Guinea-Bissau, and São Tomé and Príncipe, where Portuguese creoles are very widely spoken. In Cape Verde, which has a very high level of literacy (in Portuguese), creole is nonetheless spoken universally and is the language of everyday communication, even though Portuguese is used for official purposes, in education and in the media; African languages are no longer spoken (the Cape Verde archipelago was originally uninhabited and was used as a slaving centre). The situation in São Tomé and Príncipe is similar, although a number of different creoles are used in the two islands (São Tomense, or Forro, and Angolar in São Tomé, and Principense, or Monco, and Cape Verde creole in Príncipe), the result of different waves of settlement leading to the establishment of discrete ethnolinguistic groups. In Guinea-Bissau, a Portuguese creole (Kriol) is the majority language with an increasing number of monolingual speakers (Holm 1989:275), though several African languages are also spoken; Portuguese is still spoken only by a small minority, given that the literacy rate is low.

French was not taken to Africa until the nineteenth century. By this time, advanced linguistic codification of the language and the change of political regime within France had led to increased homogenization of speech and to the mindset that standard French was the only suitable variety for use by the ruling classes. As a result, a more uniform variety of French was diffused on this continent. In contrast to Canada, where it is possible to distinguish clear-cut, identifiable linguistic subsystems (for instance, the working-class variety of Montreal commonly referred to as *joual*), in many varieties of African French regional and social varieties are often interlinked (and sometimes also combined with fossilized interlanguage phenomena) (Drescher and Neumann-Holzschuh 2010b:16). Moreover, although French carries the status of official language in many countries of west, north and central Africa, the percentage of native speakers may be small (less than 1 percent in Chad, some 2.5 percent in Gabon, etc.). As French is frequently introduced via the education system, an imbalance also tends to occur between numbers of

3 The 'linguagem dos musseques' was used by some Angolan authors as a kind of 'stage Portuguese', and so is sometimes thought of as an African Portuguese stereotype (Perl 1994:73).

speakers in urban and in rural areas, between young and old speakers and between white- and blue-collar workers.

The situation of Portuguese in Africa contrasts with that of French: Ferreira (1988:65–68) argues strongly that the development of Portuguese in Africa is not the same as that of other colonial languages, since, unlike English and French, the adoption of Portuguese has not been the result of imposition by the colonial power so much as a tactical decision taken in the interests of raising literacy levels and achieving national unity, together with recognition of the advantages of speaking a world language.

3.4 *Romance as a vehicular language*

In Equatorial Guinea, which was colonized by Spain only in the nineteenth century, following a chequered history under Portuguese, Dutch and English influence, French and Portuguese are also official languages and the European languages are vehicular languages for a population which speaks a range of African languages natively. French is a frequently spoken vehicular language though not an official language in the former French colonies/protectorates of Algeria, Morocco, Tunisia and Lebanon, as is Spanish in the Sahrawi Arab Democratic Republic.

3.5 *Loss of contact with the metropolitan norm*

In 1492 the Sephardic Jews were expelled from Spain by the Catholic Monarchs. This resulted in their diaspora initially in several locations around the Mediterranean, especially the Ottoman Empire, and ultimately to northern Europe and the Americas. In this particular scenario, contact with metropolitan Spanish was lost. The Spanish of the Sephardic communities, known as Judaeo-Spanish or Ladino, developed in ways which are not shared with Spanish in Spain and the New World, most obviously with respect to the medieval sibilant system of Castilian (/ts:/ /dz/, /s:/ /z/, /ʃ:/ /ʒ/, which underwent mergers involving loss of the voiced term to yield /θ/, /s/, /x/ in standard modern peninsular Spanish and /s/, /x/ most generally elsewhere, but which retained the voiced/voiceless oppositions /s:/ /z/, /ʃ:/ /ʒ/ in Judaeo-Spanish).⁴ Judaeo-Spanish also missed much of the significant

4 There is a good deal of debate concerning the distinctiveness of a Jewish Castilian prior to the expulsion, though the balance of scholarly opinion is now against this (Séphiha 1986:23–24; Harris 1994; Penny 1996); the same view has been developed for Judaeo-French, which was not preserved outside the Romània (most recently Banitt 1963:253; Kiwitt 2007; 2010).

learnèd influence of Latin and Greek exerted by the Renaissance and the technical terms necessitated by the Industrial Revolution. Judaeo-Spanish speakers borrowed words extensively from the communities with which they came into contact (heavy borrowing from Arabic in north Africa created a now moribund hybrid language (see §5.3) usually distinguished by the name of *haketia*), but Balkan Judaeo-Spanish also looked towards other Romance sources, especially French, the language of the *Alliance israélite universelle*, which provided education for Jewish families in the nineteenth century, to supply some 'learnèd' and technical words. Words such as *dezvelopamiento* 'development' < Fr. *développement*, *reushir* 'to succeed' < Fr. *réussir* and *buto* 'aim' < Fr. *but* are in common use, even though in more recent times there has been a rapprochement between Judaeo-Spanish and standard Castilian. The successful maintenance of a non-national language outside its mainstream linguistic community for such a long period is remarkable, and is attributable to the Sephardim's tightly knit communities and sense of religious and cultural identity: to this day, Sephardic Jews often regard their language as their defining feature (Harris 1994:121).

3.6 Immigrant contact scenarios

Judaeo-Spanish and the case of Spanish in the United States, where Hispanic immigration may be thought of as having reinforced an already substantial native presence of Spanish, stand apart from what may be regarded as more typical immigrant contact scenarios, in which the general fate of immigrant languages appears to be eventual attrition.

This is clearly shown by the fate of French in the United States, where various communities of French speakers became established (apart from the Louisiana situation discussed in §3.2). Significant numbers of immigrants from New France (modern Quebec) travelled in search of work to New England during the nineteenth and early twentieth centuries. On arrival in the US, these immigrants (known as Francos) frequently organized themselves into communities, sometimes known as Little Canadas. Today, French is still spoken by significant numbers in some of these communities (in Maine, for example, some 5 percent of the population speaks French at home), while in others it is less vital (in Connecticut and Massachusetts, for example, the figure is closer to 1 percent). However, the immigrants often assimilated with the local population so that, although the percentage of people with French or French-Canadian ancestry may be quite high in some New England states, numbers actually speaking French as an everyday language are usually considerably lower: New Hampshire, 25 percent French ancestry vs. 3 percent

French speakers; Vermont, 24 percent French ancestry vs. 2.5 percent French speakers; Connecticut, 9 percent French ancestry vs. 1.31 percent French speakers (the percentages given are approximate). Another wave of immigration was the result of persecution, both religious and political. Several thousand Huguenots fled to the US after the revocation of the Edict of Nantes in 1685. The largest of these Protestant settlements were established in North Carolina and Virginia, the most well-known being in Charleston, where monolingual religious services in French continued until 1928. For the most part, however, these immigrants seem to have assimilated with the larger Presbyterian community within a few generations, in terms of both their language and their religion (Gingrich 2009; Fohlen 1990:32). In the nineteenth century, economic reasons led many French people to seek a new life in the US, as they were driven out of France by rural overpopulation in the 1840s and subsequently attracted by the Californian Gold Rush. Indeed, by 1856 the number of French people in California is estimated at some 245,000 (Fohlen 1990:36). Communities established during this period include the French Quarter of San Francisco when, in 1851, approximately 3,000 settlers, sponsored by the French government, came in search of their fortune. This community was later to be supplemented by French immigrants working in the wine-making trade. Texas also became an important centre of French immigration in the nineteenth century. Recent immigration from France has been steady but relatively small-scale, representing, for example, only 1.1 percent of the total number of immigrants to arrive in the US between 1961 and 1975. Twentieth-century immigration from French Africa has also led to newer francophone communities of several thousand strong being established in, among other places, the Florida cities of Orlando and Tampa.

It is possible to identify characteristic stages in the process of immigrant language shift: Berruto (1987:180–85) observes that immigrant Italians in the US at first show levelling in favour of the standard, with dialectal differentiation disappearing, and employ the reduced code of *italiano popolare*, even in social environments which call for more elaborated or educated language. The next stage of development is the extensive use of English lexicon and calques which are assimilated to the phonological and morphological system of Italian, and this is what is often characterized by such names as *italo-americano* ('Spanglish' in the case of US Spanish). At its most extreme this process can create a hybrid such as the Spanish/Italian *cocoliche* in Argentina (documented but now extinct; see §5.2.3.2.2). This is also the point at which subsequent generations appear to take an attitude towards their heritage language and will either progressively abandon it or learn it normatively. The consequence of a negative attitude is the

development of features associated with language death (see §7). Stronger maintenance of the immigrant language and a high degree of bilingualism may lead to code-switching (see §5.2.3).

In the case of immigrants from Italy, the preservation of languages other than the national language has to be considered. One relative success story in this respect is that of the language known as Talian, which is spoken by Italian immigrants, mainly from the Veneto, who settled in Rio Grande do Sul, Brazil, from 1875 onwards. The relatively remote location of the agricultural community they established meant that, against the odds of government prohibition of languages other than Portuguese in the mid twentieth century, a distinctive language based on Venetan, though with considerable borrowing from Portuguese, was preserved, and in 2009 it was officially declared part of the Historical and Cultural Heritage of Rio Grande do Sul state. It has achieved a degree of standardization through the creation of websites, some written literature, radio and song.⁵

4. The linguistic base of the overseas varieties

The ‘settlement colony’ scenario is generally considered to involve continuation of the language of the home country and is often characterized in terms of conservatism (preservation of features attested in earlier stages of the home country language but then modified there – it is clear that looked at from the opposite standpoint the language of the mother country might alternatively be considered ‘innovative’ in such respects) or innovation (the introduction or modification of features which is not paralleled in the mother-country language). An example of a ‘conservative’ feature is the preservation in many areas of Spanish-speaking America of *vide* and *vido* as 1SG and 3SG strong preterite forms of the verb *ver*; while such forms are amply attested in older Castilian, they have been replaced in the standard modern language by the regularized forms *vi* and *vio*. Conversely, the use of an assibilated articulation of /r/ in Chile, southern Bolivia and north-west Argentina may be seen as innovative, since such a pronunciation has never been evidenced elsewhere in the Spanish-speaking world. Yet some developments seem to involve elements of both conservation and innovation: the *voseo* of the Spanish of the River Plate area, in which the personal pronoun *vos* is used as a familiar 2SG form with such verb forms as *hablás* and *comés*, on the one hand reflects

⁵ See <http://talian.net.br> (last accessed 27 April 2012).

features of older Castilian usage (although the *hablås* and *comés* forms are found in fifteenth- and sixteenth-century Castilian as 2PL verb forms, the standard forms today are *habláis* and *coméis*; *vos*, originally a second person plural pronoun, had begun to be used at around the same time with singular reference, as a more polite alternative to *tú*, but has since disappeared); but in River Plate *voseo* the form *tú* and its corresponding verb forms *hablas* and *comes*, which have survived as standard in the peninsula and in many other areas of the Spanish-speaking world, are ousted.

The ‘innovative’ developments in the Romance languages outside the Romània often parallel historical changes elsewhere in Romance. Brazilian Portuguese has a number of such examples in its morphology and syntax. There has been a move away from the use of the definite article with possessive adjectives (*o meu livro* → *meu livro*; see also Ledgeway, volume 1, chapter 8), a change that is only incipient in European Portuguese, where absence of the definite article is restricted to vocative use and nouns which denote close kinship relations, but a phenomenon which is general in French (*mon livre*) and, with preposed possessive adjectives, in Spanish (*mi libro*). There has been a very general movement towards the proclisis of personal object pronouns (e.g., in a simple declarative sentence *João se levantou* corresponding to EurPt. *João levantou-se*), which resembles the changes undergone by Spanish since the fifteenth century – indeed, Brazilian Portuguese has gone further than Spanish in allowing proclisis to non-finite and imperative verb forms (e.g., *Me diga uma coisa* corresponding to EurPt. *Diga-me uma coisa* and Sp. *Dígame algo* ‘Say something to me’; *Meu carro ia se afastando* corresponding to EurPt. *O meu carro ia afastando-se* and Sp. *Mi coche se iba alejando* or *Mi coche iba alejándose*, but **Mi coche iba se alejando* ‘My car was moving away’) (Riiho 1994). Another much-remarked upon tendency in Brazilian Portuguese is the greater use of subject pronouns with verbs, to the extent that it would appear that the null subject parameter is being reset, which makes Brazilian Portuguese more like French, where an overt subject is obligatory and the resetting is complete, than European Portuguese, Spanish or Italian, where the pronoun remains syntactically optional (see Salvi, volume 1, chapter 7). The resetting of the parameter may have to do with the erosion of person–number verbal inflections, occasioned in Brazilian Portuguese by the loss of the second person singular and plural forms through the demise of *vós* (shared with European Portuguese) and *tú* (maintained in European Portuguese) – Brazilian Portuguese makes a second person formality distinction between *você(s)* and *o(s) senhor(es)*, etc., both of which have third person verb forms. Hochberg (1986) found similar

evidence in Puerto Rican Spanish, which she attributed to the fall of final -s in second person singular verb forms, and the French situation has long been attributed to such factors (Harris 1978:75), although this type of account is contested.

The coincidence of a number of features in Latin American and some varieties of Andalusian Spanish (most obviously the absence of an opposition between /θ/ and /s/ and between /k/ and /j/, the weakening of syllable-final /s/, the weakening of /d/ and the absence of the *vosotros* verb form) has led to an *andalucista* hypothesis for the basis of Latin American Spanish, though it is also possible that such 'innovative' developments came about independently or were the product of levelling produced by koineization (for a critical overview, see Lipski 1994:34–62). Brazilian Portuguese, by contrast, appears to be based on the speech of central and southern Portugal and is often represented as maintaining a number of 'conservative' features, such as the absence of palatalization of syllable-final /s/ and /z/, a more distinctive rendering of final /e/ as [i] and absence of centralization of the diphthong /au/ (see Teyssier 1980:99–113). The varieties of French exported to the New World contain linguistic features now moribund in the Hexagon,⁶ such as the use, in spoken Québécois, of *je vas* (rather than standard French *je vais*), the interrogative particle (*je l'achète-tu?* 'will I buy it?') and the use of *s'en venir* for 'to arrive' (*arriver* in standard French) (Lagueux 2005:59). Dialectal diversity may sometimes be accounted for by the origins of the colonizers: thus the variety of French spoken in Quebec, where the original settlers came from different territories located to the north of the Loire, often differs structurally from Acadian French, spoken by some 300,000 people in Canada's four Atlantic provinces (Prince Edward Island, New Brunswick, Nova Scotia and Newfoundland), whose original settlers came from the central-west area of France. For Spanish America, Canfield (1981) proposed a dialect classification based on three arbitrary dates (1550, 1650 and 1750) in the chronology of settlement, the highland areas belonging to the first, the coastal Caribbean area, Louisiana, Ecuador and central Chile to the last, and other areas of Central America and the Southern Cone to the second. In Brazil, the *carioca* (Rio de Janeiro) pronunciation of syllable-final /s/ and /z/ as [ʃ] and [ʒ], as in European Portuguese, has been seen as the result of the closer contact of Rio de Janeiro with European

6 The term 'Hexagon' and 'Hexagonal' are here used, following the use of the cognate words in French, to refer respectively to France (which is vaguely 'hexagonal' in shape), and to the speech varieties of France itself, sometimes also referred to as 'metropolitan' varieties.

Portuguese brought about by John VI's establishment of the capital there during the Napoleonic Wars.

5. Contact-induced change

5.1 'Substrate' influence

5.1.1 Settlement colonies: Spanish and Portuguese in America

We are here considering scenarios in which today, generally speaking, the Romance language is the main or only official language of the country and the sole language of the majority of the population.⁷ Both bilingual and monolingual speakers of indigenous languages belong to the lower socio-economic classes and tend to be inhabitants of rural areas. A particular issue in Hispanic linguistics has been the degree to which changes observable in Spanish in the Americas are due to contact with the indigenous substrates, as opposed to changes which are due to internal evolution.

The length and degree of contact between diasporic Romance and indigenous substrate languages has varied greatly from place to place. The early stages of the Spanish conquest in the Caribbean were characterized by the decimation of the indigenous population through disease and ill-treatment, and the indigenous Arawak language, Taíno, no longer survives; it has, however, provided a number of lexical borrowings, typically denoting fauna, flora and artefacts unknown to Europeans, which have diffused through Spanish to other areas of Latin America and indeed to other languages, e.g., *canoa* 'canoe', *maíz* 'maize', *hamaca* 'hammock', *barbacoa* 'barbecue', *tabaco* 'tobacco', *huracán* 'hurricane', and within Spanish-speaking America *ají* 'chili', *maní* 'peanut'.

The indigenous languages which have survived to the present day vary considerably in status vis-à-vis the colonial language. Very many have few speakers and must be seen as being in imminent danger of extinction: such is the situation in Brazil, for which the Ethnologue database lists some 100 indigenous languages spoken by a total of some 155,000 speakers (a mere 0.1 percent of the total population). Official recognition is to a certain extent an indicator of the vitality of these languages in particular countries (Quechua in Bolivia, Colombia, Ecuador and Peru, Guaraní in Paraguay and Bolivia, Aymara in Bolivia and Peru). However, this ignores some significant survivals: some of the Central American Republics recognize only Spanish officially,

⁷ See also Sala, this volume, chapter 6.

although there are significant numbers of speakers of Náhuatl and Mixteco in Mexico (Mexico recognizes no language as official) and of Mayan in Mexico and Guatemala (only Spanish is official in Guatemala). Official recognition is also sometimes politically motivated. The case of Bolivia represents recent aggressive status planning: the 2009 Constitution mentions thirty-six indigenous languages and places a requirement for at least one other language in addition to Castilian to be used by regional governments; however, the most widely spoken of these languages are Quechua and Aymara. At the top end of the scale in terms of maintenance, bilingualism and status is Guaraní in Paraguay, which is spoken by over 80 percent of the population and where about half the population show individual bilingualism with Spanish (figures based on the 2002 census; see Klee and Lynch 2002); nonetheless, the Paraguayan situation has been represented as one of diglossia, with Guaraní as the low-status domestic language and Spanish as the high-status language of public affairs (Ferguson 1959; Rubin 1968). This situation is due to a number of historical and geographical factors, especially the relative remoteness of Paraguay and the establishment of the Jesuit *reducciones*, which provided a protected environment for indigenous workers and encouraged the maintenance of Guaraní; Choi (2000) also proposed that wars in the nineteenth and twentieth centuries reduced the Castilian-speaking population. The general tendency, however, even with relatively widely spoken indigenous languages, is that of language shift in favour of the colonial language. In Brazil, a lingua franca based on Tupi, known as *lingua geral*, was quite widely used until the eighteenth century, when the expulsion of the Jesuits, who had encouraged its use as they had encouraged the use of Guaraní in Paraguay, and its proscription by Pombal led to its progressive abandonment. The process of attrition of the surviving indigenous languages has also increased in momentum in recent years. For example, while Quechua was actually encouraged as a lingua franca during the early years of Spanish conquest and there is a history of the upper classes discouraging the learning of Castilian by their indigenous workers throughout the colonial period even in the face of successive policies of Castilianization, census figures from Peru between 1940 and 1981 show an accelerating decline in the number of monolingual speakers of Quechua and a corresponding increase in the number of monolingual speakers of Castilian (Klee and Lynch 2002).

There is no convincing evidence of any indigenous substrate linguistic feature apart from lexis having gained acceptance in the standard educated language (what is called in the Spanish-speaking world the ‘norma culta’). There is some evidence of the systematic development of grammatical constructions which

appear to have a plausible basis in a substrate language and which have diffused quite widely in colloquial speech, even in that of some monolingual Spanish speakers. These developments are especially intriguing because such constructions do not appear to be straight borrowings or calques, and are very often consistent with features which already existed in Castilian. To take just one example: in the Spanish of the north Andean area the extensive use of the gerund to form what appear to be modal or aspectual paraphrases has been widely noticed; Kany (1951:211) goes so far as to say that the *dar* + gerund construction discussed below is used colloquially ‘among persons of some degree of culture’. Niño-Murcia (1995) examines six such uses:

1. The imperative of *dar* + gerund, which acts as a polite imperative: *dámelo pasando* = standard *pásamelo (por favor)*. This is the form which Niño-Murcia identifies as being most obviously parallel to Quechua:

tanda-ta	apamu-shpa	cu-hua-i
bread-ACC	bring-GERUND	give-I.BEN-IMP

2. The future of *dar* + gerund is also used as a polite imperative, with clitic pronouns enclitic to the future verb: *darámelo leyendo* = *léamelo*.
3. A ‘perfective’ paraphrase *dejar cocinando*: this has the meaning of standard *dejar cocinado* ‘to leave (something) cooked’ rather than ‘to leave (something) cooking’, which is the standard meaning of *dejar* + gerund.
4. A ‘recent past’ paraphrase *venir comiendo*: *vine comiendo* can mean either ‘I ate before I came’ or ‘I ate as I came’. As with the *dejar* + gerund paraphrase, the usage is not equivalent in function to that of the same paraphrase in the standard language, where *venir comiendo* has the meaning of ‘to have been (repeatedly) eating’.
5. A construction with *mandar* which is apparently lexically restricted to the gerunds *hablando*, *sacando* and *gritando* and has a negative meaning: *me mandó hablando* ‘(s)he scolded me’, *me mandó sacando* ‘(s)he dismissed me’, *me mandó gritando* ‘(s)he shouted at me’.
6. A construction with *poner* (e.g., *puso rompiendo*) which corresponds to standard *lo rompió al poner*. This appears to be similar to the construction with *venir* in that what in the standard language would be the main verb is rendered by the gerund in the north Andean examples.

Niño-Murcia attributes the development of these gerund constructions to parallelism with the Quechua ‘gerund’ or adverbial marker *-shpa-*, which,

crucially, is used in Quechua to indicate both simultaneity and sequentiality in related actions (though only when the subject of the two verbs is identical: a different marker is used when they are not identical, a distinction which is not made, or at least not made in the same way, in Spanish, where, for instance *ver comiendo* means ‘to see (someone else) eating’). Constructions involving such adverbial markers are reported as being very frequent in Quechua. It seems likely that the availability of a number of coreferential subject auxiliary + gerund constructions in Spanish (*estar, ir, quedar, venir, seguir*, etc.) may have propitiated the coining of further such paraphrases, maybe on the basis of *dar* + gerund, all of which show subject coreferentiality, though a sequentiality which is sometimes the reverse of that expected in Spanish (e.g., north Andean *vine comiendo* = standard *comí viniendo*). That some of these constructions are indeed coinings rather than calques is suggested by the fact that their strict Quechua equivalents are not always familiar to modern Quechua speakers.

5.1.2 Exploitation colonies: contact with African languages

There is much obvious evidence of a number of kinds of lexical influence on the Romance languages. There are, as is to be expected, large numbers of cultural borrowings (motivated by a referential gap in the native vocabulary) such as *chikwangue* (a bread-like food made from cassava tubers) (< Bantu; French of the Democratic Republic of Congo); *bifaka* ‘smoked herring’ (< Ewondo; Cameroon French), *kpwata* (beer made from maize) (< Tuki, Bulu, Ewondo; Cameroon French), *ngondo* (a traditional celebration) (< Sawa; Cameroon French); *kakusu* ‘type of river fish’, *jinguba* ‘groundnut’, *muata* ‘chief’ (< Kimbundu; Angolan Portuguese). African words may also become productive in the host language through the addition of a Romance suffix (e.g., *banabanisme* ‘a pedlar’s activities’ (< Wolof *banabana* ‘pedlar’; French of west Africa)) (Lafage 1993:235). Core borrowings (that is, borrowing despite the existence of an indigenous word for the item in question) are also evidenced: *ndjoo* for *gratuitement* ‘free’ (< Duala; Cameroon French); *nnam* for *village* (< Ewondo; Cameroon French); *sumara* for *bouche* ‘mouth’ (< Hausa; Cameroon French); *muxima* for *coração* ‘heart’ or *centro da cidade* ‘city centre’; *maka* for *problema* ‘problem’ or *disputa* ‘dispute’ (< Kimbundu; Angolan Portuguese). Loan shifts are common. These may occur via restriction: *individu* (‘horrible person’, Cameroon French), *politicien* (‘liar’, Cameroon French) (Bilola 2003:110); *graine* (‘peanut’, Senegal French, rather than standard ‘seed’), *charbon* (‘charcoal’, most of Africa) (Lafage 1993:232); via extension: *frère* (any male of the same generation with whom one feels one has

common ties such as family, tribe, district, all countries, rather than standard meaning 'brother') (Lafage 1993:232), *goudron* ('tarmac road or street', all countries, for standard 'tar') (Lafage 1993:232); or for other reasons: *tonton* (StFr. 'uncle') > 'lover' (Cameroon French; Biloa 2003:111); *trop* (StFr. 'too much') > 'very' (Cameroon French; Biloa 2003:111); *gâter* (StFr. 'to spoil') > 'to break, destroy' (Mali French; Canut and Dumestre 1993:223). Calquing is also prominent: *se coudre la bouche* (lit. 'to sew up one's mouth') for standard *se taire* 'to fall silent' (Congo French); *lever quelqu'un* (lit. 'to raise someone') for standard *injurier quelqu'un* 'to insult someone' (Cameroon French; Biloa 2003:113); *gâter l'oreille de quelqu'un* (lit. 'to spoil someone's ear') for standard *mal conseiller* 'to give someone bad counsel', *manger la tête du chat* (lit. 'to eat the cat's head') for standard *être rauque* 'to be hoarse' (Togo French; Lafage 1993:236).

Strong and plausible cases have been made for more structural influence. In Mozambican Portuguese there are a number of possibly related syntactic phenomena: (a) lack of marking of indirect objects by the preposition *a*; (b) passivization of the indirect object (as is permitted in English but not in standard Portuguese, e.g., *O irmão foi concedido uma bolsa de estudos* 'My brother was given a scholarship': standard Portuguese can passivize only on the direct object, *Uma bolsa de estudos foi concedida ao irmão* 'A scholarship was given to my brother'); and (c) constructions with the reflexive in which the subject is the 'beneficiary' or oblique object of the verb (e.g., *A senhora desconfiou-se este senhor* 'The woman was not trusted by this man': standard Portuguese *Este senhor desconfiou-se da senhora*). Gonçalves (1996:38–44) attributes these features to the different syntax of Bantu languages, in which direct and indirect object are not distinguished by nominal markings and the possibilities for what resemble passive structures are also different (in Tsonga, ?*Mali yinyikiwile vatsongwana* 'Money was-given (to-the-)children' is of doubtful acceptability while *Vatsongwana vanyikiwile mali* '(The-)children were-given money' is fully acceptable).

On the other hand, a number of features sometimes attributed to the influence of the indigenous languages may simply be the result of imperfect learning of the Romance language as a second language, with the 'foreigner errors' such a process brings: Mingas (2000:67–70), with reference to the Portuguese of Luanda (Angola), points out the frequency of lack of number agreement between subject and verb and lack of gender agreement within the noun phrase, features which have also been noted for the Spanish of Equatorial Guinea (Lipski 1985) (as also for the dying Spanish of Saint Bernard Parish, Louisiana; Lipski 1990). We can also see what some might

claim was a tendency towards simplification, as in the reduction of the complex vowel system of Portuguese through the loss of the central vowels and the monophthongization of diphthongs (Mingas 2000:64), or the general proclisis of atonic pronouns in contrast to the maintenance of enclisis and mesoclis in certain contexts in standard European Portuguese. At the same time, many of these preferences are also consistent with what we know of the movements observable overall in the history of the Romance languages. Portuguese is atypical amongst the Romance languages in having preserved vestiges of Wackernagel's Law, whereby clitics never occupy first place in a sentence, and the proclisis of clitic pronouns in African Portuguese may be seen in some ways as an expected development and indeed one which is shared with Brazilian Portuguese. The substitution of possessive adjectives by *de* + personal pronoun (e.g., *o livro dela* lit. 'the book of.her' rather than *o seu livro* lit. 'the her book'), possibly favoured by the absence of possessive adjectives in Kimbundu (Perl 1994:102–3), is consistent with the widespread adoption of analytic rather than synthetic structures in Romance (see Ledgeway, volume 1, chapter 8, §2); it also permits disambiguation of the multiple potential references of third person possessive *seu* (singular, plural, third person, polite second person, masculine, feminine) in standard Portuguese. As in the case of Spanish in Latin America, it would be inappropriate to claim that phenomena such as those described above have had any significant structural effect on the Portuguese standard in Africa, since they have low prestige (despite exploitation for literary effect by some creative writers), and European Portuguese has continued to be the reference language used in official documents and cultivated in education.

5.2 'Adstrate' contact

As we have seen (§3.2), European languages have often come into close contact as a result of imperialist expansion. Although the same languages often impacted on one another in a European context, such adstrate influence outside the Romània is different in degree and sometimes also in type.

5.2.1 French and English in Canada

Once again, lexical borrowing is the most obvious result of contact, although it is important to be wary of regarding all anglicisms as having equal status, or as having entered the language in the same way. Mougéon and Beniak (1989) present an interesting case-study of borrowings in their discussion of the arrival of French in the city of Welland (Ontario) after many Québécois left the overpopulated rural territories in search of work in developing industrial

centres. Cultural borrowings from English, which are also common in European French and which may be described as 'international' loanwords which are part of accepted usage (e.g., *football*, *hockey*, *jeans*, *cowboy*), are represented as 'pre-immigration' loanwords imported into Quebec French due to the lack of native equivalents for concepts common in the English-speaking world. Core borrowings divide into three subtypes: (a) pre-immigration borrowings which appear due to the dominance of English in areas such as commerce and industry within Quebec society, e.g., *truck* (standard *camion*), *factory* (standard *usine*), *plugger* (verb, standard *brancher*), which, unlike the cultural borrowings, are considered by speakers as non-standard; (b) pre-immigration borrowings which cannot be so easily associated with specific domains, e.g., *smart* (standard *intelligent*), *anyway* (standard *en tout cas*) – cf. Myers-Scotton and Okeju (1973) – and which reflect the widespread bilingualism at large in Quebec; and (c) post-immigration loanwords, some of which occur due to the particular (precise) connotations of the English word (e.g., *movie*, *high school*) and others which are found in non-specific domains (e.g., *but*, *so*, *dumb*). This last group, in particular, indicates the dominance of English in Welland in domains that, in Quebec, had been under French control; it also reflects the fact that the new, anglicized community in which French speakers were now living had caused an increase in bilingualism.

Less obvious than straight lexical borrowings, but no less due to contact with English, are calques, such as *fin de semaine* 'weekend', *chien chaud* 'hot dog' (Blanc 1993:245), *Ça se comprend réel bon* 'You can understand that real good' (Ryan 1998:101). Calques may also be introduced as a result of language-planning efforts in order to avoid the use of overt English borrowings: for example, the *Office québécois de la langue française* advocates the use of terms such as *gestion des risques* 'risk management', *analyse sociale du cycle de vie* 'social life cycle assessment', *empreinte de carbone* 'carbon footprint', rather than the English expressions themselves.

Data from Romance outside the Romània are a good testing ground for Thomason and Kaufman's hypothesis (Thomason and Kaufman 1988) that borrowing at a syntactic level is possible when contact is intense, as opposed to the view of Lefebvre (1985), Prince (1988) and R. King (2000) that grammatical change which arises in such situations is no more than a consequence of lexical or pragmatic influence. This section will discuss two examples of apparent grammatical borrowing: namely King's study of Prince Edward Island (PEI) French and Mougeon and Beniak's work on the French of Ontario.

Prince Edward Island is one of the four Atlantic provinces of Canada where Acadian French is spoken, the others being New Brunswick, Nova Scotia and Newfoundland. The beginnings of settlement date from the early seventeenth century but the number of French speakers has declined steadily to the point where, in 1991, only 2.3 percent of the 129,756 population still considered French as their home language. Contact with English has led to the widespread presence in PEI French of such phenomena as calquing/borrowing: *tiendre track de* 'to keep track of', *aller in pour* 'to go in for (a career)' (R. King 2000:109). However, King argues that, in this specific context, such morpheme-by-morpheme translations, especially when they involve prepositions, may be having a further, structural influence on the grammar of PEI French in that they encourage preposition stranding and use of the particle *back*. PEI French differs from Quebec French in that, although both varieties frequently borrow verbs from English (*watcher* 'to watch', *checker* 'to check', etc.), PEI French has also borrowed English prepositions which, as (1) demonstrates, commonly undergo stranding and appear also to encourage stranding in French prepositions themselves (2):

- (1) Quoi ce-qu'ils parlent about?
'What are they talking about?'
- (2) Où est-ce qu'elle vient de?
'Where does she come from?' (R. King 2000:136)

Although many intra- and extra-Hexagonal varieties of colloquial French feature 'orphan prepositions' (e.g., *la fille que je sors avec* 'the girl I'm going out with'), King argues that this differs from English preposition stranding (e.g., *Which candidate did you vote for?*), claiming that, in the former, the empty NP position appears to be filled by a null pronoun rather than by the trace of the fronted element (R. King 2000:137). Vinet (1984) also highlights differences in French and English usage, pointing to examples such as:

- (3) J'ai voté pour
(lit. 'I voted for')
- (4) Ce n'est pas facile de parler avec
(lit. 'It's not easy to speak with')

which are grammatical in colloquial French, but not in English. King argues that, although PEI French grammar is, in many respects, similar to Hexagonal French grammar, its tendency to borrow English prepositions has resulted in its preposition set containing a subset of English-origin prepositions. She claims further that the ability of those English prepositions to license

preposition stranding in English is also borrowed, leading to similar stranding with French prepositions, such as:

- (5) Quoi ce que tu as parlé à Jean hier de?
'What did you speak to Jean about yesterday?'

which is acceptable in PEI French but which never occurs in colloquial Hexagonal French. However, according to King, preposition stranding in PEI French does not simply represent direct syntactic calquing, as preposition stranding has been borrowed into PEI French without the constraints that operate in English: hence example (5) is acceptable in PEI French, whereas its direct translation in English ('What did you speak to Jean yesterday about?') is ungrammatical. Rather, then, this stranding with prepositions such as *à* and *de* is interpreted by King as evidence that the syntactic property of preposition stranding has been extended in PEI French from the prepositions it has borrowed from English to some of its other prepositions, possibly by lexical diffusion.

The widely attested presence of the particle *back* in Canadian French (for example, in Ontario (see Hull 1955) and Louisiana where, according to Rottet (2005), it is documented as early as 1932) is also generally attributed to intensive contact with English.⁸

- (6) J'ai l'intention de revenir back à Miami
'I plan to come back to Miami'
(7) I' m'ont donné mon argent back
'They gave me my money back' (R. King 2000:116)

In both of the above examples, *back* has the same meaning (namely 'to return to a former state') and syntactic position as in English. It is therefore usually considered to be an example of lexical borrowing.⁹ However, in PEI French, *back* may precede both the infinitive and the past participle. This even occurs in the speech of older informants who presumably have had less contact with English:

- (8) Puis je voulais back aller
'Then I wanted to go back' (R. King 2000:116)

8 Though see Tremblay (2005) for an alternative explanation, where *back* is interpreted as a relexification of dialectal *arrière*.

9 Massignon (1962) has noted that in Acadian French the particle *back* may also take on the role of the French prefix *re-* (e.g., in *revenir* 'come back'), where the semantic force of this prefix may be felt to be weak.

- (9) J'ai jamais back été dans un pool depuis
'I have never been back in a pool since' (R. King 2000:117)

In this variety of French, *back* can also carry the meaning of 'again':

- (10) Je l'avais assez haï que je l'ai jamais fait back
'I hated it so much I never did it again'

As in the case of preposition stranding above, King argues that this use of *back* is unlikely to represent direct syntactic borrowing since the fact that the adverb must follow the infinitive in English means that an ordering such as 'it is-necessary back to-come' would be ungrammatical (whereas *il faut back venir* is perfectly acceptable in PEI French). King therefore concludes that, in this context, French, rather than English, grammatical rules hold sway and that it is possible to account for the different syntactic distributions of *back* in PEI French via a lexical explanation, namely reanalysis from a preposition to an adverb (R. King 2000: ch. 7).

Mougeon and Beniak's study of the French of adolescents in the province of Ontario (Mougeon and Beniak 1991) gives an important insight into the mechanisms of contact-induced change. In this area, where decline in the use of French has given rise to a linguistic continuum ranging from speakers of monolingual proficiency to those whose speech is more restricted and who use French minimally, there is evidence of two distinct mechanisms of change, termed 'overt' and 'covert' interference. Overt interference produces a new usage in the variety under scrutiny, in which it is possible to observe a qualitative departure from traditional linguistic norms. This is demonstrated with regard to the substitution in Ontarian French of the preposition *sur* instead of *à* to introduce nominal complements expressing the idea of location vis-à-vis the broadcasting media, corresponding to Eng. *on*:

- (11) C'est toute de la musique **su'** la radio
'There's nothing but music on the radio'

A structure such as (11) is relatively rare in the French of Montreal but was found to correlate significantly with the variety spoken by restricted users, and seems clearly attributable to the influence of English. Unlike overt interference, covert interference does not produce a qualitative deviation from the traditional linguistic norm, merely a statistical one. It is more difficult to 'prove' outright since 'it is manifested by the *decline* of a form, which has no counterpart in the superordinate language' (Mougeon and Beniak 1991:160). Their study illustrates this type of interference with reference to the demise of

the preposition *chez* ‘at the house of’ in Ontarian French and the concomitant rise of other more transparent constructions which are more similar to English, such as *à la maison de*. Interestingly, such covert interference was found to correlate most strongly with locality rather than with French-language-use restriction, with unrestricted speakers in Hawkesbury, a strongly francophone area, using *chez* a high percentage of the time compared to unrestricted speakers in the minority francophone localities of Cornwall, Pembroke and North Bay. These results are noteworthy in that although, as expected, the French of restricted users was found to contain syntactic features that deviated from the expected norms, it was also demonstrated that, in those communities where restricted users outnumber the unrestricted users, these innovations can and do spread to the French of unrestricted speakers. Indeed, Mougeon and Beniak demonstrate that although unrestricted speakers in minority francophone localities may shy away from what they perceive as ‘incorrect’ usage (i.e., overt interference such as the use of *sur* rather than *à*), their speech is certainly not immune from covert interference.

5.2.2 Spanish and English in the United States

In the United States, Spanish is strongly maintained in a context in which it is nevertheless not an official language and has no regional standard. Those US heritage Spanish speakers who learn standard Spanish formally in school end up diglossic, regarding the Spanish spoken at home as inferior (Urciuoli 2008). US Spanish is in very close contact with English, which is the *de facto* official language and generally enjoys a higher prestige, which has led to very extensive borrowing from English, to be distinguished from the kind of borrowing from English which is visible globally in the present day. Thus in US Spanish there are many integrated lexical borrowings and calques from English which are not found in other varieties of Spanish and are fiercely disapproved of puristically, e.g., *rufó* ‘roof’, *mapear* ‘to mop’, *troca* ‘truck’, *suceso* ‘success’, *llamar para atrás* ‘to call back’. The influence of English also appears to extend to deeper levels of structure: Otheguy and Zentella (2007) attribute the more frequent use of subject pronouns in the Spanish of New York speakers to contact with English. Yet it is sometimes difficult to judge what it is appropriate to regard as contact-induced change. Silva-Corvalán (1994) found that the extension of *estar* (one of the two copulas for ‘be’) to new adjectival contexts (e.g., *estar joven* in the objective sense of ‘to be young’ corresponding to *ser joven* in the standard language) in the Spanish of Los Angeles correlated with the length of contact speakers had with English: it might therefore plausibly be thought that instability in the distribution of the

two copulas *ser* and *estar* is *prima facie* attributable to the influence of English, which does not have such a distinction in its copular verbs. However, Salazar (2007), who examined the changes affecting the use of *ser* and *estar* in New Mexico Spanish, where the use of *estar* has similarly been extended to new adjectival contexts, concluded that this movement was essentially no different, either in type or in the conditioning factors involved, from those encountered in monolingual speakers in Mexico City and other Latin American cities. Indeed, the encroachment of *estar* on contexts previously reserved for *ser* seems to have been an ongoing direction for the evolution of Spanish for many centuries (Pountain 2000).

5.2.3 Code-switching

Code-switching is another well-researched type of contact which is found widely outside the Romania. It may occur both at the intersentential level, where a switch happens at a clause boundary, thus giving, in effect, alternating passages of speech, each of which involves only one code:

- (12) INTERVIEWER: C'est juste un petit micro, il y a une clip tu peux mettre sur ton gilet là
(‘It’s just a small mike, there’s a clip you can put on your sweater’)
INFORMANT: *I’m a star!* (Poplack 1987:64)

while intrasentential code-switching typically involves shorter passages of speech, often no more than single words (or Embedded Language islands, in the terminology of Myers-Scotton 1993b):

- (13) de quel degré de connaissance **djal** la personne
‘on which degree of knowledge **of** the person’
(French/Moroccan Arabic; Bentahila and Davies 1998:38)
- (14) Faut que tu **pack your own** au ‘Basics’
‘You have to pack your own at ‘Basics’
(Canadian French/English; Poplack 1987:65)

or else donor language material occurring with bound morphemes from the recipient language which yield a mixed constituent:

- (15) arrête de **pick-er** le **ear**
‘stop picking the [your] ear’ (Canadian French/English; Heller 1989).

5.2.3.1 Code-switching in the French-speaking world

Within the French-speaking world, code-switching can have social motivation, helping to index what Myers-Scotton (1993a:7) terms a rights and obligations set

between the participants of a conversation. It may also neutralize potential language conflict (cf. Heller 1988:82, for Montreal, Canada (Canadian French/English)) or have the phatic function of changing the tone of the conversation (cf. Myers-Scotton, 1993a:106, for the Congo (French/Lingala)). However, in contrast to these 'marked' functions, so-called 'unmarked' code-switching may also be found in a stable bilingual community as an integral part of that community's linguistic repertoire where, in Poplack's words, it functions 'as a mode of interaction similar to monolingual language use' (Poplack 1987:53). This provides bilingual speakers who feel they have a genuinely mixed identity with the ability to encode this identity through the use of two languages in the same discourse. 'Unmarked' code-switching therefore contrasts with socially motivated code-switching in that it is not the switches themselves that necessarily have a particular indexicality but rather it is the overall pattern of switching which carries the communicative intention (Myers-Scotton 1993a:117). Two examples of unmarked code-switching are given below.

Swigart (1992:100) predicts that Urban Wolof (a variety of Wolof containing profuse Wolof–French code-switching and spoken by city-dwellers in Dakar (Senegal) 'will become the native language model of many Dakar children in years to come'.

- (16) Mësoon naa lire been *Jeune Afrique Magazine*, non pas *Jeune Afrique Magazine*, benn *Famille et Développement*. Journalistes yu journal boobu ñoo demoon def ay kii, quoi, enquête la, je ne sais pas, ci prison Sénégal. Mais soo gisee, soo ko lire-e, dangay bëgg jooy sax, dangay wax ne jamais, jamais duma def loo xamene dinañu ma japp, yobu ma ci kaso bi. Mais, benn article bu lamentable laa ia wax!

'I once read a *Jeune Afrique Magazine*, no, not *Jeune Afrique Magazine*, [one called] *Famille et Développement*. The journalists from that paper went to do some, uh, well, it was an enquiry. I don't know on Senegalese prisons. Well, if you saw, if you read it, you would just want to cry you would say never, never will I do something that they will get me for and take me to prison. I'm telling you about a really pitiful article [Note that the noun *article* is an established loan and not a code-switch!]

(personal communication between Swigart and Myers-Scotton;
see Myers-Scotton 1993a:124)

The second example of unmarked code-switching comes from Morocco where, although some domains (e.g., the home) are exclusively Arabic and others (e.g., science education) are exclusively French, Bentahila and Davies demonstrate that the phenomenon is frequently found in casual conversation:

- (17) *ʕla:s liʕana les moustiques daba ʕrfti fajɲ huma mxbʕi:n taħt le lit et là, donc pour pouvoir être sûr qu'il n'y a pas de moustiques . . . c'était la poubelle xSSha tkun vide wtanqbD ana tanxarʒ hadʕi kullu et tandir lma alors automatiquement la kajɲ ʕi moustique hnaja mxbaf elle cherche l'ombre, elle fout le camp f la journée.*

'Why? because mosquitoes, now you know where they are hiding, under the bed and there, so to be able to be sure that there are no mosquitos . . . It was the bin, it must be kept empty. I took everything and I put it outside and I poured water over it, so automatically if there is a mosquito here hiding it looks for the shade, it goes away in the daytime.'

(Bentahila and Davies 1983:304)

5.2.3.2 Code-switching in the Spanish-speaking world

5.2.3.2.1 *'Spanglish'* The high degree of Spanish/ English bilingualism in the United States has given rise to both code-switching and language mixing (more frequent switches, often within the same sentence or phrase and certainly within the same linguistic turn; see Auer 1999). These phenomena have attracted a great deal of attention from purists, who often fail to distinguish them from integrated borrowing from English into the Spanish of the US, branding any manifestation of Spanish which contains large numbers of anglicisms as 'Spanglish' (see Pountain 2007). Language mixing has also been extensively cultivated by creative writers for stylistic effect, and it is sometimes claimed that a new language is being constituted (see especially Stavans 2003). However, such mixing is very far from having reached 'fused lect' status in Auer's sense, and choices remain idiosyncratic.

5.2.3.2.2 *Cocoliche* Another well-known case of language mixing, though it is no longer directly observable, is that of Italian and Spanish in Argentina in the late nineteenth and early twentieth centuries, which gave rise to a mode of speech known as *cocoliche*. This is not at all the same kind of phenomenon as Spanish/English language mixing in the US, however. In the first place, the Italian involved was not a well-defined system (Italian arrivals in the River Plate before the Second World War tended not to be native speakers of standard Italian, but of southern Italian dialects). Second, the intention of speakers of *cocoliche* was not to celebrate their bilingualism or to engage in playful invention but to communicate with Spanish speakers within an established Spanish-speaking society. Another consideration is that since the languages involved are cognate Romance languages, the degree of mutual comprehensibility and formal parallelism is high, and so movement between one and the other is correspondingly eased. *Cocoliche* is perhaps therefore

more appropriately characterized as ‘foreigner Spanish’ with a large number of Italian features. But like Spanish/English mixing in the US, this was an unstable system: Meo Zilio (1989:208) speaks of there being as many *cocoliches* as speakers, with a continuum corresponding to the degree of maintenance of Italian in successive generations of immigrants. Although *cocoliche* has now disappeared, it is one of the many ingredients in the lexical stock of *lunfardo*, the urban slang of Buenos Aires, which has a relatively high prestige because of its use in popular tango culture (e.g., *manyar* ‘to eat’ < It. *mangiare*, *laburo* ‘job’ < It. *lavoro*).

5.2.3.2.3 *Fronteiriço* In northern Uruguay, Portuguese is in a situation which partially resembles that of Spanish in the US. When the state of Uruguay was created in 1825, resolving a long frontier dispute between Spain and Portugal, many Brazilian Portuguese speakers remained on the border areas of the new country, just as Spanish-speaking Mexicans were to remain in those parts of Mexico annexed by the US. Portuguese was strongly maintained throughout the nineteenth century, and only with public education and with the improvement of communication with the rest of Uruguay did Spanish come to be spoken. A mixed Portuguese–Spanish mode of speech called *fronteiriço*/*fronterizo* has developed: it does not appear to have the stability characteristic of a hybrid, in that choice between Portuguese and Spanish is unsystematic and unpredictable in purely linguistic terms, but the degree of switching is very high even for a case of language mixing, and it breaks some of the constraints which have been observed for Spanish/English code-switching in the US. For example, the free morpheme constraint (see Sankoff and Poplack 1981:5), which states that a switch may not occur between a bound morpheme and a lexical form unless the latter has been phonologically integrated into the language of the bound morpheme, is overridden in cases where a verb stem from one language combines with a verb inflection from the other, e.g., *chegó* ‘(he/she) arrived’, which has the Portuguese stem *cheg-* ‘arrive’ and the Spanish 3SG.PRT inflection *-ó* (see Elizaincín 1992 and Pountain 2003).

5.3 Hybridization

The most extreme product of contact-induced change is hybridization, or the creation of ‘fused lects’, where there has been a systematic merger of one language with another.

5.3.1 Mixed code / split language

5.3.1.1 Media Lengua

Media Lengua, spoken as a first language in Salcedo, Ecuador, has Quechua syntax and morphology but Spanish lexis (speakers of Media Lengua no longer speak Quechua, but speak Spanish as a second language).

- (18) unu fabur-ta pidi-nga-bu bini-xu-ni
a.favour-ACC.ask-NOM-BEN.come-PRG-ISG
'I come to ask a favour'
- (19) yo-ga awa-bi kay-mu-ni
I-THEM.water-LOC.fall-'come after some event'-ISG
'I have come after falling in the water' (Muysken 1977)

5.3.1.2 Michif

The North American variety known as Michif, spoken, though no longer transmitted as a first language, in the scattered Métis communities of North America (primarily in Saskatchewan and Manitoba in Canada and in North Dakota and Montana in the US) has an overall structure which is closer to Cree than to any other linguistic variety. A split has occurred between, on the one hand, its nouns and adjectives, almost all of which are (Métis) French, and are governed by the rules of French morphology and syntax and, on the other, its verbs, which are derived from Plains Cree, and are governed by Cree morphology and syntax. In his comprehensive description of Michif, Bakker (1997:5) cites a story (an extract of which is reproduced below), that was recounted to him by a native speaker, as a representative example of the language:

- (20) *Un vieux ê-opahikê-t ê-nôcihcikê-t*
An.M old trap-he.CONJ COMP-trap-he.CONJ
êkwa un matin ê-waniskâ-t âhkosi-w,
and an.M morning COMP-wake.up-he be.sick-he
but kêyâpit ana wî-nitawi-wâpht-am ses pièges.
But still this.one want-go-see.it.he.it his.P[OSS] trap.
Sipwêhtê-w. Mêkwât êkotê ê-itasihkê-t, une tempête.
Leave-he. Meantime there CONJ-be.busy-he. A.F storm
Maci-kîsikâ-w. Pas moyên si-misk-ahk son shack. Wanisi-n.
Bad-weather-3. No way COMP-find-he.it. his cabin be.lost-he
Pas moyên son shack si-misk-ahk.
No way his cabin COMP-find-he.it.¹⁰

¹⁰ Like most of the Michif speech community, this speaker is not able to speak or to understand Cree. Unusually, however, and unlike most of his fellow speakers, he has quite a good command of French (Bakker 1997:6).

As Bakker (1997:6f.) comments, from the extract above it is clear that the components of the noun phrase are (Métis) French and that French nouns, articles and possessive pronouns are used in exactly the same way as they would be in French with respect to gender and number. Elsewhere in the same text, French prepositions are also evident and adjectives are inflected for masculine or feminine. In contrast, the verbs are all Cree, displaying, for example, Cree agreement markers (such as *-t* and *-w* for 'he' and *yân* for 'I') and the Cree tendency to differentiate stem forms according to whether the subjects and objects are animate or inanimate, exactly as would occur in Plains Cree dialect, e.g., *wâpaht* 'see it'; (transitive inanimate); *vâpam* 'see him/her' (transitive animate). Strikingly, French elements are pronounced with French phonology and Cree elements with Cree phonology. Clearly, for such a split language to have come into existence, long-term contact must have taken place between the French and Cree speech communities. Bakker describes the thousand or so remaining modern speakers as the descendants of European (often French Canadian) fur traders and Cree-speaking Amerindian women, and suggests that, in all probability, the speech community was never more than two or three times its current size, even in its heyday. Papen (2004) agrees that the mixed code was unlikely to have been spoken by the entire Métis nation, describing it, rather, as a language spoken only by the descendants of bison hunters, and unknown in, for example, Métis communities such as St Lawrence, Manitoba, which had never participated in the bison hunting.¹¹

Although other mixed languages exist (such as Ma'a, spoken in northern Tanzania, which has Bantu morphosyntax and a Cushitic lexis) no other case has been documented of such systematic and regular 'linguistic intertwining' (to use Bakker's term). Myers-Scotton (2002:254–58) suggests that Michif may have developed through a combination of convergence and fossilized code-switching, with Cree as the Matrix Language and Métis French as the Embedded Language, via the gradual fossilization of Embedded Language islands from Métis French. However, Bakker's view is that Michif's precise linguistic structure is likely to have arisen because of the particular typological properties of the two languages in contact (Bakker 1997:4, 214–47). This seems to be substantiated by evidence from the code-mixing that has occurred between Montagnais and French in the community of Betsiamites, Quebec

¹¹ According to Papen, non-bison-hunting communities spoke, rather, a variety of Métis French (described by Papen 2004:112–27) which has phonetic similarities to Louisiana French.

(Winford 2003:192), although, in this particular case, French verbs are also incorporated.¹²

5.3.2 Other mixed varieties

Similar claims for a 'mixed code' have been made for a variety of Acadian French spoken by young people in the Moncton area of New Brunswick. This variety, termed Chiac, has a heavily anglicized lexicon and is described variously as a 'métissage français/anglais' (Perrot 1995, cited in King 2008) and a 'code mixte' (Gérin 1984). Chiac is sufficiently well known to feature in the work of the Acadian writer Régis Brun (for example, in his 1974 novel *La Mariécorno*). Gérin (1984) considers Chiac to be a relexified variety of French with additional structural borrowing, although King (2008) argues that, in terms of contact, it differs little from some other Acadian varieties and should certainly not be discussed in the same framework as Michif. Other so-called 'mixed' varieties of French documented outside the Romània include Frangache (Madagascar; Bavoux 2000:22) and Camfranglais (Cameroon; Biloa 2003:247–77; Chia 1990; Fosso 1999; Efoua-Zengue 1999). However, these varieties seem to be characterized more by their lexis than Michif.¹³

6. Maintenance of language community, standard

The kinds of change described in the preceding sections which have been undergone by those Romance languages which diffused widely outside Europe have naturally tended to result in the creation of new regional varieties, a process which invites comparison with the fragmentation of Latin (cf. Varvaro, this volume, chapter 1), which was similarly a language of empire. It is therefore interesting to observe with what success a sense of language community and a common standard have been maintained.

6.1 Spanish

The impression of unity within the Spanish-speaking world which is given today is in no small measure due to the public statements of its language planning body, the *Real Academia Española*, acting in concert with and on

12 In the mixed code spoken in Betsiamites, French verbs are inserted as infinitival forms which combine with an auxiliary verb *tut* 'do' bearing all necessary inflections.

13 Féral (2010:62), for example, comments that Camfranglais does not constitute a different linguistic system from French and would certainly not be classed by linguists as a distinct language of Cameroon.

behalf of the *Asociación de Academias de la Lengua Española*. The *Academia* describes the Spanish standard quite explicitly as pluricentric: the shared educated usage of any area is to be admitted as standard provided that the unity of the language is not threatened.¹⁴

The latest of the *Academia's* projects, a compendious grammar of the language (Real Academia Española 2009) in fact reads more like a descriptive than a prescriptive document, and even the tone of its prescriptions is gentle. This official attitude has been propitiated by certain linguistic realities, as well as by historical circumstance. In the first place, it has proved possible for a common spelling system to be maintained (with the exception of the use of *x* in Mexico, where this letter represents /x/ in borrowings from the indigenous languages (e.g., *México* itself) as well as the /*(k)s*/ of universally used words such as *examen*, etc.) because it reflects the phonemic distinctions made by educated speakers. For some of these speakers it is not maximally simple – for the majority who make no distinction between /*ʎ*/ and /*j*/ and between /*s*/ and /*θ*/, the orthographic distinctions between *ll* and *y* and between *s* and *c* or *z* respectively are redundant, but they are perpetuated in the interests of those who do. In the second place, there are no really marked syntactic differences between educated speakers; the differences which do exist are readily perceptible and do not create ambiguity (for example, the widespread use of the *cantara* verb form as a pluperfect and general past tense in subordinate clauses in Latin America (Kany 1951), or the use of the *voseo* familiar second person singular in Argentine Spanish, which is represented in printed dialogue and advertisements). Yet such convergence in the educated written level of language is the product of co-operation amongst the national *Academias*, which in turn reflects a conscious decision by influential Latin American language planners, beginning with Andrés Bello's 1847 *Gramática de la lengua castellana destinada al uso de los americanos*, not to permit excessive fragmentation of the language (in the interests of Latin American unity); Bello explicitly referred to the precedent of the fragmentation of Latin which was consequent on the fall of the Roman Empire. Conversely, the *Real Academia Española* has in recent years been at great pains to reflect American usage in its dictionary. The greatest threat to the future unity of Spanish would

14 [...] 'the norm of Spanish does not have a single axis, that of its realization in Spain, rather, its character is polycentric. Therefore, the various usages of the language regions are considered entirely legitimate, with the sole proviso that they should have general currency amongst the educated speakers of the area and should not imply any breakdown of the system as a whole, that is, they should not imperil its unity' (www.rae.es/rae/%5CNoticias.nsf/Portada4?ReadForm&menu=; last accessed 2 June 2012).

seem to lie in the Spanish of the US, where, as we have observed, Spanish does not have official status and is not subject to formal corpus planning in the same way as other areas of the Spanish-speaking world.

6.2 Portuguese

By contrast, many factors in the Portuguese-speaking world militate against unity, or at least unity between Brazilian and European Portuguese (as we have seen, the African Portuguese-speaking countries have tended to follow the European norm as a means of national unification). There is no tradition of co-operation between the Brazilian and Portuguese Academies: there is no Academic dictionary or grammar, although all the Portuguese-speaking countries have recently agreed to implement the common spelling system which was originally proposed in 1990. Between Brazilian and European Portuguese there are quite major and obvious syntactic differences, most notably with regard to personal pronouns, which differ in position and sometimes in form, possessives and the continuous aspect:

- (21) João se levantou (BP) / João levantou-se (EP)
 ‘João got up’
 Me parece que ... (BP) / Parece-me que ... (EP)
 ‘It seems to me that ...’
 Me diga uma coisa (BP) / Diga-me uma coisa (EP)
 ‘Tell me something’
 Não conheço ela (BP) / Não a conheço (EP)
 ‘I don’t know her’
- (22) Meu carro (BP) / O meu carro (EP)
 ‘My car’
- (23) Está escrevendo (BP) / Está a escrever (EP)
 ‘He/she is writing’

Pronoun position has a particularly interesting status in that it has become a powerful sociolinguistic marker which is strongly associated with Brazilian nationalism; awareness of this feature, more than any other, has provided a focus for the insistence on the distinctive nature of Brazilian Portuguese.

The Portuguese situation was different from that of Spanish in a number of significant ways. European learning and culture were not extended to Brazil as they were to the Spanish Empire: there is only a sketchy history of printing prior to the arrival of the royal court in 1808, whereas printing in Spanish America dates from the sixteenth century; while the University of San Marcos in Lima was founded in 1551, followed by many other institutions during the colonial period, it was not until 1808 that a Faculty of Medicine was founded in

Rio. Politically, the Portuguese Empire in the Americas did not fragment in the same way as the Spanish Empire, and so the nineteenth-century argument for a unity based on the European standard advocated by Bello for Spanish-speaking America did not apply. In the twentieth century, literary authors of the Modernismo movement, which originated in São Paulo in 1922, deliberately cultivated a language which drew on spoken Brazilian Portuguese and emphasized the differences with the European standard, thus giving the language cultural prestige.

6.3 French

Despite their French-speaking heritage, the countries of the French-speaking world cannot be thought of as a 'community' any more than those of the English-speaking world. Their different histories of settlement, the different status that French can occupy within these territories and indeed the different nature of the ties maintained with France (or lack thereof) mean that there is little to unite them. The *départements d'Outre-Mer*, for example, are considered as much a part of France as the *départements* in a metropolitan region such as Brittany or Aquitaine, and are governed by the same laws; and, until their relatively recent independence, France also had had a say in the government of its former colonies. In contrast, the era of colonization is more distant in Canada, and the country's greater economic clout means that there is much less apparent reason for it to seek to establish such connections with the Hexagon.

The past century, however, has witnessed the establishment of a number of international francophone organizations and associations. For example (among many others) *L'Association des écrivains de langue française*, the post-colonial successor to the *Société des écrivains coloniaux* (1926); *L'Association des universités partiellement ou entièrement de langue française* (1961), which fosters interaction between academics; and *L'Association de co-opération culturelle et technique* (1970), which encourages cultural and technical co-operation between more than thirty French-speaking states. This ideology of a global French-speaking community (*la francophonie*), which gained momentum in the 1960s, shortly after a number of former French colonies had become independent, was warmly supported by the then presidents of two of these former colonies, Léopold Sédar Senghor (Senegal) and Habib Bourgiba (Tunisia), the latter stating in 1965: 'A nous francophones, seul un Commonwealth à la française pourra donner les énormes moyens nécessaires au progrès de nos élites, c'est-à-dire de nos Etats' ('For us French-speakers, only a French version of the Commonwealth can give us the immense means which are necessary in order for our élites, that is, our

states, to make progress'). Although the movement represented a convenient way for former colonies to maintain connections with France, it is noteworthy that, in order to allay any suspicion that this was an attempt of sorts to re-establish colonial-type links, constituent associations are not juridically based nor do they carry compulsory membership. Nonetheless, even these semi-formal ties inevitably lead to the expectation that cultural aims should be supported by pragmatic assistance and technical aid on the part of France towards, for example, Third World members of *la francophonie*.

Since the 1980s, these organizations and associations have been brought together under the umbrella of *l'Organisation internationale de la francophonie*, which was founded with the aim of uniting and co-ordinating the actions of French-speaking countries around the world. In 1984, President Mitterrand of France set up the *Haut conseil de la francophonie*, an international think-tank generating new ideas for the promotion of the French language and, in 1986, France marked its creation of a Secretary of State for *la francophonie* by a summit meeting at Versailles, *La Conférence des chefs d'états et de gouvernements ayant en commun l'usage du français* 'the conference of Heads of State of governments who share the use of French'. This first meeting has been followed by regular biennial summits of francophone heads of state, hosted not only by countries from all parts of the French-speaking world but also by more peripheral territories such as Romania, Egypt and the Lebanon which, although not traditionally francophone, have long-standing political or cultural ties with France.¹⁵ *L'Organisation internationale de la francophonie* currently comprises fifty-six member states and governments, three associate members and sixteen observers.

French has the reputation of being one of the most homogeneous languages in the world. However, different political realities (the existence of autonomous French-speaking nations with different educational systems) and extensive diatopic variation mean that the actual situation is somewhat different (cf. Lüdi 1992:149). In Canada, for example, two centuries of isolation from France and extended contact with English have led to clear linguistic differences between the local variety of French and that of the Hexagon, such as (i) words being coined in Quebec (e.g., *magasiner* 'to shop', *char* in the

15 The complete list of summits reads as follows: 1987: Quebec (Canada); 1989: Dakar (Senegal); 1991: Paris (France); 1993: Grand-Baie (Mauritius); 1995: Cotonou (Benin); 1997: Hanoi (Vietnam); 1999: Moncton (New Brunswick, Canada); 2002: Beirut (Lebanon); 2004: Ouagadougou (Burkina Faso); 2006: Bucharest (Romania); 2008: Quebec (Canada); 2010: Montreux (Switzerland); 2012: Kinshasa (Democratic Republic of Congo).

meaning of ‘car’) or in the Hexagon (e.g., *frangine* ‘sister’) that are used exclusively within these territories; (ii) words being extended semantically in one or other of the territories (e.g., *chaud* ‘hot’ carrying the additional meaning of ‘drunk’ in Quebec French); (iii) innovations in morphology and lexis, such as the development of a distinct imperfect form of *être* in Ontario (*sontaient*) (Mougeon and Beniak 1991:111f.) or the use of the preposition *clair de* (= *en dehors de* ‘outside of’) in the Acadian variety spoken in L’Isle Madame (Hennemann 2007); and (iv) a divergent pronunciation and morphology (Juneau 1987; Blanc 1993). Such differences have led certain commentators to suggest that ‘it is quite legitimate to call [the French spoken in Quebec] Québécois, whereas it would be incongruous to speak of Belgian or Swiss’ (Rézeau 1987:201). A marked change in attitude has occurred in this respect: in the 1960s and 1970s speakers of Quebec French tended to stigmatize their own variety with regard to Hexagonal French (cf. D’Anglejan and Tucker 1973), and the *Office québécois de la langue française* seemed to be encouraging the alignment of Quebec French with the international standard. However, recent views about Québécois linguistic identity have been more positive, especially in the wake of Bill 101 ‘La Charte de la langue française’ which, in 1977, made French the sole official language of Quebec, and led to an increasingly widespread view of the French of Quebec as an autonomous linguistic community standing in opposition to the French of France, rather than as a North American *francophonie* standing in opposition to English. Baggioni’s list of lexicographical works on Quebec French clearly shows the sea-change that has taken place. Early titles such as *Dictionnaire de nos fautes contre la langue française* (1896) and *Dictionnaire correctif du français du Canada* (1968) become replaced by works such as *Dictionnaire du français québécois* (1985), *Dictionnaire québécois d’aujourd’hui* (1992) and *Dictionnaire historique du français québécois* (1998) as the North American/Quebec lexis becomes increasingly viewed as justifying independent description from ‘European French’ in the same way as Webster’s 1829 dictionary defined American English as a separate norm from British English (Baggioni 1998:263). Local television programmes and usage in schools are now both also being standardized towards the norms of Quebec (Lüdi 1992:164).¹⁶

The more recent implantation of French in Africa, its more widespread status as a second or shared language rather than as exclusively a first language, and the fact that it was brought to the continent as the language of the ruling classes (i.e., the ‘high’ variety in a diglossic situation and typically

¹⁶ For a more detailed account, see some of the chapters in Brasseur and Falkert (2005).

a vehicular or official language rather than a vernacular) mean that, by and large, fewer morphosyntactic differences from Hexagonal French are apparent than with Canadian French. Regional African norms tend to be mainly lexical in nature, arising for the most part from contact with local languages (cf. Diallo 1993:238) and, although one norm generally remains in place for writing, especially within the education system, the linguistic situation of each individual country differs so considerably that it is meaningless to generalize.¹⁷ However, from publications such as the *Dictionnaire du français de Côte d'Ivoire* (Duponchel 1975), it is clear that distinct norms are now being recognized for different varieties of African French.¹⁸

The francophone community is thus unified through the use of a common language and separated through the development of national norms, indices and linguistic variables with which its speakers identify (cf. Clyne 1992:1). For many, this would seem to imply that French may be becoming pluricentric in the same way as has been officially recognized for Spanish (§6.1).

7. Language death/decline in favour of other languages

Although the history of the Romance languages outside the România has generally been one of expansion and diversification, there are also a number of examples of attrition or death. There appear to be two contrasting scenarios: (a) where use of a Romance language is declining because of large-scale contact with another language in a community which is isolated from normative pressures (for example, French in Louisiana and Massachusetts) and the Romance language exhibits many of the linguistic correlates often associated with language death; and (b) where the Romance language starts to become less significant within a former colonial territory following independence: in this scenario (for example, French in the Seychelles, where French may have represented no more than a second or third language for most of the population, only really serving as a family/home language for the ruling classes) its gradual shrinkage and disappearance can occur without manifesting the linguistic symptoms of language death.

The first scenario is typical of the New World. Many North American varieties of French are declining in speaker numbers due to large-scale contact

17 For a convenient one-volume treatment, see, for example, Robillard and Beniamino (1993).

18 According to Lafage (1993:228), Bal (2006) considered more than 12,000 lexical items.

with English. For example, according to Rottet (1998:65), the present generation of speakers of the Cajun French of Terrebonne-Lafourche, whose decline started in the 1940s, will probably be its last (cf. also Valdman 1998); Szlezák (2010) estimates that, despite the fact that it was widely spoken until the 1970s, the French of Massachusetts is likely to disappear in the space of two generations; and the French of Missouri, once heard widely across the Illinois County, is now spoken by only a few elderly individuals (Golembeski and Rottet 2004:141). This type of ‘gradual death’ (Campbell and Muntzel 1989) is also seen in, for example, some Indian Ocean speech communities (see, for instance, Bollée 1993:127), where French is being replaced by English for everyday informal conversations. The disappearing varieties frequently display many ‘hallmark’ features that characterize dying languages. For example, Rottet describes how the French of Terrebonne-Lafourche seems to be becoming structurally closer to the replacing language, English, via an increase of ‘English patterns’ such as non-finite subordinate clauses with overt subjects, with usage increasing according to a clear age continuum:

- (24) Je veux les enfants de êt’contents; mon z’aimerais vous-autres venir
back à 7h30.
‘I want the children to be happy; I’d like you to come back at 7.30.’
(Rottet 1998:84)

In Massachusetts, most speakers of French are already semi-speakers and, although Szlezák (2010:111) states that the variety is disappearing too rapidly to show much sign of morphosyntactic breakdown, influence from English is clearly demonstrated, for example by preposition stranding (see §5.2.1), which occurs to a far wider degree than in Hexagonal French:

- (25) la femme que j’ai parlée à hier
‘the woman who I spoke to yesterday’ (R. King 2000:118)

or by the use of disjunctive pronouns in the place of conjunctive pronouns:

- (26) J’ai rencontré **elle** dans dix-neuf quarante-sept.
‘I met her in 1947.’ (R. King 2000:118)

Similar phenomena are observable in the Spanish-speaking Isleño community of St Bernard Parish, Louisiana, which was the result of emigration from the Canaries (hence the name of the community: *isleño*) between the late eighteenth century and the early twentieth century, at which point emigration ceased. During this time, maintenance of Spanish was favoured by successive waves of settlement by Spanish speakers. But the community’s comparative

isolation from the rest of the Spanish-speaking world, both in Europe and the Americas, improvement of communications and public education, which brought the imposition of English through schooling, together with the migration of many Isleños in search of work, have all militated against its survival. Isleño Spanish displays erosion of verbal inflections, non-observance of gender agreements and simplification of complex syntactic constructions in the direction of parataxis:

- (27) Hay mucha manera loh muchacho salí
 There-are many way(s) the boys leave
 'There are many ways for the boys to leave' (Lipski 1990)

In the above sentence there is no plural marking of the noun phrase *mucha manera*, and the verb *salí* appears to be the infinitive form (minus final *-r*); and there is no subordinating conjunction expressing the relationship between the two clauses. The standard Spanish rendering of this notion would involve a subordinate clause introduced by *para que* and requiring a subjunctive verb: *Hay muchas maneras para que los muchachos salgan*.

The education system may also contribute to the death of distinctive local varieties. In many North American speech communities where French is becoming obsolescent, French-language instruction programmes are used to try to make good the dwindling number of speakers. However, as occurs in Massachusetts, for example, the variety of French that is disappearing and the variety being taught in the schools are often not linguistically identical, with the latter more likely to be closer to the norms of standard French as the school takes over the primary role of language transmission (cf. Thomas 1982, cited in Mougéon 2004:178).¹⁹ A similar trend may be observed in some of France's recent former colonies, where local terms are increasingly becoming replaced with those of standard French, the variety that predominates in the media and in the schools. Charpentier (1993:309) notes how, in Vanuatu, teaching standard French in schools has led to the disappearance of local terms, which now only survive in the context of a lexical gap (for example, with regard to local flora and fauna). *Emboucaner* 'to bewitch' (standard *ensorceler*), *achards* 'condiments' (standard *condiments*), *torcher* 'to light by electricity' (standard *éclairer à la lampe électrique*) represent, among many others, examples of terms which are currently disappearing (cf. Bavoux 2000:13, who reports on the same phenomenon in the French of

¹⁹ Mougéon (2004:179) states that this is not always the case and that the local forms may be preserved if, for example, they are more regular than the standard form.

Madagascar). (Conversely, however, in countries such as the Côte d'Ivoire, the standard French diffused during the colonial period may be replaced by a local variety; cf. Ploog 2010.)

The case of the death of Judaeo-Spanish (§3.5) has more in common with the second scenario than with the first. Where Judaeo-Spanish is preserved (today, mainly Turkey and Israel) it does not exhibit the usual linguistic features of attrition. Numbers of speakers have been reduced by the Holocaust (the city of Salonika (Thessaloniki), which had a population of 50,000 Sephardim prior to the Second World War, now has a Sephardic community of just over 1,000²⁰) and by language shift in favour of the national language as a result of better education: to French (by virtue of the *Alliance israélite universelle*) and then Turkish in Turkey, to Hebrew in Israel, and to English in the United States (Spolsky and Shohamy 1999:228–31). There is also today some sense of a reference standard: the *Autoridad Nacional del Ladino i su Kultura* was created by the Israeli government in 1997 and publishes the journal *Aki Yerushalayim*, which sets one of the standards for spelling (though different spelling systems are adopted by the French *Association Vidas Largas* and the Belgian *Los Muestras*, which have been published since the last years of the twentieth century). In Spain, the *Instituto Arias Montano*, primarily an academic institution, which was founded just after the Civil War, publishes the journal *Sefarad* and Joseph Nehama's *Dictionnaire du Judéo-Espagnol* (Nehama 1977).

²⁰ Foundation for the Advancement of Sephardic Studies and Culture (www.sephardicstudies.org/thes2.html; last accessed 25 April 2012).

II

Creoles

IRIS BACHMANN

I. Introduction

This chapter aims to describe the origins and development of Romance creoles. I shall discuss different theories of creole genesis, including different variations of superstratist, substratist and universalist theories. In accordance with more recent approaches to creole genesis, due attention will be given to socio-historical circumstances of creole genesis and its relation to other forms of contact-induced change such as koineization and partial restructuring. This will raise questions about the nature of language change and the importance of language contact as opposed to language internal change.

Throughout the chapter, I shall adopt a Romance creole perspective drawing mostly on research on French, Portuguese and Spanish creoles. However, as will become clear in the discussion on the history of Pidgin and Creole Studies, one must adopt a comparative creole perspective for some of the discussion on creole genesis in order to reflect developments accurately.

I shall start by discussing the history of Creole Studies and its close relation to the discipline of Romance philology and the paradigm of historical-comparative grammar in the nineteenth century. In this context I discuss claims that creole genesis mirrors the genesis of the Romance languages, an issue under discussion from the early academic descriptions of Romance creoles. The emergence of the discipline of Pidgin and Creole Studies in the context of American Structuralism in the 1960s, however, largely eclipsed the relation to the lexifier language and focused on creoles as nativized pidgins and the characteristics of pidgins and creoles as a typological group. The historical background provides a basis for a critical discussion of different theories of genesis and for the question of whether Romance creoles are to be considered Romance languages. In this context, the notion of creoles as ‘advanced’ representatives of Romance varieties will be discussed together with similar claims for non-standard forms of modern Romance. I shall focus

on recent claims that creolization has been involved in the formation of Brazilian Vernacular Portuguese and non-standard Caribbean Spanish. This discussion leads us to a re-examination of recent models of contact-induced language change through dialect levelling in English and Spanish. I shall show that Romance creoles should be related to these processes of change instead of being set apart from the Romance languages. Their genesis, when compared to other developments in Romance, demonstrates a need to take language contact and contact-induced change more seriously in all languages. What emerges from recent discussions is a need to relate language change to the socio-historical settings in which languages are transmitted and develop. Moreover, our discussion of the beginnings of Creole Studies in nineteenth-century Romance philology and the subsequent naturalization of the concept of language will lead us to a discussion about the process of standardization and its role in language change. I shall argue that models of language change need to take standardization and the respective contact situations between the standard and the coexisting vernaculars such as described under the notion of continuum and diglossia for the creole context into account in order to determine language change as a situated historical process.

2. The history of Creole Studies and its beginnings in Romance philology

2.1 *The Romance model*

The linguistic phenomenon called creole languages came to prominence as a consequence of European expansion from the fifteenth century. The languages of the European colonial powers generally provide most of the lexicon of these languages,¹ while the affiliation of the grammatical structure has been a hotly debated issue in theories of creole genesis. Besides Portuguese, Spanish and French as the lexifier languages for the Romance creoles analysed in this chapter, English, Dutch and Danish also served as a basis for creole languages which emerged from European expansion. In recent decades the discussion has been opened to include pidgin and creole languages, which have emerged from contact between non-European languages alone (cf. Versteegh 2008 for an overview). While descriptions of, and writing in, creole languages date

¹ Therefore these languages are often called the 'lexifier'; another term used is 'base language'. Compare Chaudenson (2001:23f.) for a discussion of the theoretical and ideological implications of both terms in their relation to creole genesis. See also the section on genesis theories. I shall use both terms throughout the chapter.

back longer, academic interest in creole languages emerged in the second half of the nineteenth century in the context of historical-comparative grammar and particularly in the field of Romance philology. I argue that this did not happen accidentally, but followed from the research paradigm of the discipline, in which Romance creoles occupied a clearly defined space. Analysing this research configuration will shed light on the much-debated relation between creole languages and their lexifiers, namely the question of genetic relation, types of language change and the supposed exceptional status of a shift with ‘abnormal transmission’ that they are said to represent (Thomason and Kaufman 1988:146).²

Since Hall (1966) the standard view on creole genesis has been that creole languages arise from a pidgin through nativization, although this view has been challenged in recent years, as will be discussed later. Pidgins are defined as drastically reduced languages that provide a basic communication tool in a contact situation between speakers with different native languages. The massive reductions that characterize the non-native pidgin are thought to be partly reversed by the process of creolization, which involves re-elaboration of grammatical complexity. Interestingly, the distinction between native creole and non-native pidgin was not such an important issue in nineteenth-century linguistic descriptions of creole languages.³ Rather the term ‘creole’, originally used in the Romance context for people born in the New World, was later applied to other things, including some language forms particularly in the French colonial context (Chaudenson 2001:1–13;⁴ Mufwene 1997:37–39; Stein 1998:612). The term ‘pidgin’ on the other hand originated in the English colonial context in Asia and was initially used for speech forms in that region (Mühlhäusler [1986] 1997:1–3). What we find instead in early attestations of these linguistic phenomena is a reference to these varieties as degenerate forms or corruptions of the original Romance language (including *português adulterado*, *corrompido*, *corrupto*, *deturbado* ‘adulterated, corrupted, corrupt, distorted Portuguese’; *español corrompido*, *chapurreado* ‘corrupted, broken

2 Cf. DeGraff (2003; 2005b) for a critical discussion of this theoretical stance, which he calls ‘creole exceptionalism’. While I subscribe to the general critique of DeGraff, I believe that his historical argument relies excessively on the discussion of anthropological, non-linguistic sources for the nineteenth century and that he therefore overlooks the specific place that Romance philology accorded to creoles, as well as its difference from the foundation of Pidgin and Creole Studies in the twentieth century.

3 While Schuchardt (1909) distinguishes between *Handelssprachen* (‘trade languages’) and *Sklavensprachen* (‘slave languages’), he does so by characterizing the languages by their different usages rather than in terms of life-cycles.

4 I cite the 2001 translation, which is a revised version of Chaudenson (1992).

Spanish'; *français bâtard, corrompu, grossier, mauvais* 'bastard, corrupted, vulgar, bad French'). In addition to these terms, the more general terms *gíria, jerga, jargon* (all meaning 'jargon'; see also Trumper, volume 1, chapter 14) were used, and were also applied to the speech forms of delinquents or other socially marginalized persons (Stein 1998:614). These terms were not exclusive to creole languages, but were also applied to other non-standard varieties.⁵ Furthermore, the corruption metaphor was commonplace for the description of the development of the Romance languages from Latin, which were often seen as degenerate language forms in relation to Latin. Fuchs (1849:50), who describes views on the development of the Romance languages since the Renaissance, introduces it as a widely held opinion:⁶ 'The almost general opinion, that is all but forced upon us from most of the earlier views, is that the Romance languages are on a far lower level of development and linguistic perfection than the Latin language, because they are considered mostly as mixed languages and as mere corruptions and deliberate deformations of Latin.'

Apart from the reference to ideas about the superiority of inflectional over analytic languages as propagated by Humboldt ([1827–29] 1996), Fuchs's use of the terms 'mixed languages' and 'corruptions' (*Verstümmelungen*) is striking in its similarity to the stereotypes about creole languages, including the supposed violent nature of their genesis. Of course, Fuchs only cites these views to debunk them as misperceptions that have been proved wrong by the insights of the new discipline of Romance philology, and particularly by Friedrich Diez's (1836) pioneering work which demonstrated that vulgar Latin is the missing link to establish a continuous development of the Romance languages from Latin (cf. Oesterreicher 2000; Lüdtke 2001), thus successfully applying the model of historical-comparative analysis developed by Jacob Grimm for the Germanic languages (Auroux *et al.* 2000:159–62).

It is precisely in this context of Romance philology that the first linguistic analyses of creole languages emerge and the comparison to the Latin/Romance scenario is frequent (cf. examples given in Bachmann 2005:31–33). However, we can note that in most cases some difference is underlined to introduce a fine distinction between the two cases. Hugo Schuchardt,

5 Compare, for example, the debate surrounding a possible fragmentation of Spanish in Latin America after independence and Andrés Bello's, and subsequently Rufino José Cuervo's, use of terms such as *corrupción* 'corruption' and *corruptelas* 'corruptions'. They also draw the comparison with the Latin case (cf. Scharlau 2004).

6 Here and throughout, all quotations from non-English sources will be given in English translation only.

considered the founding father of Creole Studies with the most substantial body of work on creole languages at the time, is an interesting example for this qualified comparison (Schuchardt 1883a:236): ‘This consideration brought me to creole studies or rather brought me back to them; because after Scholle (1869) had rightly rejected the term “daughter languages” for Romance, I looked for real daughter languages and thought to find them in creoles, whose development indicated a break to me, and thus formed an instructive contrast to the development of Romance. In recent times, an interest developed in establishing the influence of pre-Romance languages on vulgar Latin which had been uncritically exaggerated by older scholars and was ignored in comfortable precaution by younger ones, and it seemed to me that one ought to look particularly outside Europe to determine the characteristics and the limits of such influence.’

As can be seen from Schuchardt’s words, the continuity of Romance languages with vulgar Latin is already considered a given at this time; however, the idea of a break in transmission implicit in the corruption metaphor is now applied to creole languages. Even though Schuchardt’s interest lies in establishing the influence of language mixture in all language change against the doctrine of the Neogrammarian sound laws, which are thought to operate purely on an internal level, the idea of continuous sound change was too central to language studies at the time to be tackled head on (cf. Baggioni 2000). Creole languages thus provided a testing ground at the margins of the principal object of research, which lay in the reconstruction of the major European national languages.

2.2 *Creoles as vernacular languages: the context of dialect studies*

However, despite the difference established by Schuchardt between the essential continuity from Latin to Romance as opposed to a break in development from Romance to Creole, Romance creoles were analysed, as the term ‘daughter languages’ indicates, as derivations of Portuguese, Spanish and French, while acknowledging substrate influence. Adolpho Coelho (1880–86) openly blurred this distinction in a series of articles by simply referring to the different overseas varieties of Portuguese including creole languages in his title as ‘The Romance or neo-Latin dialects of Africa, Asia and America’. Schuchardt’s analyses also consist of elaborate derivations of sound changes for which he compares creole languages not only with the standard form of their European lexifier, but also with other overseas varieties as well as older stages of the base languages to situate the creoles in a web of interrelated

language changes (Bachmann 2005:42–45, 55f.). For São Tomense, for example, he lists systematically occurring sound changes of which he says (Schuchardt 1882:896):⁷ ‘Here, as elsewhere, African tendencies have met those of Romance languages, resuscitating old features in a new setting.’ With respect to /r/, he notes the following changes (examples all taken from Schuchardt 1882):

- | | | |
|-----------------|---------------------------------|--|
| (1) Deletion | a. between vowel and consonant: | <i>bodo</i> < <i>bordo</i> ‘bank’ ⁸ |
| | b. between vowel and glide: | <i>stoa</i> < <i>storia</i> ‘story’ |
| | c. word finally: | <i>bendê</i> < <i>vender</i> ‘to sell’ |
| (2) becomes /l/ | a. between vowels: | <i>calo</i> < <i>caro</i> ‘expensive’ |
| | b.1 between consonant and vowel | <i>complá</i> < <i>comprar</i> ‘to buy’ |
| | b.2 with metathesis | <i>flime</i> < <i>firme</i> ‘firm, fixed’ |
| | c. word initially | <i>leno</i> < <i>reino</i> ‘kingdom’ |

Note that Schuchardt does not usually give the base word from which the creole word has been derived through the listed sound changes unless it is an unusual derivation, for which intermediate reconstructed stages are given (*djelo* ‘money’ < **diyelo*, **dielo* < Pt. *dinheiro*); he assumes his audience of Romance philologists to be familiar with Romance languages as a point of comparison. In another case, he describes nasalization in São Tomense and compares it to other creoles he lists (*frominga* < *formiga* ‘ant’) in Papiamentu, giving both old Spanish and Portuguese as possible sources for *formiga*.⁹ In the same article he describes the shift of stress with slight vowel lengthening for São Tomense (*cumó* < *como* ‘how, like’; *cablá* < *cabra* ‘goat’) and compares this with the Portuguese spoken in Angola and the region of São Paulo in Brazil. When discussing the frequent sound changes /ʎ/ > /j/, /θ/ > /s/ and /v/ > /b/, he notes that these are ‘common for Spanish and particularly established in the overseas territories’ (Schuchardt 1884a:141).

This anchoring of Creole Studies to Romance philology can also be seen in the fact that Gröber’s *Grundriss der romanischen Philologie* ‘Manual of Romance

7 I cite the translation of Schuchardt’s text (2008:136) by Tjerk Hagemejer and John Holm in Holm and Michaelis’s edition of key texts in *Contact Languages* (2008).

8 He notes that this change is common in French creoles as well as in many French dialects.

9 In later examples, he gives Spanish base words for Papiamentu forms such as the palatalized dental fricatives in *djes* < *diez* ‘ten’, *djente* < *diente* ‘tooth’. Lenz (1928) follows Schuchardt in acknowledging some Portuguese influence on Papiamentu, while also expanding Schuchardt’s Spanish derivations with a more systematic comparison to varieties of American Spanish (cf. Bachmann 2005:45f.). Lenz, a German Romance philologist who worked at the University of Chile, was in a good position to develop this American dialect comparison.

Philology' intended to publish a section on Romance creoles. While the preface to the first edition mentions a chapter with the aim of explaining the relation of some Romance languages to creole, thus focusing on the genetic status of creoles (Gröber 1888, I:VI), the second edition (Gröber 1904, I:V) announces a chapter dedicated to the description of creole vernaculars (*creolische Mundarten*). Although neither chapter was ever finished, the explicit mention by the editor as well as the change from a more general genetic discussion to the description of individual creoles seems to indicate that creoles were increasingly recognized as a valid object of research and one which belonged to the realm of Romance philology.

This interest in creole languages followed the emergence of dialect studies across Europe from the 1870s (cf. Morpurgo Davies 1998:289) and needs to be seen as complementary to this development. As research on sound change advanced, philologists were looking towards dialects as repositories for possible sound change through which they could demonstrate past developments in real time. Similarly, creole languages could indicate possible future developments in the base languages, which after all remained the main focus of historical reconstruction.

This intimate relationship between creoles and other varieties of Romance manifests itself in yet another way that indicates a different conceptualization of language in relation to its communicative functions. In the French colonial context, the term *patois* was initially also used for creoles as a generic term for the local vernacular (Stein 1998:614; cf. also Mufwene 1997:39). While the term 'creole' was only used in the Romance context, and chiefly for French and to some extent Portuguese creoles, the French term *patois* was used for English creoles in the Caribbean,¹⁰ where the term 'creole' was only introduced in the twentieth century by linguists, who used the term in the context of Pidgin and Creole Studies (Stein 1998:613f.; Mufwene 2000:79; Le Page 1988:30). This usage points towards a conceptualization of creole languages as similar to regional vernaculars, which were undergoing changes in their communicative functions due to the increasing importance of emerging standard languages.

The term *patois* is attested as early as the thirteenth century and initially referred simply to language or a way of speaking, sometimes to natural languages in an affectionate way. As Thomas (1953) shows, the negative connotations of the term *patois* coincide with the progressive standardization of the French language and the concomitant institutional support from the

¹⁰ The use of the term *patois* in its general meaning of vernacular was first attested in English in 1643 according to the *OED*.

sixteenth century onwards, which relegated other regional languages increasingly to the realm of purely oral usage and moreover to being confined to the use of the uneducated as the local elites adopted the Île de France standard. However, it is important to keep in mind for our discussions of creole genesis that until the nineteenth century the acquisition of standard French was restricted to an educated élite, as was passive knowledge of the standard to varying degrees (Posner 1997:83–86). Certainly, in the colonial settings in which creole languages emerged, the metropolitan standard played a minor role (cf. the detailed discussion in Chaudenson 2001:145–53). Mufwene (2000:69) talks about a diffuseness of language practices in the historical settings in which creoles emerged. More detailed studies to determine exactly how the term *patois* was used in these contexts would be needed to trace its semantic changes in different local situations.

Important for our discussion is the conceptual distinction contained in the term *patois* which marks the relationship to the developing new standard: language is conceptualized in terms of the distinction both between written and spoken codes and between educated and uneducated language users. These parameters must have been of particular importance in linguistic communities that were being reconfigured by increasing centralization towards a national standard, and where, therefore, the regional languages were losing some of their communicative functions. Specifically, it meant that these languages became associated with the lower, uneducated classes instead of marking a regional distinction.

Interestingly, it is precisely in this sense that the equivalent term *volkstaal* ‘vernacular’ (literally ‘folk language’) was applied in the ABC islands (Lesser Antilles: Aruba, Bonaire and Curaçao) to Papiamentu at the turn of the twentieth century. It was primarily used in discussions about the spelling systems for Papiamentu and the more general debate about the insufficient knowledge of Dutch among the islands’ inhabitants, and hence the status of Papiamentu vis-à-vis the colonial language Dutch (cf. Bachmann 2007). Putte (1999:21f.) relates this increasing interest in language matters to Dutch nation-building and the role of recently standardized Dutch as the national language. It seems that by the late nineteenth century the important change in the conceptualization of the term *patois* described by Thomas (1953) had gained prominence in relation to underlying changes in language practices such as the increasing distribution of national standard languages through the education system. However, it is interesting to note that it was also central for the discipline of Romance philology from its beginnings in the first half of the nineteenth century. There, the term ‘vulgar Latin’ and the older term

römische Volkssprache ‘Roman vernacular’ (the same term as *volkstaal*), indicate the same conceptual distinction between the language spoken by the common people in the Roman colonies and the literary language widely documented historically. The problem for philologists was how to get hold of these spoken varieties that allowed the reconstruction of continuous language change, but which were by definition not faithfully recorded in writing. The solution lay in documents that deviated from the written standard and thus gave a glimpse at underlying variation.

Scholars working on creole languages at the time encountered a similar problem. While they were able to base their studies on documents written in creole, those were often recorded using the spelling conventions of the base language and often documented by people with different degrees of knowledge of the creole language. Therefore we find ample discussion in Schuchardt, who was a classicist by training and had worked on vulgar Latin, on the nature of the creole texts that formed the basis of his analyses. Consider the following example from Schuchardt (1883b:810): ‘Above all, the orthography is usually extremely inconsistent even within one and the same text. The effort to stick to the Portuguese spelling, for example, prevails to an excessive extent in the translation of the New Testament of 1826. Where they deviate from the tradition, they often do so only half-heartedly, and thus we come across many forms that are neither genuinely Portuguese nor genuinely creole. We have a similar case in our documents. One cannot expect valid results from the attempt to peel off the capricious masquerade covering the spoken sound.’

What is reflected in this citation is a fairly common situation for a non-standardized language, which is written down using the spelling conventions of another language, in this case the related base language Portuguese. Schuchardt must have been familiar with these ‘contaminated’ documents from his work on vulgar Latin and his task as a philologist consisted precisely in establishing the specificity of the language used in the document through critical text analysis. However, in his complaint he points towards a further problem underlying historical-comparative analysis with its increasing focus on sound change. While alphabetic writing in general bears a relation to the sound of language, the Romance spelling systems also follow the morphemic and etymological principle to varying degrees. This creates obstacles for an analysis whose prime objective is the analysis of sound change, even though inconsistencies of spelling can obviously also contribute important information about the actual pronunciation. In addition, the discussion of the creole data point again to the close relation of creoles with their lexifiers, particularly

in those contexts where the two languages coexisted. And even where they did not coexist immediately, a related standardized language could play the role of providing a source for lexical and grammatical expansion (cf. Martinus 1990, and Joubert 2002, for Papiamentu, in which case both the related Spanish and the official language Dutch have played this role). For the philologists, access to creole data was generally guaranteed through local attempts to write down oral creole sources with the help of an established literary language.

This discussion of vernacular versus literary languages in philology points to an underlying tension in the make-up of the discipline that became more prominent towards the end of the nineteenth century, including the increasing drift between literary studies and linguistics. The explanations in Gröber's manual of the discipline *Grundriss der romanischen Philologie* (1888) can be analysed as exemplary evidence. Gröber points out that the principal objective of Romance philology, analogous to that of Classical philology, is to analyse the outstanding works of literature thus tracing the cultural history of a nation whose culture manifests itself in the national language. However, Gröber is aware that in order to trace the history of the Romance literatures one has to deal with their more humble beginnings as vernacular literatures. He therefore discusses in a chapter the so-called oral sources of Romance philology to bridge the gap: 'By oral sources of Romance philology we mean the languages and ways of speaking used for communication among the lower classes, as well as those manifestations of the *Geist* of the Romance peoples dressed in literary forms, which are not disseminated through writing but proliferate and are passed on orally, from one generation to the next. Together with artefacts, practices, customs, pastimes and the way of life of the Romance peoples, they form the basis of Romance folklore and convey knowledge of Romance popular culture' (Gröber 1888, I:197).

Again, we find the distinction between written and spoken code on the one hand and between high and popular culture on the other, the latter being associated with the lower strata of society. Gröber tries to neutralize this difference by relegating the popular culture to the realm of folklore, while reserving the realm of philology for high literature. However, the distinction cut across the literary field, as we can see in the way the Dutch scholar Hesseling, also a classicist by training, and his Curaçaoan collaborator Cohen Henriquez (Cohen Henriquez and Hesseling 1940:162) used the popular character of Cervantes's literature to compare its proverbial nature to their collection of creole proverbs, thus justifying their literary character.

2.3 *The naturalization of language: the vernacular/standard divide revisited*

While literary studies tried to neutralize the distinction between the literary and the vernacular in favour of the literary, linguists began to criticize the value of the literary at the expense of the natural character ascribed to spoken dialects. Not only did they become wary of the written representation of language as discussed, but they were aware of the encroaching character of the standard languages gaining ground and beginning to replace the traditional dialects in some areas.

Hermann Paul's discussion of the German language context in the final chapter of his monumental *Prinzipien der Sprachgeschichte* ('Principles of the History of Language'), the so-called bible of the Neogrammarians, is exemplary in that respect (cf. Paul 1880:266–88). He discusses the standard language mostly in terms of its artificiality as opposed to the natural dialects, which he describes as 'infected' by the standard to varying degrees (p. 274). However, Paul is aware of the distinct quality of the standard language, in which, if learned through formal instruction and used mostly for writing, the impact of sound change on word formation is minimized, according to his view. However, he claims that ultimately, the artificial language may be naturally acquired as the standard gains wider distribution, thus blurring the line between traditional dialects and the written standard. Interestingly, Paul (p. 282) assumes that language mixture between varieties, the dialects and the standard will take place as the standard spreads, thus leading to numerous face-to-face interactions between dialect and standard speakers. This leads, according to Paul, to a number of intermediate forms between the standard and the vernacular forms. On the other hand, he discusses social situations in which the distinction between the standard form and the vernacular remains more stable, because it has not yet penetrated everyday vernacular language practices and is thus acquired only in an artificial way, namely through formal instruction (his example being Swiss German). It is interesting to note that in both cases, Paul refers to the vernacular varieties and the standard as different languages, and assumes underlying bilingualism in these cases (*Doppelsprachigkeit*), which reflects the systemic idea of language proposed by Saussure and which was already present in the later Neogrammarian writings (see Bachmann 2005:121–33). We also cannot fail to notice the similarity to modern discussions of diglossia (Ferguson 1959) and continuum (DeCamp 1971b), which have been used to describe creole/lexifier situations and which are

seen to be marked by a similar difference between stability and increasing interpenetration.¹¹

However, Paul introduces a further distinction to salvage the ultimate dominance of the unconscious language use over the artificial and to minimize the theoretical impact of writing on language development: he draws a distinction between *Schriftsprache* ('written language') and *Gemeinsprache* ('common language'), thus making it clear that the standard language is not simply equivalent to the written standard, but that the 'high' variety exists in both forms. While the written code is always learned in an artificial way, that is through instruction, the spoken standard may be acquired naturally as it gains wider distribution. We therefore need to distinguish on an individual basis how the person has acquired a language in order to know if it is natural or artificial to them. The mixing he describes then only takes place as the standard gains wider distribution as a natural language, so that the standard and the dialects come into contact in everyday communication more and more frequently. The social and media distinction between literary and vernacular language discussed in the previous section is replaced here by a distinction between the acquisition processes anchored in the individual speaker. While Paul maintains the distinction between standard and vernacular (*Gemeinsprache* and *Mundart*), the terms gain a different theoretical status through the emphasis on the mental process of acquisition and the dominance of face-to-face communication in its effect on language change. Paul acknowledges the importance of the written form and its wide distribution through printing (1880:286) in the way it protects the language material from the erosions typical for sound change, which can lead to a reorganization of structural properties. However, he does not see this influence as a productive force, but merely as a slowing-down factor.

Hunnius (1988:346) therefore seems right in tracing the beginnings of the turn towards natural language at the expense of standard languages to the Neogrammarians. However, our discussion shows that they were aware of the conceptual problems involved in delimiting the object of research to natural language and to processes of sound change triggered by face-to-face interaction, while at the same time basing their research on written documents exhibiting a range of language and writing practices.

¹¹ Cf. Auer (2005) for a more recent discussion in the context of European dialects and the phenomena of destandardization.

2.4 *Creoles as fully-fledged natural languages*

The focus on the natural qualities of language and the embedding of processes of language change in the mental processes of individual speakers, which is at the heart of Hermann Paul's *Principles of the History of Language* (1888), despite his problematization of writing and standard languages in the last two chapters of his book, is fully established by the time creoles gain a more prominent place in language studies in the context of American Structuralism and plays a fundamental role in the emergence of Pidgin and Creole Studies as an academic discipline. As Hall's book title *Pidgin and Creole Languages* (1966) indicates, the focus on the natural characteristics of languages goes hand in hand with the distinction between pidgins and creoles as defining stages in a pidgin–creole life-cycle (Hall 1962), which presupposes the existence of a non-native pidgin, drastically reduced in its grammatical and lexical properties, which is then expanded through a process of nativization when the creole becomes the native language of a speech community (Hall 1966:xi–xiii). While creole languages are recognized according to this view as fully-fledged languages in their function as native languages, they remain in a state of exception through their relation to non-native pidgins and therefore to their supposed unnatural beginnings (cf. also DeGraff 2005b:558–62).

It is illuminating in this context to look at some of Hall's other writings of that period, which highlight the role of the mother tongue for the psychological development of a person, and which relate the concept of a speech community to a homogeneous group of native speakers of the same variety (cf. Bachmann 2005:174–79). While Paul talks of the awareness of the communicative function of the standard language and its impact on language practices, Hall's conception of language largely eclipses the material and social distinction between literary language and vernacular, which were still operational in historical–comparative grammar. Although Hall (1952:14), a Romance linguist himself with training in historical linguistics, acknowledges the existence of standard languages, he perceives them as yet another variety, spoken naturally by privileged groups of society. The difference with respect to other varieties lies only in the associated value judgements of the social status of their respective speakers, which are considered extra-linguistic factors. The marginalization of the material distinction between standard and vernacular, which began with the shift in focus on to natural language by the Neogrammarians, is complete in American Structuralism by the time Pidgin and Creole Studies emerges as a discipline.

This wider change in the conceptual underpinnings of language studies allowed the reshaping of the study of pidgin and creole languages as a group rather than as appendices of their base languages. However, this came at the price of conceptualizing them as a special group of languages, which could be defined by their peculiar developmental status.¹² This paradigmatic change led to a number of descriptions of creole languages from a structural perspective, particularly in academic theses, which focused purely on the internal grammatical functions of the analysed languages without making reference to the lexifier and to the creole's historical development (cf. Bachmann 2005:166–74 for a discussion of Papiamentu). Silva-Fuenzalida (1952:12–19), for example, analyses the verbal morphology of Papiamentu (cf. (3a–d)):

- (3) a. Mi **ta** bini
 I *ta* come
 'I come, I'm coming'
- b. Mi **tabata** bini
 I *tabata* come
 'I was coming'
- c. Mi **a** bini
 I *a* come
 'I came'
- d. **Lo** mi bini
 lo I come
 'I will come'

Unlike Lenz (1928:120f.), Silva-Fuenzalida does not consider the derivation of these preverbal markers from Spanish or Portuguese (viz. *ta* < *está* 'stay.3PRS.SG', *tabata* < *estaba* 'stay.PST.IPF', *a* < *ha* 'has.3PRS.SG', *lo* < *logo* 'later').¹³ Instead, she takes all verbal forms she has identified and distinguishes between the inflected forms *ta* 'am, is, are' and *taba* 'was, were', the prefixes *lo* (future prefix) and *a* (past complete prefix), as well as infinitival forms (the invariant verb form) such as *bini* 'to come', *tembla* 'to shake', *tende* 'to understand' and the respective present participles *binyéndo*, *temblándo*, *tendyéndo*.¹⁴ She justifies

12 Mufwene (2000:70f.; 2001:82) speaks of creoles as disenfranchised varieties, for which the distinction of a different group of speakers has led to claims of exaggerated difference from their lexifiers as well as of the non-applicability of ordinary processes of language change.

13 While the three verb forms can be derived from Spanish or Portuguese, *lo* seems to be of Portuguese origin (viz. *logo* 'later'; cf. Sp. form *luego*).

14 She does not mention the past participle formed through stress of the last syllable of the invariant verb forms (*duna* > *duná* 'give') or the prefix *di-* (*fangu* > *difangu* 'catch') or the combination of other verbal markers with *lo* (cf. Maurer 1998:166, 168).

the different treatment of the preverbal TMA (tense, mood, aspect) markers, which even requires her to assume a zero suffix for the past complete and the future aspect, because her whole analysis takes as its starting-point the different status of the verb *ta* (p. 16). She argues that it is the only verb showing inflection and that its inflected forms thus function as ‘auxiliaries to form the present and past incomplete’ (p. 16), as well as serving as copulas (e.g., *es ta malu* ‘this is bad’; *es tabata malu* ‘this was bad’). She assumes *-ba* to be the inflection by which the past incomplete is formed, although the form generally employed is *taba+ta* (cf. Maurer 1998:164). She then derives the present participle by postulating the suffixes *-ando* and *-endo* and an *-i* suffix for verbs ending in a consonant (a stem alternation) with a zero allomorph to derive patterns such as *por* ‘to be able’ and *pudyéndo* ‘being able’. Due to her focus on derivational patterns, she does not consider the different syntactic status of *lo*, which can precede non-stressed subject pronouns (see above).¹⁵

In addition, we see in studies on pidgins and creoles an increasing reliance on English as the academic metalanguage instead of the respective lexifiers, which allowed for a comparison of the data with other, completely unrelated languages. Hall (1966:53) gives an example with his discussion of the Tok Pisin (Neo-Melanesian) verbal marker *-im* together with the transitivity markers on the verb in Hungarian. Of course, these typological comparisons require a common terminology, which is based increasingly on English. Note also the following typological comparison including Romance creole examples from Michaelis (2000:178f.) for copulas developing from subject pronouns, which is done in a similar typological spirit.¹⁶

- (4) a. Joan **e** nha pai (Kabuverdianu)
 Joan cop my father
 ‘Joan is my father’
- b. Ami **I** pursor (Kriyol (Guiné-Bissau))
 I cop teacher
 ‘I am a teacher’
- c. Nem bilong dispela man em Dabi¹⁷ (Tok Pisin)
 name poss this man he Dabi
 ‘the name of this man is Dabi’

15 Compare Maurer (1988) for a detailed analysis of the tense and aspect system in Papiamentu.

16 Michaelis (2000:177f.) convincingly argues for the pronoun origin (*ele* ‘he’) of the copula instead of an origin in the Portuguese copula forms *é* ‘is’ and *era* ‘(it) was’, since the copula precedes the negative particle unlike other verbal particles and triggers the use of a stressed pronoun instead of the clitic forms used in predicates.

17 There is also a verbal copula (*stap*) in Tok Pisin.

d. Malta	hi	gżira (Maltese)
Malta	COP/3SG.F	island
'Malta is an island'		

She argues that the copulas diachronically developed from a resumptive pronoun in a topic–comment structure such as *John, he/this man* and shows that they still show some morphosyntactic properties of the original construction, such as the syntactic position typical of pronouns (Portuguese creoles) and the limitation to third person singular subjects and gender marking (Maltese), while in Tok Pisin the grammaticalization process is just starting from a resumptive pronoun and coexists with the general copula *stap*.

Interestingly, there was also a shift in topographic focus in Pidgin and Creole Studies at the time. While early Creole Studies emerged in the context of Romance philology and focused on Romance creoles, the emergence of Pidgin and Creole Studies in the context of American Structuralism focused initially on Caribbean creoles and their pan-Afro-American dimension. DeCamp's (1971a) historical account in the proceedings of the second conference on Creole Studies held at the University of the West Indies underlines this American focus. While he acknowledges the early work of Schuchardt and Hesselning as precursors, he focuses on the period after the Second World War, when the emergence of Pidgin and Creole Studies makes it possible to define these languages as an independent group and thus as a separate object of research. Instead, within the historical–comparative paradigm, creole languages had to be analysed as derivatives of their base languages in order to enter the research paradigm. For the first time, creole languages based on English became the focus of research, as one can see from the *Proceedings of the Conference on Creole Language Studies* (Le Page 1961), the first of its kind taking place on the eve of the independence of the British West Indies.¹⁸

This focus on the pan-Afro-American dimension was related to efforts to justify an independent status for African-American English as opposed to White American English. Consider the following quotation from Bailey (1965:172): 'I would like to suggest that the Southern Negro "dialect" differs from other Southern speech because its deep structure is different, having its origins as it undoubtedly does in some Proto-Creole grammatical structure. Hence,

18 Van Name (1869–70) even claimed that there were no English creoles, due to the already highly mixed character of English. Statements like these underline the need for more historical research on the perceptions of individual creole languages in their local contexts.

regardless of the surface resemblances to other dialects of English – and this must be expected, since the lexicon is English and the speakers are necessarily bi-dialectal – we must look into the system itself for an explanation of seeming confusion of persons and tenses.’¹⁹

The connection of African-American English to a creole grammatical system serves to give it an independent status from English and attempts to explain supposed deviance from the English language by its roots in a different grammatical system. This line of argument and its relation to language attitudes is also the focus of DeCamp’s (1971a:33) explanation: ‘If American Negro English is indeed a creole with varying degrees of acculturation to standard English, rather than a divergent dialect with varying degrees of deviance from standard English, the social and political implications will be great indeed. [...] If the creole-origin hypothesis for Negro English is confirmed and if the Negro, like the Chicano, really does have a genuine linguistic heritage rather than merely a substandard deviant dialect of English, a great many people, both black and white, will have to revise drastically their attitude toward Negro English.’

Note that the distinction DeCamp makes depends on the possibility of claiming a creole connection for African-American English since that acknowledges its status as belonging to a different language system. Only the status as a different language with its own grammatical system seems to guarantee a ‘genuine linguistic heritage’. It is interesting to compare this discussion with that of Paul, who distinguished the standard language and the dialects as different language systems with the possibility of language mixture occurring between them. DeCamp’s and Bailey’s discussion, on the other hand, seems to take the historical ‘roofing’ of different varieties through standard languages as a purely structural matter. Therefore, they interpret dialects as lying within the same language system, while African-American English could be proved to be fundamentally different in its structural make-up through its relation to a creole grammar and could hence be recognized as a separate linguistic entity and consequently indicative of a different ethnic identity.

It is clear that the distinction between dialect and language is no longer considered in its relation to processes of standardization. The existing

19 The term ‘Negro English’ is common in these early works until the late 1970s. In Valdman (1977), the term ‘Black English’ substitutes for it, to be replaced later by ‘Afro-American Vernacular English’ (AAVE) or ‘African-American English’ (AAE).

categories are taken for granted and translated into structural differences; hence the importance of claiming that creoles are independent language systems, which led to a focus on describing the maximally basilectal varieties. A variety such as African-American English causes problems in this scenario because it rather seems to display a continuum in relation to English and hence goes against the scenario of maximal difference.

Milroy (2001) notes that it is precisely when the standardization of the major European languages has reached a high level and the standard languages were widely distributed across society that linguistics started to ignore the processes that had led to standardization and their impact on the linguistic ecology. This is precisely what my discussion of the history of Pidgin and Creole Studies has demonstrated. However, Milroy (2001) shows convincingly that such views on 'normal' language use are deeply influenced by our own language ecologies in which standard languages play an important role, but that this model is by no means universal. Furthermore, standardization itself brings about changes in language through the elaboration of lexicon and grammatical structure to fit the pragmatic needs of the new fields of usage (cf., for example, Schlieben-Lange 1983; Auroux 1992) and the existence of standard languages combined with widespread literacy have led to a replacement of traditional dialects by varieties that are heavily influenced by the standard (Koch 2004:614), which is a process that Paul (1888) was able to see at work in the mixture of standard and dialect varieties, even if he did not explore its precise implications.

The different perceptions of creole languages discussed here give testimony to the historical grounding of conceptions of language and their importance for the study of languages. Similarly, Le Page (1988:28) points out that both the notion of standard language and of native language are cultural stereotypes, which change historically and from one society to another. He sees the two stereotypes as two sides of the same coin, since they both depend on a belief in largely homogeneous monolingual societies (p. 29f.). Interestingly, while the emergence of creole languages, arguably with different degrees of autonomy vis-à-vis coexisting languages (lexifier or not), falls in the period of a slowly-increased stabilization and promotion of national standard languages in the metropolitan colonial societies, they were yet far from being widespread in their home countries, let alone in the colonies. The colonial societies then at the time of the formation of creole languages had very different language ecologies from today, with many dialects and languages in contact without a clear overarching standard language, allowing for considerable degrees of variation and approximations (Mufwene 2000:73).

3. The genesis of Romance creoles

In this section, I shall show how theories of creole genesis are underpinned by the change in the perception of language as primarily natural language, before I move on to discuss more recent developments in creole genesis theories that tap into a wider discussion of language contact and change. As previously discussed, research on creoles in the nineteenth century largely followed the historical-comparative paradigm and therefore analysed creoles essentially as derivatives of the European Romance languages with a certain degree of influence from their substrate languages acknowledged. While a discussion about language mixture did exist (Bachmann 2005:50–56), this discussion was largely constrained by the assumption of language change as driven by regular and continuous sound change. Thus, predictably, the American linguist Whitney concludes in an article ‘On mixture in language’ (1881) that, in the case of English, language mixture with respect to grammatical structure could not be proved and that any derivational features seemingly borrowed from French such as *pure*, *purity* should be attributed to lexical borrowing of derived forms. However, this point of view begs the question of how else morphological processes could be interpreted by speakers if not through the analysis of repeating structures, thus recognizing the same derivational pattern as in French. What this example shows is that the mainstream of the discipline tended to focus on language-internal processes of change instead of mechanisms of change by external influence from other languages. This was the reason why Schuchardt (1885), in his debate with the Neogrammarians, had to argue such a fine line to demonstrate that, while external influence was possible, it was to be sought in the way meaning was attributed to word-forms in the form of calques or possible elimination of grammatical distinctions where the structure of the superstrate and the substrate language were too distinct. Schuchardt used Humboldt’s term of ‘inner language form’ (*innere Sprachform*), to indicate that these mechanisms belonged to the core processes of language change in a similar way to well-established processes of sound change (Bachmann 2005:53–55). In relation to Philippine Spanish creole (Chabacano), Schuchardt (1884a:127) discusses the origins of the anterior marker *ya*. Comparing the Spanish form *escribió* ‘he wrote’ (stem+3SG. PRT) with the Tagalog form *su-ng-múlat* (conjugated verb form + infix -ng-), he argues that the difference in grammatical structure is too great and that therefore the language learner only recognizes the semantic identity of the expression of anteriority or past tense. But due to the dissimilarity of the forms, he argues that there is a need for the creation of a new form

(*Neubildung*) replacing the infix *-ng-* with the prefix *na-*, which is then transferred to Spanish *ya* ‘already’ and used as a preverbal marker with the invariant verb form as in *ya escribi* ‘he wrote’.

More mechanical and radical theories of language mixture such as Adam’s *Les Idiomes négro-aryen et malé-aryen: essai d’hybridologie linguistique* (1883), in which he proposes that creoles are made up of European lexical items grafted on to an African or Malayan grammar, were the views of an outsider, who had relatively little impact on the academic profession of historical-comparative grammar, and Schuchardt himself rejected Adam’s view on language mixture (Schuchardt 1883a; 1909:443).

3.1 Français avancé or système optimum

Hjelmslev (1938) is the first to put forward the idea of creoles as optimal systems genetically related to their lexifier language. The yardstick for genetic relation is not so much the vocabulary *per se*, but the grammatical building material, namely the lexical elements used to form verb and noun phrases. Hjelmslev argues that there is a complete continuity with respect to these *formants* of the grammatical system because they are all derived from the *langue initiale* (although there are examples of functional substrate morphemes such as the 3PL pronoun *nan* in Papiamentu, which is also used to mark plural). On the other hand, Hjelmslev acknowledges that the underlying creole system is different from the system of the initial or base language. We take his discussion (p. 283f.) of the noun phrase as an example. He notes that Mauritian creole has essentially preserved the form and function of the French indefinite article (*éne*²⁰ < *une* (FSG)), while the definite article is ignored, in his words, often forming part of the noun as in *éne léroi* = ‘a king’ (cf. Fr. *le roi* ‘the king’). He concludes that this leads to a system of two articles as well; however, the distinction is not between definite and indefinite article, but between an indefinite and a zero article, which exhibit the following paradigm, given that obligatory number on the noun does not exist:

- (5) a. *éne dizéf* (cf. Fr. *un oeuf* ‘an egg’)
‘an egg’
- b. *dizéf* (cf. Fr. *(un) oeuf* ‘(an) egg’, *l’oeuf* ‘the egg’)
‘egg, the egg’
- c. *dizéf* (cf. Fr. *les oeufs* ‘the eggs’, *(des) oeufs* ‘some eggs’)
‘the eggs, eggs’

20 Generally spelled in a more phonemic, less ‘Frenchified’, spelling as *en* today.

Hjelmslev postulates the same continuity of building material, but difference of grammatical system for Latin vis-à-vis French. This perception is in line with Saussure's ([1916] 1996) idea of language as a closed system, in which the signifiers derive their value through their difference from other signifiers within the system. With respect to the building materials, this idea implies that while they can be diachronically traced to the lexifier language, their exact function or value within the system (and Hjelmslev reserves the term 'morpheme' for that systemic function) has necessarily changed, since it can only be determined within the system as a whole. If we go back to the example above, we can see how he points with his translations towards grammatical distinctions of noun phrases such as generic/non-generic, specific/non-specific, but notes that the way these distinctions are expressed within the system is different in French compared to Mauritian creole: while French uses the indefinite and definite articles and a partitive construction with *de* 'of', Mauritian creole expresses these distinctions with the difference between indefinite and zero article.²¹

What is difficult to explain in this theoretical framework is the change from one stable system to another, and Hjelmslev's (1938) comments are vague in this respect. While, in general, he argues that genetically related systems are diachronic continuations of each other, which follow along the lines of the inherent dispositions of the initial system (p. 281), he then claims in relation to the change from Latin to French that the latter is 'a system different from the abandoned system born from an abrupt reaction, which led to the extreme opposite' (p. 285). In the case of creoles, interestingly he resorts to a *tabula rasa* metaphor (p. 286), from which a so-called virginal system was created, which presents an optimum of the base language and therefore is related to its essential properties, the 'disposition inherent in that system' in Hjelmslev's words (p. 281). The two examples of the Latin/French and French/Creole genetic relations seem to suggest a weakening of the explanatory power and thus theoretical role of genetic affiliation if it can subsume cases of 'extreme opposite' and '*tabula rasa*' within the idea of genetic continuity. Instead, his discussion seems to indicate that the investigation of language as a closed system gains increasing prominence and thus establishes a focus on how grammatical relations interact within that system.

21 There is also a determiner developed from the postposed demonstrative *-là* (e.g., *sat-là* 'the cat'), which exists in a number of other French creoles. It is generally used as a definite and anaphoric NP marker and often serves to delimit the noun phrase (cf. Posner 1985:180). In American French creole, but not in the Indian Ocean creoles, it can also be used to delimit restrictive relative clauses (cf. the discussion in §4.1).

The discussion of creole languages as a continuation of the structures inherent in the French system, is similar to more recent discussions of popular French as representing the natural tendencies of language change inherent in the French system, while the standard has essentially preserved an older, obsolete state of the language (cf. Hunnius 1988; 2003; Koch 2004). I return to this question in the last section.

I have argued in the previous section that it is precisely the systemic approach of structuralism with its focus on natural language which is at the heart of the emergence of Pidgin and Creole Studies as a discipline allowing for creoles to be perceived as independent of their lexifiers. However, we also saw that the recognition of creoles as fully-fledged autonomous languages went hand in hand with defining them as a group by their development from a previous non-native pidgin. Accordingly, what theories of creole genesis had to explain is the supposed similarity of creoles as a group, instead of any derivational patterns from the lexifier to the creole, as argued by Hjelmslev. A special focus of comparisons lay always on the system of preverbal TMA markers, which are combined with invariant verb forms. Compare the following examples from different Romance creoles for progressive/habitual aspect:²²

- (6) a. I ka manjé (Lesser Antilles French creole)
 3SG TMA eat
 ‘he is eating, eats’
- b. Li pé gèt lisien (Haitian creole)
 3SG TMA watch dog
 ‘he is watching the dog’
- c. Bo ta kaminá (Palenquero)
 2SG TMA walk
 ‘you are walking/walks’
- d. Mi ta lesa (Papiamentu)
 1SG TMA read
 ‘I am reading/read’
- e. N ta bibi binyu (Kriyol (Guinée-Bissau))²³
 1SG TMA drink wine
 ‘I drink wine’

22 Examples from Pfänder (2000:89), Valdman (1978:213), Schwegler (1998:256), Maurer (1998:162) and Kihm (1994:93). Cf. Pfänder (2000:200–9) for detailed tables for French-, Iberian- and English-based creoles and for critical discussion.

23 Kihm (1994:93) notes that Kriyol has two imperfective markers (*ta* and *na*). *Ta* seems to express more the habitual aspect and *na* the progressive aspect. However, he analyses the two forms in terms of a difference in the specificity of the event structure (cf. pp. 93–96).

Furthermore, most creoles show invariant verb forms used without a preverbal marker. Those usually show a distinction between stative and non-stative verbs, whereby stative verbs without preverbal markers are generally interpreted as a present state, while non-statives receive a punctual, completed action interpretation (cf. Holm 1988:150f.).²⁴

Interestingly, Mufwene (1996:87f.) points out that, in the 1970s and 1980s, the genesis debates evolved around substrata versus universals (cf. the title of the volume edited by Muysken and Smith 1986), while superstratist theories were overlooked, despite their prominence between the 1920s and the 1960s. This should not surprise us, given our discussion of the paradigmatic changes underlying the emergence of Pidgin and Creole Studies. Indeed, we can see that Hall (1958), who cemented the pidgin–creole life-cycle in his manual *Pidgin and Creole Languages* (1966), was among the superstratists adopting a similar line to Hjelmslev on their genetic relation, while also acknowledging substrate influence. However, his insistence on Bloomfield's (1933) distinction between non-native pidgins and nativized creoles formed the basis for theories seeking autonomous beginnings for languages that were increasingly seen as a group apart.

3.2 *Ur-Kreole: the monogenesis theory*

The first genesis theory to attempt to define creoles as a distinct group is the monogenetic theory (Whinnom 1956; 1965), which tries to explain perceived similarities across creoles by a common ancestor of all European creole languages in a Portuguese pidgin widely used in West Africa and in Asia, which was then relexified, hence the term 'lexifier',²⁵ by the respective European colonial languages.²⁶ Whinnom (1965:517) mentions particularly the system of preverbal markers with invariant verb forms (cf. the examples above), an intensifier with the meaning 'very' derived from 'too much' (*tu machi* in Pidgin English, *demasiado* or *masiao* in the Philippines and *masiá* in Papiamentu (cf. Sp. *demasiado*), *tro* in Dominican French creole (cf. Fr. *trop*)), as well as some Portuguese functional lexical items (the preposition *na* (cf. Pt.

24 This is, however, not a clear-cut pattern and European languages show the distinction to some degree as well, as Holm (1988–89:151f.) points out.

25 Cf. Goodman (1987:361f.), who gives a good account of the history of the monogenesis theory.

26 This is not to say that Schuchardt had not commented previously on the need to compare Romance and Germanic creole languages. However, there is no systematic attempt at a genesis theory to be found in his work, which is driven, rather, by the overarching aim to define the working of language mixture across languages.

na 'in.the') and the postverbal particle *kabá* to emphasize the completion of an action (cf. Pt. *acabar* 'to finish')).²⁷

The theory originated in Whinnom's analysis of the Philippine Spanish Creoles (1956), in which he establishes a link between four varieties and a Portuguese pidgin, which he thought to be the basis for the first contact between Spanish garrison soldiers in the Moluccas, thus providing a model for the developing pidgin then rapidly creolized and brought to the Philippines. The historical evidence is extremely scarce and Whinnom (1956:10f.) acknowledges as much in the following quotation, which describes the stipulated scenario: 'I have devoted a good deal of space to the hypothetical history of Ternateño, a language of which no text or sample exists, whose name I have coined, and of which no mention is made in any save one document. But Ternateño is important, simply because in the history of any contact vernacular the most important consideration is the problem of genesis, and one fact at least is certain: that the first of the Spanish contact vernaculars in the Eastern Seas arose in Ternate, and had already achieved creolization before the evacuation of the island by garrison and Christian inhabitants.'

His historical argument depends on one document mentioned in this citation, in which reference is made to a 'corrupt Spanish' spoken on Ternate (Whinnom 1956:4). Whinnom interprets this as an indication that there was indeed a Spanish creole spoken in Ternate, since today people from Manila still refer to the Spanish creoles there as 'corrupt Spanish'. There are, however, a number of problems with this hypothesis, which relies heavily on an interpretation of one historical source in the light of the pidgin-creole life-cycle. While Whinnom first cautiously refers to the language in question as a 'contact vernacular' (p. 4), he then adds a scenario by which he assumes the existence of a Spanish pidgin, modelled on the supposedly still existing local Portuguese pidgin, used between the Spanish soldiers and their indigenous wives, which in turn would have been creolized by their children. However, reference to the existence of a 'corrupt Spanish' certainly does not warrant such far-reaching conclusions, as the previous discussion demonstrates. While it is not unreasonable to assume some knowledge of Portuguese in these Christian communities, whose religion, combined with their low caste, had isolated them from their surroundings, it does not seem necessary to seek extraordinary explanations for a rapid language transmission given the intimate family relations of the soldiers with local women. There is also no proof that the children born into these relationships radically altered the structure of

27 Cf. critical discussion in Chaudenson (2001:48f.) in relation to French creoles.

the vernacular spoken between their parents, since the only reference we have is to the existence of a 'corrupt Spanish', which is a term, as we have mentioned before, that was also applied to varieties of Spanish not considered creole. In addition, Whinnom (p. 2f.) himself highlights the social difference between the uneducated garrison troops, often Mexican *mestizos*, and members of the lesser aristocracy, which formed the class of administrators and colonists. These circumstances probably sufficed to explain linguistic and other prejudices, as well as linguistic differences between the two different groups, resulting in descriptions such as 'corrupt Spanish'. What is clear from our discussion is that quite scarce historical documentation has been interpreted in the light of contemporary theoretical notions of the pidgin-creole life-cycle.

Nevertheless, Whinnom (1965) expands his theory of monogenesis eventually to all creoles, adding the mechanism of relexification to account for the existence of French, English, Portuguese or Spanish creoles. He concludes that a stable Portuguese pidgin must have existed in all the Portuguese possessions by the end of the sixteenth century, which was also used as a trade language by other foreigners in the area. While Whinnom (p. 513) acknowledges that many of the documents 'speak simply of "Portuguese"', he still holds that the reasonable assumption is that the documents refer to a 'stable Portuguese contact language'. However, this is the same reinterpretation of historical descriptions in the service of contemporary theory noted before. Furthermore, there is little evidence to propose the identity of the Asian Portuguese variety with that spoken in West Africa other than Whinnom's claim (p. 513) that it spread too rapidly and in too stable a form. He acknowledges that the first texts in a Portuguese creole from India date only from the end of the seventeenth century, while direct evidence of the form of the African and Asian pidgin does not exist. Despite any direct lack of evidence, he further adds Sabir or the so-called lingua franca to the genealogy to explain the pidgin's presumed simplicity and uniformity.²⁸

The sweeping claim of this hypothesis did not attract too many followers, maybe not least because of the change of paradigm in the discipline, which made for a stronger focus on synchronic description rather than historical derivation. In addition, more detailed knowledge of the historical situations in which creoles arose cast doubt on the possibility of general widespread

28 Cf. Schuchardt's (1909) analysis of the historical records available on the lingua franca, which he discusses as yet another example of language mixture without claiming any direct relation to pidgin or creole languages.

knowledge of a Portuguese pidgin among the slaves shipped to the New World. Alleyne (1971:179) cites historical sources pointing out the need for interpreters on the slave ships. However, a number of people have argued for a revised version of the monogenetic theory by pointing out the importance of a Portuguese pidgin for the Spanish creoles in the Caribbean, as well as for the possible existence of a formerly widespread restructured variety of Spanish particularly in the Caribbean region (cf. Schwegler 1999 for detailed discussion).

Goodman (1987) makes a detailed assessment of the Portuguese lexical influence that had been taken as evidence for the monogenesis, and points out how limited it was in most instances or that it could be better explained by other more local influences, including the role of Dutch plantation owners settling in other parts of the New World after their expulsion from Brazil, and the particular contribution of Sephardic Jews with Portuguese in their linguistic repertoire in this context (cf. also de Granda 1974).²⁹ This affects particularly Papiamentu, Sranan Tongo and Saramaccan, where the Dutch had a clear presence in the development of the creole languages. His conclusion is that there might have been some influence of the Portuguese pidgin through the currency it had among people involved in the slave trade. However, these influences certainly do not justify the wide claims of the monogenesis theory.

Furthermore, as Mufwene (1986:130f.) notes, the monogenesis theory does not ultimately explain the emergence of the Portuguese pidgin or the lingua franca on which the development of all creoles ultimately rests in this theory. Naro (1978) seemingly attempts to fill this gap in his 'A study on the origins of pidginization' by claiming that the beginnings lie with the Europeans, who deliberately modify their speech when speaking to foreigners. In a careful analysis of literary depictions of Africans speaking Portuguese, he concludes that the documentation shows clearly that the beginnings of the Portuguese pidgin lie in the hands of the European speakers, who deliberately simplified their speech in interactions with African slaves, who they encountered in Portugal as a result of their coastal explorations. His evidence consists of early records of a stereotyped 'corrupt' Portuguese, which is used in Portuguese plays to depict Africans. However, the lack of similar documents in Africa does not necessarily prove that there was no language transmission at work there, nor that the Portuguese were the ones who invented this pidgin (Naro

29 Chaudenson (2001:48f.) points out in his criticism that *na* is not attested for the French creoles and that the equivalent of the word 'thing' as an interrogative is not used in most of them.

1978:339; '[...] it was the 15th-century European who taught the African "how to speak" [...]', but rather gives testimony to the social conditions of the colonial enterprise, which unsurprisingly did not directly feed into any artistic and literary activities by the mostly illiterate Portuguese involved on the ground. In his analysis of the procedures used by Europeans to modify their speech, Naro (pp. 340f.) then ultimately invokes behavioural principles underlying the formation of pidgins, which he calls the factorization principle, whereby all basic units of meaning are expressed by phonologically separate, stress-bearing units, thus accounting for invariant verb forms, stressed invariable pronouns, a single stressed negative phrase and the loss of the 'essentially meaningless' copula. It is interesting to note that while Naro follows the historical argument and cites circumstantial historical evidence to retrace the Portuguese pidgin, he ultimately invokes universal principles to account for the concrete make-up of the pidgin, which he says are characteristic of the unnatural change that pidgins represent.

3.3 *Nativization runs full circle: the bioprogram and the relexification hypothesis*

The next attempt at a big theory of creole genesis was Bickerton's bioprogram (1981; 1984) and it picks up precisely where Naro's theory of pidginization stops by explaining the natural workings that lead from a non-natural pidgin to a natural, fully-fledged language, which the creole was considered to be. Bickerton's explanation rests on the defining characteristics of creoles as nativized pidgins by claiming that the qualitative differences found in creoles vis-à-vis the pidgins from which they developed were due to the first generation of children who expanded a pidgin into a fully-fledged language by following a bioprogram that represents the minimal characteristics of natural language. This set of default structures is activated when children are faced with the supposedly rudimentary and incomplete linguistic input of their parents' speech.

Bickerton's analysis (1981:7–9) rests on data from Hawaii, where, according to his calculations, a pidgin emerged in the first decades of the twentieth century and a creole developed from that pidgin in the next generation of speakers. Interviews made in the early 1970s would therefore guarantee access to the pidgin variety of older speakers and the creole variety of younger speakers. Bickerton tries to show the fundamental difference between the pidgin and the creole by analysing word order and movement rules, the determiner system, *for-to* complementization, relativization and the system of preverbal markers.

An important point of the analysis was the claim that the exact distribution of preverbal particles (Bickerton still calls them verbal auxiliaries) found in all creoles is part of the universal bioprogram. The preverbal particles express the marking of tense, mood and aspect (therefore TMA) and can be combined – so Bickerton claims – only in that order.

On a structural level, Muysken (1988), among others, raised doubts about the homogeneity of creoles by showing differences between some of the creoles cited by Bickerton in relation to serial verbs and predicate cleft-constructions, which he shows to pattern with substratum influence. However, he also demonstrates (p. 300) that the existing phenomena are more restricted than in the possible substrate languages, thus pointing out the need for clearer mechanisms on substrate influence. Even the prototype of the TMA system did not stand up to closer examination: there are postverbal markers (for example, *-ba* and *-ja* in the Upper Guinea Portuguese Creoles and *-za* in the Gulf of Guinea Portuguese Creoles; cf. Holm 1988:267), and one of these, *-ba*, is a tense marker, thus clearly violating the presupposed order TMA when combined (cf. also Maurer 1998). Moreover, more detailed analyses of the verbal systems and their semantics for expressing the anchoring of an event indicate that these systems show much more variation than a minimal bioprogram hypothesis would suggest and rather point towards an intricate relation between substrate and superstrate influence, as well as independent grammaticalization processes that follow more general trends of language change.³⁰

A more general criticism levelled against the theory concerns the way in which it reifies the notions of pidgin and creole. Mühlhäusler ([1986] 1997) demonstrates that expanded pidgins such as Tok Pisin show a remarkable degree of stability and structural complexity, which matches that of creoles. This proves that the structural developments of creoles cannot depend purely on first language acquisition by children. In addition, the socio-historical circumstances in many plantation economies suggest that children could not have played the determining role, given that the constant influx of new slaves and the high mortality rate suggest limited possibilities for child rearing (Arends 1995; Mufwene 1996). Thomason and Kaufman (1988:164) highlight the improbability of the whole linguistic scenario of generationally divided speech communities as follows: ‘When Bickerton poses the question of how a

30 Compare, for example, Pfänder (2000) for French creoles; Kihm (1994) for the Upper Guinea Kriyol; also Baker and Corne (1986), who relate the complex system of distant and proximate past and future in Mauritian creole to Bantu influence; and Detges (2000) for a discussion of different processes of genesis for preverbal markers in French creoles.

child can “produce a rule for which he has no evidence” (1981:6), he is, in our view, asking the wrong question. We prefer to ask how a child can create grammatical rules on the basis of input data received which is much more variable than the input data received by child in a monolingual environment.’

This unlikely scenario not only overestimates monolingual scenarios as the norm (cf. Le Page 1988:29), but also stipulates an essential stability of adult language behaviour, which seems to be incorrect (Thomason 2008:248).

While universalist and substratist have often been on opposite sides theoretically (cf. the introduction in Muysken and Smith 1986), the relexification theory forms the perfect complement to the bioprogram hypothesis in making differences between non-natural and natural language acquisition as solely responsible for the genesis of creole languages. While Bickerton sees the natural L1 acquisition of children with insufficient input from their parents’ non-natural pidgin as the trigger for creole genesis through the activation of the bioprogram, Lefebvre claims that the carry-over of substrate structure is a necessary result of the restrictions of L2 acquisition, in which the foreign lexical input is processed strictly through L1 structures. She thus claims that in L2 acquisition, which creole genesis ultimately represents, learners will acquire the lexical forms of the target language only (hence relexification) while maintaining the syntactic and semantic properties of their native language (Lefebvre 1998:9). While relexification is not the only process in creole genesis according to this theory (reanalysis and dialect levelling are assumed to act on the lexical entries created by adults through relexification), she does claim that all the essential grammatical properties of the creole are ultimately derived from the substrate. On a par with the bioprogram hypothesis, this strong claim seems too exclusive, and has also led to criticism (cf. DeGraff 2005a; Mufwene 1996; 2001).

Most other proponents of substrate theories had more modest aims (cf., for example, Holm 1988–89) and it seems right to acknowledge that their perspective has long been ignored (cf. also the discussion on substrate influence in the historical–comparative paradigm). Alleyne (1986:303), one of the proponents of substrate influence in the genesis of creole languages, points out that the focus on ‘concepts such as degenerate pidgins preceding creoles, superstrate languages as historical points of departure, children as the generators of creoles, the disuse of indigenous/substrate language, and Universal Grammar as some kind of tyrannical force rather than an inferred abstraction’ have put researchers arguing for substrate influence on the defensive. While Alleyne is certainly right about the long-standing neglect of substrate influence, he ignores the appeal of Bickerton’s theory, as well as the monogenetic theory,

which lies in the fact that it tries to explain the existence of creole languages as a group, which strengthens the coherence of the discipline of Pidgin and Creole Studies. Substrate theories generally go against this uniformity since they must predict important differences between Pacific and Atlantic creoles.

While Alleyne (1980) is most interested in the African contribution to Caribbean creolization (*Comparative Afro-American* is the title of his influential study) and thus continues the focus of the early twentieth-century Pidgin and Creole Studies and their local focus on the Caribbean perspective, Holm (1988–89) relates the study of substrate influence to systematic differences between Atlantic and Pacific varieties. Similarly, subgroupings for creole language with the same lexifier according to the differing substrate influence have been made for French creoles (Baker and Corne 1986), for Portuguese creoles (Stolz 1998) and for Spanish creoles (Lorenzino 1993).

3.4 *Modifications of the pidgin–creole life-cycle*

In more recent years, it seems that creole studies have shown less appetite for grand schemes,³¹ and Alleyne (2000:125) summarizes the general mood neatly as follows: ‘One must admit that there has been of late a waning in what was a relentless search for a single hypothesis to account for the genesis of “creole” languages. There has been a growing awareness that these languages, at least those of the Caribbean, demonstrate well-known, “normal” processes of language change, both internal non-motivated and externally motivated change; rather than, or at least in addition to the special exclusive process or processes that traditionally have been assumed to be involved.’

Detailed historical research has shown that the strong hypothesis of a deterministic pidgin–creole life-cycle largely driven by a one-generational language creation process has little foundation in most of the socio–historical settings where creoles emerged. Alleyne (1986:306) points out that the idea of a simple pidgin of adults which children transform according to their cognitive needs is much too simplistic and goes against the historical records, which paint a much more nuanced picture of different levels and styles of language use by different groups of slaves (p. 307). Both the bioprogram and the relexification hypothesis seem to stem from a conceptualization of speakers as essentially monolingual: on the one hand Bickerton assumes that the slave populations immediately abandon their native languages and are therefore monolingual in pidgin, while the relexification hypothesis assumes that the native language

31 See, however, McWhorter’s (1998; 2001) discussion about the creole prototype and its supposed simplicity.

essentially determines the structural make-up of the emerging creoles, therefore assuming that speakers can only gain very limited access to languages other than their native language. Both ideas seem to place too much emphasis on the native language and ignore existing multilingualism in many parts of the world, which makes language contact a much more frequent scenario than previously acknowledged. This has led to an increase in research on language contact, in which pidgin and creole languages are seen as part of a range of contact phenomena from borrowing and code-switching to language shift and contact-induced change (cf., for example, Thomason and Kaufman 1988).

Other research has demonstrated the importance of more accurate socio-historic contextualizations and the importance of demographics. Chaudenson (2001:96–129) introduced the important distinction between the early home-stead phase and the later plantation phase, which marked quite different levels of population distribution and hence access of slaves to the language of the colonizers. It was thus only after the plantation economies were fully established that the ratio between black and white people changed dramatically. Chaudenson (2001) concludes from this distinction that the creoles were formed as approximations of approximations of the superstrate language as direct access to it became increasingly more difficult.

Furthermore, Arends (1995) points out the internal hierarchies within slave plantations, which explain that there were different levels of access to European culture by different slave populations, and that the more integrated house slaves and overseers probably played an important role in the transmission of language to field slaves. Similarly, Chaudenson (2001:89–94) describes the cultural and linguistic adaptation of new slaves, which clearly shows that linguistic integration was an important aim and an organized process, at least in some areas. Mufwene (1996) uses detailed information on population statistics in the making of Caribbean creoles to discuss different outcomes of the restructuring process.

Even the widespread claim that the slaves had no way to communicate in African languages due to the multiplicity of languages and the policy of ethnic mixing of slave populations to avoid revolts has received some correction. Arends draws on historical work (Eltis 2001) identifying systematic geographical patterns among the slave populations destined for different regions, observing how ‘rather than drawing slaves from all over coastal West and West Central Africa, specific regions in the New World recruited their slaves from specific regions in Africa’ (Arends 2008:313).

This gives us a clearer idea when looking for likely substrate influence. On the other hand, it remains a point of debate how long African and other

languages coexisted in exogenous plantations. Bonvini's (2008) analysis of the Brazilian situation demonstrates the lengthy retention of African languages in Brazil, partly due to missionaries' language policies both in Angola and Brazil to use general languages, similarly to the general Amerindian languages used in Ibero-America. This adds a further factor to determining different outcomes of creolization and highlights the need for a more realistic picture of the multilingual beginnings of creole languages.

4. Scales of creoleness: decreolization, partial restructuring and theories of language change

In this section, I discuss recent issues in Pidgin and Creole Studies which seem to me most relevant to the history of Romance languages and to general theories of language change. Despite Thomason and Kaufman's (1988) insistence on the impossibility of defining the genetic relation of pidgin and creole languages due to their irregular pattern of transmission, the most interesting contributions to Pidgin and Creole Studies in recent years have been investigations which relate language change in creoles to that of other languages and thus ask questions about the nature of language change and the role of contact therein. In fact, this is precisely what Thomason and Kaufman (1988) do with their book-length presentation of contact phenomena, even if they try to define just how much contact-induced change a normally transmitted language can undergo before it becomes genetically untraceable. Thomason (2008:257) explicitly discusses the difficulty in defining 'the borderline between a mixed language, on the one hand, and on the other hand a language that is not mixed under this definition but that has undergone quite a bit of ordinary contact-induced change'. However, she ultimately maintains the position that mixed languages such as pidgins and creoles cannot be genetically traced and are therefore unclassifiable to historical linguistics. It seems, however, unsatisfactory that contact-induced change, which is frequent, as Thomason and Kaufman (1988) show, remains an exceptional phenomenon that does not fit the model of historical linguistics. It seems to me that historical linguistics requires a theory of language change that is able to account for all instances of language change. Recent developments in research appear to focus precisely on the borderline areas discussed by Thomason by looking at 'intermediate' forms of Romance and other creoles, as well as contact-induced change in 'non-mixed' languages. I shall examine some recent developments which try to shed new light on processes of language change as well as research on 'intermediate' forms of Romance creoles. This discussion will

lead us to reconsider the question of whether creoles can be considered Romance languages, or rather, what their status in the history of the Romance languages is.

4.1 *Koineization and degrees of restructuring*

In a provocative article entitled 'Jargons, pidgins, creoles, and koines: What are they?', Mufwene (1997) calls into question the fundamental difference between pidgins and creoles, as opposed to other varieties of European languages for which some contact-induced change has been claimed. In particular, he questions the usefulness of the term *koiné*, which has had a revival in recent socio-historical research showing the influence of dialect mixing for processes of levelling and simplification in some urban and colonial varieties (cf., for example, Trudgill (1986; 2004) for English or Tuten (2003) and Fontanella de Weinberg (1992) for peninsular and American Spanish respectively).

Tuten (2003:25) also discusses the relationship between koineization and creolization and tries to mark a clear difference between them. He suggests that, unlike creolization, difficulties in communication were not an issue in koineization due to the mutual intelligibility of the dialects involved. However, mutual intelligibility is a quality which is hard to pin down. It has often been taken as a criterion for differentiating the lexifier from its creole varieties by claiming mutual intelligibility of the latter, but not of the creole varieties with the lexifier, but this claim has likewise been repeatedly disputed. Chaudenson (2001:40f.), for example, notes that often the French standard has to serve as a medium of communication for creole speakers from different countries, which is precisely the function one would expect from a shared standard language.³² Mufwene (1997:45) points out that some of the so-called *koinés* have arisen from the contact between speakers who clearly do not consider their languages as related to each other and hence certainly not as mutually intelligible.

Mufwene (1997) argues that the possibility of distinguishing between *koinés* and pidgins and creoles depends on the possibility of drawing a line between languages and dialects, which has proved notoriously difficult on purely structural grounds. Furthermore, Mufwene (2001) points out that unequivocal multilingualism, and therefore contact between clearly unrelated languages, has also played a role in the formation of European languages and their colonial varieties due to population movements and migration; he therefore

³² Cf. Mufwene (2001) for a genealogy of different varieties of English according to their communicative functions.

proposes a unified approach to all these varieties as contact-induced restructuring. Interestingly, very similar principles for different types of contact-induced change have been proposed independently in the past. Siegel (1997) compares a number of approaches to pidgin and creole languages, second language learning and interlanguages, indigenized varieties and *koinés* and shows that they converge on principles such as unmarkedness, regularity, transparency and frequency. The mixing process is mostly seen as a pool of competing features, from which some are selected or levelled out according to the above-mentioned criteria. This idea is probably easier to conceive of for phonetic and phonological change; however, research on substrate transfer of morphosyntactic structures (Boretzky 1983) shows that there needs to be some match of features or structures for them to be selected into the contact variety. Furthermore, theories of koineization and creolization also coincide in claiming a process of autonomization (Chaudenson 2001) or focusing (Trudgill 2004). Chaudenson interprets the fact that these varieties were termed *créoles* at some point, rather than the previous label of 'corrupt French', as an outer sign of that process. Trudgill (2004) takes up Le Page and Tabouret-Keller's (1985) insight that a variety needs to become a point of communication and identification for a community in order to jell as an independent language.

Taking Mufwene's argument seriously means that the historical comparison with the varieties of the base language that were present in the language mix at the time of contact comes to the fore, but the possible contribution from substrate languages also present in the contact situation is not neglected. Universal principles such as transparency and unmarkedness would come into play with respect to selection of some features over others. Posner (1985) follows this path of research with an analysis of complementizer and relative clauses in French creoles compared to seventeenth-century popular French and popular varieties of Canadian French. Her analysis shows that with respect to complementizers, French creoles clearly follow the tendency of popular seventeenth-century French for deletion:

- (7) On disait c' était un vieux bâtard (Posner 1985)
 One said it was an old bastard
- (8) Moin coue i bon (Jourdain 1956, cited in Posner 1985)
 Me think he good

With respect to relative clauses, she notes that French creoles have followed the tendency of popular French and other Romance languages for generalized relative complementizers. However, while Montreal French shows only occasional deletion of the relative complementizer, French Caribbean creoles

show quite systematic deletion except in subject position. Posner (1985:179) argues that this development can be explained by the Accessibility Hierarchy, and is a change which is in no way specific to creolization (cf. modern English). The Indian Ocean creoles seem closer to popular French with the widespread use of the general *ki* complementizer with only occasional deletion in non-subject position.

In addition, she points out a further grammaticalization process in the Caribbean creoles (including Guyana), but not in the Indian Ocean creoles. In the Caribbean varieties the definite postverbal marker *la*, derived from the French distal adverb *-là*, is used to delimit relative clauses by analogy with the demarcation of the NP. Compare the three examples below, cited in Posner (1985:176, 180):

- (9) a. Nõm la moin ka pale a ... (Martinique; Jourdain 1956:309)
man =la me PRG speak to
'the man I'm talking to'
- b. Fãm -lã blãf (Haiti; Sylvain 1936:41)
woman =la white
'the woman is white'
- c. Fãm blãf -lã
woman white =la
'the white woman'

Posner (1985:182) notes that this grammaticalization process is furthest developed in Haitian creole, where the construction is used now for all restrictive relative clauses substituting for the more 'Frenchified' model. She cites (pp. 18f.) substrate influence from Ewe, which has a postverbal definite marker *lá* also delimiting restrictive relative clauses, in combination with the widespread use of the deictic *lá* in Quebec to mark the end of sentences as possible sources for this development.

This type of careful cross-dialectal historical analysis, which takes substrate influence into account, was also at the focus of a volume edited by Neumann-Holzschuh and Schneider on *Degrees of Restructuring in Creole Languages* (2000). A number of the contributions in this volume deal with restructuring processes in subsystems of the grammar of creole languages in a similar way to Posner (1985) and carefully delineate similarities and differences of scale in individual processes. Neumann-Holzschuh (2000), for example, shows the relatively conservative character of Louisiana creole by analysing historical and modern data from different French creoles. Compare the following examples (p. 394) on the position of the negative particle *pa*:

- (10) a. Chimen té pa bon (Guyana creole, 1872)
 way TMA NEG good
 ‘the way was difficult’
 b. afin i té pas laide (Guadelope creole, 1923)
 at-last 3SG TMA NEG help
 ‘in the end, he did not help’
 c. mé roi là té pas olé (contemporary Lousiana creole)
 but king =DET TMA NEG want
 ‘the king did not agree’

In modern Lousiana creole the anterior marker precedes the negative particle and this form is also attested in older texts from Guyana and Guadelope. The generalization whereby the negative particle precedes all TMA markers in most modern French creoles therefore seems to be a later innovation, which Lousiana creole does not share.

Michaelis (2000) shows different stages of a grammaticalization process for subject pronouns developing into predicate markers in Seychelles creole and Réunion creole. She sees the starting-point for this type of development in left-dislocation constructions, where *Pierre, il chante* ‘Pierre, he sings’ develops into *Pierre i chante* ‘Pierre sings’, and cites historical evidence for examples with a resumptive pronoun (p. 167). In Seychelles creole, the marker is obligatory in all third person contexts with a full NP subject, but not with personal pronouns (pp. 164f.). Compare the following examples:

- (11) a. Lakaz i zoli
 house TMA nice
 ‘the house is nice’
 b. Marcel i dir (. . .)
 Marcel TMA says
 ‘Marcel says (. . .)’
 c. Bann dan mi danse
 PL woman TMA dance
 ‘the women dance’

However, she notes that the marker *i* does not occur in the context of negation or together with other verbal particles (p. 165).

In Réunion creole, on the other hand, *i* combines with all persons and with both full NPs and pronouns, some of which contract with the particle (cf. the following paradigm from Michaelis 2000:168):

1SG	<i>mi manz</i> (<i>mi</i> < <i>moin</i> + <i>i</i>)	1PL <i>nou i</i> (<i>ni</i>) <i>manz</i> ‘I eat’, etc.
2SG	<i>vi manz</i> (<i>vi</i> < <i>vou</i> + <i>i</i>)	2PL <i>zot i manz</i>
3SG	<i>li manz</i>	3PL <i>zot i manz</i>

As in Seychelles creole, the particle cannot be used together with other verbal particles. She argues that the distribution of the particle *i* in the two French creoles can be interpreted as ongoing grammaticalization which functionalizes a high-frequency pronoun (3SG) as a verbal marker to indicate present tense and finiteness. While in Seychelles creole this process has not expanded to all persons yet and is restricted to affirmative contexts, in the creole of Réunion the marker has lost all traces of agreement and has developed into a fully-fledged present-tense marker.

Neumann-Holzschuh (2000:403) concludes her analysis by pointing out that the problem of scale of difference from the base language is similar in the case of creole languages and Romance languages, where the distance to the base language can only be defined in relation to individual parts of the grammar. Defining a prototype is elusive (cf. Thomason's criticism (2008:244) of McWhorter 1998), since the changes take place in different subsystems of grammar and can have different extensions across dialects. However, careful analysis and comparison with developments in related and unrelated languages certainly provides a rich testing ground for historical developments in the Romance creoles.

4.2 *Partial restructuring, decreolization and standardization*

Holm (2004) takes on a similar issue, but from a comparative creole perspective, by comparing African-American English, Afrikaans, Brazilian Vernacular Portuguese, Non-standard Caribbean Spanish and Vernacular Lects of Réunionnais French (mostly called Réunionnais creole or simply *créole*; cf. Chaudenson 2001:30). The creole status of all the above varieties has been disputed in the past.³³ Holm (2004: preface) acknowledges the controversy surrounding the terms: while 'semi-creole' was first introduced by Reinecke (1938) for Afrikaans, the big debate surrounding African-American English and decreolization occupied an important place in the American beginnings of the discipline of Pidgin and Creole Studies, as discussed earlier (cf. also Holm 2004:1f.; Mufwene 2001:86–98). According to this view African-American English developed from a formerly widespread English creole spoken across the southern parts of the US, which subsequently underwent decreolization, which is understood as a loss of creole features and hence an approximation towards the lexifier. However, the term was used by Reinecke in a different

³³ Compare for example Naro and Scherre (1999) for Brazilian Vernacular Portuguese, Lipski (2005) for Afro-Hispanic including non-standard Caribbean varieties, and Baker (2000) for Réunionnais.

sense, as Holm (2004:7) indicates, namely for the result of an incomplete creolization process. Holm adopts the term ‘partial restructuring’ in this sense and in line with the discussion on degrees of restructuring (p. 10), but the use of the term ‘partial’ seems to indicate that restructuring is a process with a clearly defined end-product, while the discussion above of Neumann-Holzschuh and Schneider (2000) and Posner (1985) highlighted the possibility of viewing the restructuring processes of creole languages in a wider perspective of language change (cf. also Mufwene 1997:58f.). I shall return to this question after discussing some of the data Holm presents on Brazilian Vernacular Portuguese and relating them to the paradigmatic changes that Brazilian Portuguese is said to have undergone according to Roberts and Kato (1996) and Kato (2000).

I now return to relative clauses (all data are taken from Holm, 2004:125–27, unless otherwise indicated). Brazilian Vernacular Portuguese exhibits the relative complementizer discussed above for Romance and was first systematically addressed by Tarallo (1983). He discusses the following three existing forms:

- (12) a. A menina **com** **quem** falei
 the girl with whom I-spoke
 b. A menina **que** eu falei **com** **ela**
 the girl that I spoke with her
 c. A menina **que** eu falei^{34,35}
 the girl that I spoke
 ‘The girl I talked to’

While the first variant exists almost exclusively in writing, the variant with a resumptive pronoun is highly stigmatized. The most widely used option in educated speech is the variant with the general complementizer and no resumptive pronoun (cf. Bagno 2005:185–98). Holm (2004:125f.) cites examples from Brazilian Vernacular Portuguese for the variant with a resumptive pronoun, and data from Baxter (1987), who notes that relative clauses with an invariant relative complementizer and a resumptive pronoun are common

34 Data cited in Kato (1996:223). Note the subject pronouns in (12b) and (12c); as Kato (2000) discusses, Brazilian Portuguese has lost its positive null-subject parameter setting similar to French, and thus shows overt subject pronouns as do creole languages and other forms of vernacular Spanish, particularly those of the Caribbean (cf. also the discussion in Holm, 2004:128).

35 This version is generally referred to as a relative clause with a deleted resumptive pronoun.

in Portuguese and Spanish creoles;³⁶ in fact they seem to be the general strategy for relativization of PPs:

- (13) a. ome **ku** zõ sa ka fla **n-** e (São Tomense)
 man who John PRG talk about him
 'the man that John is talking about'
- b. OmE **ki** m ba **kw** E (Angolar (São Tomé))
 man that I went with him
 'the man who I went with'
- c. kel ome **ke** n fala **k'** el (Kabuverdianu)
 that man that I spoke with him
 'that man who I spoke with'
- d. N mora na kasa **ku** bu mora-ba **n el** (Kriyol (Guiné-Bissau))
 I live in house that you live=ANT in it
 'I live in the house that you used to live in'
- e. homber **ku** m' a papia **kun'** e a papia malu (Papiamentu)
 the man that I PST speak with him PST speak bad
 'the man that I spoke with spoke badly'

Holm notes that the generalized complementizer with and without a resumptive pronoun does exist in colloquial European Portuguese as well. It seems that the creole languages have simply gone furthest in generalizing the complementizer with a resumptive pronoun as the relativization strategy for PPs. On the other hand, Brazilian Vernacular Portuguese has only partly gone down that path while keeping the variant without resumptive pronoun, which is in fact the most widely diffused variant (cf. Kato 1996:224).³⁷

There is indeed a further complication: Brazilian Vernacular Portuguese differs also in the distribution of relative clauses with or without a resumptive pronoun and invariant complementizer. While both European and Brazilian Portuguese exhibit the deleted version with a PP, this does not seem to be the case for the creole languages. On the other hand, neither European Portuguese nor Portuguese creoles show resumptive pronouns for the direct object position, as Brazilian Portuguese does (cf. Holm 2004:125; Maurer 1998:181; Kihm 1994:177):

36 I cannot discuss here the debate surrounding the origins of Papiamentu and its status as a Spanish- or Portuguese-based creole (cf. Munteanu 1996; Martinus 1997). Papiamentu and Palenquero do share a number of features with some West African Portuguese creoles. However, this might not be wholly unexpected given the structural proximity of Spanish and Portuguese and, in some cases, the shared substrate.

37 She reports Tarallo's historical data. A quantitative study, particularly in isolated areas, would be interesting, in order to relate this phenomenon to similar restructuring in the Portuguese creoles.

- (14) Esse rapaz, **que** eu conheci **ele** (Amaral 1928:78)
 this guy that I know him
 'This guy, who I know'

Holm cites this example from Amaral's dialect study on the *Dialecto caipira* spoken in the hinterlands of the state São Paulo. He marks the example with a comma, presumably indicating a pause. Kato (1996:230) cites a similar example without a comma:

- (15) O buraco que taparam ele outro dia
 the hole that covered._{3PL} it other day
 'The hole that they covered the other day'

Interestingly, she relates this structure to the left-dislocation structure in the following example:

- (16) Esse buraco_i , taparam ele_i outro dia
 this hole they-covered it other day
 '(as for) this hole, they covered it the other day'

Referring to Pontes (1987), she notes that this structure, which represents a topic/comment organization, is very frequent in spoken Brazilian Portuguese. Tarallo (1983; 1989) and Kato (1996; 2000) relate these phenomena to wider changes that Brazilian Portuguese has undergone, namely the loss of the null-subject parameter, which has resulted in overt subjects (Duarte 1996) and the loss of clitics, which are substituted by full pronouns or by PPs in the case of indirect objects (Cyrino 1996). Moreover, this change has led to a general weakening of the object position, so that the object can often be deleted. While direct object deletion does occur in European Portuguese as well (Raposo 1986), it is much more widespread in Brazilian Portuguese and is part of a set of compensatory strategies to make up for the loss of the clitic (Duarte 1989; Bachmann 2011; see also Ledgeway, volume 1, chapter 8, §3.4).³⁸ Compare the following examples:

- (17) a. Achei o livro que você me indicou, mas ainda não o li
 I-found the book that you me=showed but still not it= I-read
 b. Achei o livro que você me indicou, mas ainda não li ele
 I-found the book that you me=showed but still not I-read it
 c. Achei o livro que você me indicou, mas ainda não li Ø
 I-found the book that you me=showed but still not I-read

³⁸ Schwenter and Silva (2003) have shown that animacy plays a role in the distribution of full pronouns. Note that the resumptive pronoun in the example above also refers to a person.

- d. Achei o livro que você me indicou, mas ainda não li esse
 I-found the book that you me=showed but still not I-read this
 livro
 book
 'I found the book you recommended to me, but I haven't read
 it/that book yet'

Duarte (1989) shows that the use of these variables is highly dependent on school education as well as monitoring in speech. Clitic use and that of compensatory lexical NPs rise dramatically for educated speakers when they know that the use of clitics is being tested. Bachmann (2011) develops this idea further and shows that in quality evening news programmes on television, news presenters and reporters avoid the use of clitics, considered pedantic in spoken language, but make wide use of lexical NPs and passive constructions, which eliminate direct objects, to construct direct object anaphoric reference.

This discussion of the interaction of linguistic norms with variation in language brings us back to the issue of decreolization. While the discussion above showed the validity of the idea of degrees of restructuring for creole languages, particularly when we take historical and current popular forms of the base languages as a point of departure, the discussion about Brazilian Vernacular Portuguese also reveals a lacuna of current research on creole languages and partly restructured varieties, to stay with Holm's terminology for the moment. While Holm (2004:125) refers to colloquial European Portuguese as opposed to standard Portuguese, the term 'Brazilian Vernacular Portuguese' remains vague. On the one hand, some examples are taken from isolated areas (Helevécia, Ceará), while he acknowledges that the general pattern of resumptive pronouns and deletion does exist more widely, even though we do not know the exact quantitative patterns. Interestingly, Holm (p. 126) compares the deletion strategy to decreolization: 'Such deletion of a stigmatized construction resulting in a structure that is still nonstandard [viz. constructions with prepositional phrases; I.B.] is characteristic of decreolization.'

If we accept that creolization, just like creole languages, is hard to pin down by structural features alone, it might also be better to discuss issues of decreolization as a more general phenomenon of influence from a coexisting standard variety. The continuum model tried to address this issue by relating existing sociolinguistic variation to decreolization by showing the possibility of progressive assimilation to the superstrate (DeCamp 1971b; Rickford 1987). DeCamp's and Rickford's research showed a continuum in Jamaican creole between an acrolect (Jamaican Standard English) and a basilect (the deep

creole variety), on which idiolects patterned systematically around clusters of features. On the other hand, Ferguson (1959) cites Haiti as a classical case of diglossia, because the functions of French and Haitian creole are clearly distinct and therefore prevent the type of assimilation process (or decreolization) typical for a continuum situation. Fleischman (1986) notes that part of the reason for the stability of the Haitian diglossia is rather related to the fact that many speakers in Haiti are monolingual in Haitian creole and that there is therefore no possibility of assimilation due to a lack of exposure to French (cf. also the discussion of diglossia in creoles in comparison to Romance in Schlieben-Lange 1977:88–90). On the whole, however, sociolinguistic studies have remained the exception in Pidgin and Creole Studies, despite De Rooij's (1995:53) comment that 'there is no escape from addressing the problem of variation'. This lacuna might be related to the disciplinary nature of the field. As Bickerton (1976:180) points out: 'In particular, they [sociolinguistic studies] have shed extensive light on the hitherto puzzling process of "decreolization" – that by which a creole in contact with its superstrate may progressively lose creole characteristics and eventually come to appear as no more than a rather deviant dialect of the superstrate. In turn, an understanding of decreolization has helped to change radically the prevailing opinion about the origins of Black English [. . .].'

The sociolinguistic research related to decreolization served the aim of rehabilitating African-American English as a language related to a creole rather than being a 'deviant dialect' of English. Similarly, part of the aim of Pidgin and Creole Studies was to rehabilitate these varieties as fully-fledged languages independent of their base languages. However, the discussion of a possible continuum of processes of decreolization seemed to call into question that independence by putting the creole language back into a space of variables blurring the line between the lexifier and the creole.

However, if we think in terms of degrees of restructuring resulting from contact-induced change and accept that the language/dialect boundaries are harder to determine than we might wish, we can think of decreolization as contact-induced change through coexistence with the superstrate language. This process does not seem different in quality, but rather in degree, from the situation of, for example, French in France or Portuguese in Brazil, where popular varieties, restructured to varying degrees, are in contact with a written standard, which penetrates language practices to different degrees.

We can also see an example of this in the use of relativization variants in Papiamentu, which has undergone a process of standardization. Maurer (1998:181) points out that besides the general complementizer *ku*, which also

serves as a preposition and a conjunction for coordination, other more hispanized variants are in use such as *kende* ‘who.SG’ (cf. Sp. *quien*), *kendenan* ‘who.PL’ (cf. Sp. *quienes*), *lokual* ‘which’ (cf. Sp. *el cual*), *loke* ‘that which’ (cf. Sp. *lo que*), which can be combined with prepositions. In that way relativization with a resumptive pronoun can be avoided, as the following examples from Maurer (1998) show:

- (18) a. Nan no konosé e hende *ku* m’ a papia kuné
 they not know the person that I PRV speak with.PRON.3SG
 b. Nan no konosé e hende *ku ken* m’ a papia
 they not know the person with whom I PFV speak
 ‘They don’t know the person who I spoke with’

It seems likely that these constructions have entered Papiamentu through writing, since the syntactic expansion associated with creating a written *Ausbau* variety often has recourse to the lexifier Spanish, with which many literate people are familiar, and to the national language Dutch (cf. Martinus 1990; Joubert 2002). Likewise, Pountain (2006) mentions the more ‘exact’ relative pronouns *cuyo* ‘whose’ and *el cual* ‘which’ as areas of syntactic borrowing from Latin,³⁹ and a similar process in fact seems likely for the high-register Portuguese variants discussed above. The important question is in how far these variants have spread more widely. This discussion seems to indicate that, similarly to creolization, which was discussed in the context of more general phenomenon of contact-induced restructuring, decreolization can be seen as a phenomenon not specific to creole languages, but related to changes brought about by contact with a standard language.

As example (18) shows, this requires some further specification. Pountain (2006) shows how syntactic borrowing needs to be related to a distinction of register in a language and concludes that the ease with which features can migrate downwards seems to depend on the relative distance between the written and the spoken code. This brings us back to the process of focusing, described by Trudgill (2004) for new dialect formation and as acts of identity by Le Page and Tabouret-Keller (1985) for Caribbean creole languages. It seems that while focusing leads to an autonomization of language forms, the processes described in decreolization can lead to a destabilizing of that very autonomy by leading to progressive merger with a prestige variety (the acrolect). It seems that we do not yet clearly understand what is involved in

39 The quotation marks are his, because relative constructions with a resumptive pronoun are obviously just as precise.

these processes of jelling versus a gradual loss of autonomy and, indeed, if the contact phenomena described above are different in nature from the contact-induced restructuring described earlier. However, it seems clear that a deeper understanding of the communicative situation, including the use of standard varieties in different media, is crucial to determine its nature.

Pfänder (2000), for instance, relates the more aspectual use of the preverbal 'imperfective' marker *ka* and the non-marked invariable form in Guyana French creole to the way the narrative is constructed in oral discourse in this specific society, although also related obviously to a universal tendency of marking tense and aspect. Interestingly, Pfänder (2000:233) suggests that writing seems to play a role in changes in the verbal system in the coastal areas of Guyana, where people have more access to French as the written language. There the particle system seems to lose some of its aspectual force and develops the more abstract anchoring of the temporal dimension, similarly to the French creole of Martinique, where access to the French language is generally more widespread. The conditions for this type of language contact seem to imply extending the scenario of language change from the face-to-face communication which was at the heart of Hermann Paul's conception of sound change, to other forms of language circulation.

5. Conclusions: creolization as a heuristic term

I have outlined the change of perspective in the research agendas for creole languages from nineteenth-century Romance philology and historical-comparative grammar, to the emergence of Pidgin and Creole Studies in the context of twentieth-century American Structuralism. In the nineteenth century, creole languages were analysed essentially as derivations of their base language, and their study emerged in the context of a wider interest in dialectal variation. Pidgin and Creoles Studies, on the contrary, focused on creoles as independent language systems and the comparison of creole languages as a group. This allowed for a conceptualization of creoles as fully-fledged languages, but at the expense of losing sight of the important function of their lexifiers, particularly in contexts where they still coexist with the creole, often serving as the written standard and prestige variety.

It has become clear that the way creole languages were conceptualized depended on their insertion in these different research paradigms which shaped the research focus to a certain extent. In particular, I have shown how theories of creole genesis have been shaped by these underlying conceptualizations. If we therefore ask the question of whether creoles are to be

considered Romance languages, the answer is necessarily related to disciplinary aspects as well. Neumann-Holzschuh and Schneider (2000:3) touch on this when they identify ‘different traditions and different lines of thinking’ for English versus Romance creoles. This might be partly related to the fact that a substantial proportion of the research on Romance creoles is still done by researchers within the discipline of Romance linguistics or the modern languages departments of the respective lexifiers, often with training in historical linguistics. On the other hand, the clearer move towards general linguistics departments in the United States certainly favoured the comparative creole perspective at the expense of the role of the lexifier. Finding researchers with a solid expertise in the relevant substrate languages who focus on the study of creole languages has long been a challenge.

My discussion in the previous section should have demonstrated that I believe that the study of creole languages can greatly contribute to research questions that are important for the development of the Romance languages, particularly the importance of contact-induced change. Since creole languages formed part of European colonial expansion, and consequently arose at the same time as other European varieties in the Americas, Africa and Asia, it seems essential to understand the similarities and differences in the development of all these Romance varieties.

As Schlieben-Lange had already noted in her (1977) article ‘L’origine des langues romanes – un cas de créolisation’, a contribution to the old and controversial debate about whether the origins of the Romance languages lie in creolization, this seems to be asking the wrong question. Rather, she concludes, we should understand creolization as a heuristic process that helps us ask the right questions about processes of change. She distinguishes the following factors that research on creole languages has taught us to consider when it comes to language change (pp. 94–99): the historical situation, the types of linguistic contact, the consequences of the contact situation (coexistence with base language (diglossic or not), complete shift, commercial language as go-between, etc.), and types of development (genesis and transmission). This still seems a reasonable conclusion, and we have seen that some of these suggestions are already being put into practice.

References and bibliographical abbreviations

- Adam, Lucien 1883. *Les Idiomes négro-aryen et maléo-aryen. Essai d'hybridologie linguistique*. Paris: Maisonneuve.
- Adams, James 1977. *The Vulgar Latin of the Letters of Claudius Terentianus*. Manchester: Manchester University Press.
1991. 'Some neglected evidence for Latin *habeo* with infinitive: the order of the constituents', *Transactions of the Philological Society* 89: 131–90.
2003. *Bilingualism and the Latin Language*. Cambridge: Cambridge University Press.
2007. *The Regional Diversification of Latin, 200 BC – AD 600*. Cambridge: Cambridge University Press.
- Adams, Marianne 1987. *Old French, Null Subjects, and Verb Second Phenomena*. Unpublished Ph.D. dissertation, UCLA.
- AGI = Archivio glottologico italiano.
- Agostiniani, Luciano 1983. 'Aspirate etrusche e gorgia toscana: valenza delle condizioni fonetiche dell'area toscana', in Agostiniani, Luciano and Giannelli, Luciano (eds), *Fonologia etrusca e fonetica toscana: il problema del sostrato*. Florence: Olschki, pp. 25–59.
- Ahlqvist, Anders (ed.) 1987. *Les Premières Grammaires des vernaculaires européens*. Saint-Denis: Presses universitaires de Vincennes.
- AIS = Jaberg and Jud 1928–40.
- ALF = Gillieron, Jules and Edmont, Edmond 1902–10. *Atlas linguistique de la France*. Paris: Honoré Champion.
- Alinei, Mario 1981. "'Dialecto": un concetto rinascimentale fiorentino', *Quaderni di Semantica* 2: 147–73.
- Alleyne, Mervyn 1971. 'Acculturation and the cultural matrix of creolization', in Hymes (1971), pp. 169–86.
1980. *Comparative Afro-American*. Ann Arbor: Karoma.
1986. 'Substratum influences – guilty until proven innocent', in Muysken and Smith (1986), pp. 301–15.
2000. 'Opposite processes of creolization', in Neumann-Holzschuh and Schneider (2000), pp. 125–33.
- Alonso, Amado 1968. *Castellano, español, idioma nacional*. Buenos Aires: Losada.
- Altieri Biagi, Maria Luisa 1965. *Galileo e la terminologia tecnico-scientifica*. Florence: Olschki.
- Alvar, Manuel 1953. *El dialecto aragonés*. Madrid: Gredos.
1978. 'Espagnol et langues indigènes: attitudes en Amazonie colombienne', *RLiR* 42: 1–43.
- Alvar Ezquerro, Manuel 1992. 'Spanisch: Lexikographie. Lexicografía', in *LRL* (VI.1): pp. 636–51.

- Amaral, Amadeu. [1928] 1976³. *O Dialeto Caipira*, São Paulo: HUCITEC (Secretaria da Cultura, Ciência e Tecnologia).
- Amoretti, Giovanni 1992. “L’oracolo di casa Manzoni”: Emilia Luti e la revisione dei *Promessi Sposi*, *Otto/Novecento* 16: 5–21.
- Andler, Daniel (ed.) 2004. *Introduction aux sciences cognitives*. Paris: Gallimard.
- Andresescu, Ana 2002. *Arta cărții. Cartea românească veche 1508–1700*. Bucharest: Univers enciclopedic.
- Andrieux-Reix, Nelly 2005. ‘Aspects nouveaux de la recherche en français médiéval’, in Valette, Jean-René (ed.), *Perspectives médiévales. Trente ans de recherches*. Paris: Société de langue et de littérature médiévales d’Oc et d’Oïl, pp. 9–35.
- Angelopoulos, Athanasios 1979. ‘Population distribution of Greece today according to language, national consciousness and religion’, *Balkan Studies* 20: 123–32.
- Angenendt, Arnold 1999. ‘Die Christianisierung Nordwesteuropas’, in Stiegemann, Christoph and Wemhoff, Matthias (eds), 799 – *Kunst und Kultur der Karolingerzeit: Karl der Grosse und Papst Leo III.* in Paderborn, vol. II. Mainz: von Zabern, pp. 420–33.
- Anipa, Kormi 2002. *A Critical Examination of Linguistic Variation in Golden-Age Spanish*. New York: Peter Lang.
- Anselmo, Artur 1981. *Origens da imprensa em Portugal*. Lisbon: Imprensa Nacional.
- Appel, Carl 1930 (6th edn). *Provenzalische Chrestomathie*. Leipzig: Reissland.
- Appel, René and Muysken, Peter 1987. *Language Contact and Bilingualism*. London: Arnold.
- Arends, Jacques 1995. ‘The socio-historical background of creoles’, in Arends, Muysken and Smith (1995), pp. 15–24.
2008. ‘A demographic perspective on creole formation’, in Kouwenberg and Singler (2008), pp. 309–31.
- Arends, Jacques, Muysken, Pieter and Smith, Norval (eds) 1995. *Pidgins and Creoles*. Amsterdam/Philadelphia: Benjamins.
- Ariza, Manuel 2005. ‘El romance en Al-Andalus’, in Cano Aguilar (2005), pp. 207–35.
- Arnaldi, Francesco and Smiraglia, Pasquale 2001. *Latinitatis Italiae Medii Aevi Lexicon*. Tarnunze: SISMEI - Edizioni del Galluzzo.
- Arneodo, F. 2003. ‘La place du provençal et l’influence de la culture provençale, aujourd’hui, sur son aire linguistique en Provence d’Italie’, in Blanchet, Philippe and Pons, Paul (eds), *Les Langues et cultures régionales ou minoritaires de l’Arc alpin. Actes du colloque international (Gap, 12–13 juillet 2002)*. Aix-en-Provence: Unioun Prouvençalo, pp. 23–32.
- Arquint, Jachen Curdin 1964. *Vierv ladin. Grammatica elementara dal Rumantsch d’Engiadina bassa*. Tusan: Lia rumantscha.
- Arrighi, Laurence 2007. ‘L’interrogation dans un corpus de français parlé en Acadie’, in Gadet, Françoise and Guerin, Emmanuelle (eds), *Français parlé, français hors de France, créoles (Linx 57)*, pp. 47–56.
- Arvinde, Vasile 1989. ‘Rumänisch: externe Sprachgeschichte. Histoire externe de la langue’, in *LRL* (III), pp. 288–305.
- Ascoli, Graziadio Isaia 1864. ‘Lingue e nazioni’, *Politecnico* 21:77–102.
1873. ‘Saggi ladini’, *AGI* 1:1–556.
1876. ‘Paul Meyer e il franco-provenzale’, *AGI* 2:385–95.
1878. ‘Schizzi franco-provenzali’, *AGI* 3:61–120.
1881. *Lettere glottologiche*. Milan.

1882. 'Lettere glottologiche: prima lettera', *Rivista di filologia e di istruzione classica* 10:1–71.
1975. *Scritti sulla questione della lingua* (ed. Grassi, C.). Turin: Einaudi.
- ASLEF = Pellegrini, Giovan Battista (ed.) 1972–86. *Atlante linguistico-etnografico friulano*. Padua/Udine: Dipartimento di Linguistica/Istituto di Filologia della Facoltà di Lingue e Letterature straniere.
- Atanasov, Petar 1984. 'Meglenoromâna', in Rusu (1984), pp. 476–550.
2002. *Meglenoromâna astăzi*. Bucharest: Editura Academiei.
- Atsma, Hartmut 1989. *La Neustrie. Les Pays du Nord de la Loire de 650 à 850*. Sigmaringen: Thorbecke.
- Atsma, Hartmut and Vezin, J. 1981–82. 'France I; France II', in Bruckner, Albert and Marichal, Robert (eds), *Chartae latinae antiquiores. Facsimile Edition of the Latin Charters prior to the Ninth Century*, XIII; XIV. Zurich: Graf.
- Auer, Peter 1999. 'From codeswitching via language mixing to fused lects: towards a dynamic typology of bilingual speech', *International Journal of Bilingualism* 3:309–32.
2005. 'Europe's sociolinguistic unity, or: a typology of European dialect/standard constellations', in Delbecq, Nicole, Auwers, Johan and Geeraerts, Dirk (eds), *Perspectives on Variation: Sociolinguistics, Historical, Comparative*. Berlin/New York: Mouton de Gruyter, pp. 8–42.
- Auer, Peter, Hinskens, Frans and Kerswill, Paul (eds) 2005. *Dialect Change. Convergence and Divergence in European Languages*. Cambridge: Cambridge University Press.
- Auerbach, Erich 1958. *Literatursprache und Publikum in der lateinischen Spätantike und im Mittelalter*. Berne: Francke.
- Auroux, Sylvain (ed.) 1992. *Histoire des idées linguistiques, vol. II: Le Développement de la grammaire occidentale*. Liège: Mardaga.
- (ed.) 2000. *Histoire des idées linguistiques, vol. III: L'Hégémonie du comparatisme*. Liège: Mardaga.
- Auroux, Sylvain, Bernard, Gilles and Bouille, Jacques 2000. 'Le développement du comparatisme indo-européen', in Auroux (2000), pp. 155–71.
- Auroux, Sylvain, Puccinelli Orlandi, Eni and Mazière, Francine (eds) 1998. *L'Hyperlangue brésilienne*. Paris: Larousse (*Langages* 130).
- Auzanneau, Michelle 1998. *La Parole vive du Poitou. Une étude sociolinguistique en milieu rural*. Paris: L'Harmattan.
- Avalle, D'Arco Silvio 1965. *Protostoria delle lingue romanze*. Turin: Giappichelli.
1983. *Latino «circa romançum» e «rustica romana lingua»*. Padua: Antenore.
2002. *La doppia verità. Fenomenologia ecdotica e lingua letteraria del medio evo romanzo*. Florence: Galluzzo.
- Avolio, Francesco 2001. 'Confini, isoglosse, varietà linguistiche, regioni storiche in aree abruzzese e napoletana', in Marcato, Gianna (ed.), *I confini del dialetto*. Padua: Unipress, pp. 107–26.
- Ayres-Bennett, Wendy 1987. *Vaugelas and the Development of the French Language*. London: Modern Humanities Research Association.
- (ed.) 1994a. *La Grammaire des Dames (Histoire Épistémologie Langage 16(2))*. Vincennes: Presses Universitaires de Vincennes.
- 1994b. 'Le rôle des femmes dans l'élaboration des idées linguistiques au XVII^e siècle en France', *Histoire Épistémologie Langage* 16:35–53.
1996. *A History of the French Language through Texts*. London: Routledge.
2004. *Sociolinguistic Variation in Seventeenth-Century France. Methodology and Case Studies*. Cambridge: Cambridge University Press.

- Bachmann, Iris 2005. *Die Sprachwerdung des Kreolischen. Eine diskursanalytische Untersuchung am Beispiel des Papiamentu*. Tübingen: Narr.
2007. 'Negertaaltje or volkstaal: the Papiamentu language at the crossroads of philology, folklore and anthropology', *Indiana* 24:87–105.
2011. 'Norm and variation on Brazilian TV evening news programmes: the case of third-person anaphoric reference', *Bulletin of Hispanic Studies* 88:1–20.
- Badia i Margarit, Antoni 1981. *La formació de la llengua catalana. Assaig d'interpretació històrica*. Barcelona: Abadia de Montserrat.
- Baggioni, Daniel 1998. 'Norme(s) du français au Québec et dictionnaires "québécois"', in Brasseur (1998a), pp. 261–77.
2000. 'La naissance de la créolistique', in Auroux (2000), pp. 253–61.
- Baglioni, Daniele 2010. *L'italiano delle cancellerie tunisine (1590–1703). Edizione e commento linguistico delle «Carte Cremona»*. Rome: Accademia nazionale dei Lincei and The Philological Society.
- Bagno, Marcos 2005. *Dramática da língua portuguesa*. São Paulo: Loyola.
- Bailey, Beryl Loftman 1965. 'Toward a new perspective in Negro English dialectology', *American Speech* 40:171–77.
- Baker, Philip 2000. 'Theories of creolization and the degree and nature of restructuring', in Neumann-Holzschuh and Schneider (2000), pp. 41–63.
- Baker, Philip and Corne, Chris 1986. 'Universals, substrata and the Indian Ocean creoles', in Muysken and Smith (1986), pp. 163–83.
- Bakker, Peter 1997. *A Language of Our Own*. New York/Oxford: Oxford University Press.
- Bal, Willy 1979. *Afro-romanica studia*. Albufeira: Poseidon.
2006. *Inventaire des particularités lexicales du français en Afrique noire*. Paris: EDICEF.
- Baldinger, Kurt 1958. 'Die hyperkorrekten Formen als Konsequenz der Scripta im Altgaskognischen', in Lausberg, Heinrich and Weinrich, Harald (eds), *Romanica. Festschrift für Gerhard Rohlf's*. Halle: Niemeyer, pp. 57–75.
1971. *La formación de los dominios lingüísticos en la Península Ibérica*. Madrid: Gredos.
- Banitt, Menachem 1963. 'Une langue fantôme: le judéo-français', *Revue de Linguistique Romane* 27:245–94.
- Banniard, Michel 1975. 'Le lecteur en Espagne wisigothique d'après Isidore de Séville: de ses fonctions à l'état de la langue', *Revue des Etudes Augustiniennes* 21:112–44.
- 1980a. *Le Haut Moyen Age occidental*. Paris: PUF.
- 1980b. 'Géographie linguistique et linguistique diachronique: essai d'analyse analogique en latin tardif et en occitano-roman', *Via Domitia* 24:9–43.
1985. 'Vox agrestis: quelques problèmes d'élocution de Cassiodore à Alcuin', in 'Etudes Antiques, d'Hippocrate à Alcuin', *Trames* (special issue), pp. 195–208.
1988. 'Saint Jérôme et l'elegantia d'après le De optimo genere interpretandi (ep. 57)', in Duval, Yves-Marie (ed.), *Jérôme entre l'Orient et l'Occident*. Paris: Etudes augustiniennes, pp. 305–22.
1989. *Genèse culturelle de l'Europe, Ve-VIIIe siècle*. Paris: Seuil.
- 1991a. 'Naissance et conscience de la langue d'oc (VIIIe/IXe siècles)', in Barral i Altet, Xavier (ed.), *La Catalogne et la France méridionale autour de l'an mil*. Barcelona: Generalitat de Catalunya, pp. 351–61.

- 1991b. 'Normes culturelles et réalisme langagier en Lusitanie au VI^e siècle: les choix de Martin de Braga', in *Actes du XIV Centenario del Concilio III de Toledo 589–1989*. Toledo: Arzobispado de Toledo, pp. 661–76.
- 1991c. 'Rhabanus Maurus and the vernacular languages', in Wright (1991), pp. 164–74.
- 1992a. *Viva Voce. Communication écrite et communication orale du IV^e au IX^e siècle en Occident latin*. Paris: Institut des Etudes augustinienes.
- 1992b. 'Latin et communication orale en Gaule: le témoignage de la Vita Eligii', in Fontaine, Jacques and Hillgarth, Jocelyn (eds), *L'Europe au septième siècle. Changement et continuité*. London: Studies of the Warburg Institute, pp. 58–86.
- 1992c. 'La rouille et la lime: Sidoine Apollinaire et la langue classique en Gaule au VI^e siècle', in Holtz, Louis and Fredouille, Jean-Claude (eds), *De Tertullien aux Mozarabes. Mélanges offerts à Jacques Fontaine, à l'occasion de son 70^e anniversaire, par ses élèves, amis et collègues*. Paris: Institut d'études augustinienes, pp. 413–27.
- 1993a. 'Latin tardif et français pré littéraire: observations de méthode et de chronologie', *Bulletin de la Société de Linguistique de Paris* 88:139–62.
- 1993b. 'Les deux vies de saint Riquier: du latin médiatique au latin hiératique', *Médiévales* 25:45–52.
1994. 'Seuils et frontières langagières dans la Francia romane du VIII^e siècle', in Jarnut, Jörg, Nonn, Ulrich and Richter, Michael (eds), *Karl Martel in seiner Zeit*. Sigmaringen: Thorbecke, pp. 171–90.
- 1995a. 'La cité de la parole: théorie et pratique de la communication latinophone chez saint Augustin', *Journal des Savants* 2:283–306.
- 1995b. 'Latin tardif et latin mérovingien: communication et modèles langagiers', *Revue des études latines* 73:213–30.
- 1996a. 'Oralité et formes marquées: expressivité et changement langagier', in Moussy, Claude (ed.), *Lingua latina*, 5, *L'Oralité en latin*. Paris: PUPS, pp. 69–83.
- 1996b. 'Latin tardif et langue d'oc: de quelques témoignages sociolinguistiques', in *Actes du colloque Languedoc et langue d'oc*, Toulouse. Paris: Société de langue et de littérature médiévales d'oc et d'oïl, pp. 33–46.
1997. *Du Latin aux langues romanes*. Paris: Nathan.
- 1998a. 'Diasystèmes et diachronie langagières du latin parlé tardif au protofrançais (III^e–VIII^e siècles)', in Herman (1998a), pp. 131–53.
- 1998b. 'Variations langagières et communication dans la prédication d'Augustin', in Madec, Goulven (ed.), *Augustin prédicateur (395–411)*. Paris: Institut d'Etudes augustinienes, pp. 73–93.
- 1999a. 'Latin vulgaire ou latin parlé? Question de nom, question de modèle', *Cahiers d'études romanes* 10:57–69.
- 1999b. 'Conflits culturels et compromis langagiers en Occident latin: de la crise culturelle à l'invention linguistique (III^e–IX^e siècles)', in Chrysos, Euangelos and Wood, Ian (eds), *East and West: Modes of Communication*. Leiden: Brill, pp. 223–42.
- 2000a. 'Apport de la phonologie diachronique à l'histoire des formes poétiques des IV^e/IX^e siècles', in Stella, Francesco (ed.), *Poesia dell'alto medioevo europeo. Manoscritti, lingua e musica dei ritmi latini*. Florence: SISMEL edizioni del Galluzzo, pp. 139–55.
- 2000b. 'Le latin mérovingien, état de la question', in Goullet, Monique and Parisse, Michel (eds), *Les Historiens et le latin médiéval*. Paris: Sorbonne, pp. 17–30.

- 2000c. 'Credo et langage: les missions de saint Boniface', in Dierkens, Alain and Sansterre, Jean-Marie (eds), *Voyages et voyageurs à Byzance du VI^e au XI^e siècle*. Geneva: Droz, pp. 133–56.
- 2001a. 'Normes et usages mérovingiens: graphie et parole, 650–750', *Liaisons-HESO/AIROÉ* 33:1–32.
- 2001b. 'Action et réaction de la parole latinophone: démocratisation et unification (IIIe-Ve siècle)', *Antiquité tardive* 9:115–29.
- 2001c. 'Changement de langue et changement de phase (VIIe/VIIIe s.) en Occident Latin', in Moussy, Claude (ed.), *De lingua latina nouae quaestiones*. Louvain: Peeters, pp. 1021–31.
- 2002a. 'Diasystème latinophone et interactions communicationnelles (IIIe-VIIIe s.)', *Mémoires de la Société de Linguistique de Paris (nouvelle série)* 11:47–64.
- 2002b. 'Les verbes de modalité en latin mérovingien', in Fruyt, Michèle and Moussy, Claude (eds), *Les Modalités en latin*. Paris: Presses de l'Université de Paris-Sorbonne, pp. 173–83.
- 2003a. 'Latinophones, romanophones, germanophones: interactions identitaires et construction langagière (VIIIe-Xe siècle)', *Médiévales* 45:25–42.
- 2003b. 'Changements dans le degré de cohérence graphie/langage: de la notation du phrasé à la notation de la phonie (VIIIe-XIe siècle)', *Medioevo romanzo* 27:178–99.
2004. 'Continuité et discontinuité langagières: autour de la notion d'inversion des hiérarchies (IIIe-VIIIe s.)', *Aemilianense* 1:13–31.
- 2005a. 'Questions de métalangage en linguistique diachronique', in Molinier, Christian, Choi-Jonin, Injoo, Bras, Myriam, Dagnac, Anne and Rouquier, Magali (eds), *Questions de classification en linguistique. Méthodes et descriptions*. Berne: Lang, pp. 1–17.
- 2005b. 'Prérequis de réceptibilité du latin tardif en période de transition', in Kiss, Sándor, Herman, József, Mondin, Luca and Salvi, Giampaolo (eds), *Latin et langues romanes. Études de linguistique offertes à József Herman à l'occasion de son 80ème anniversaire*. Tübingen: Niemeyer, pp. 105–13.
- 2005c. 'Niveaux de langue et communication latinophone', in *Settimana di Studio del Centro Italiano di Studi sull'Alto Medioevo LII. Comunicare e significare nell'alto medioevo*. Spoleto: CISAM, pp. 155–208.
- 2005d. 'Prototypes latins de migration à gauche des morphèmes suffixés', *L'Information grammaticale* 107:3–7.
- 2006a. 'La construction du passé langagier: invention du clivage de la parole. Normes bourgeoises et pensée linguistique', in Claudes, Pierre, Bernard-Griffiths, Simone and Vibert, Bertrand (eds), *La Fabrique du Moyen Age. Représentations du Moyen-Age dans la culture et la littérature françaises du XIXe siècle*. Paris: Champion, pp. 191–203.
- 2006b. 'Langue des Vies, langue des chartes aux VIe-VIIIe siècles: questions sur la réceptibilité de l'écriture en Occident Latin', in Bremer, Ernst and Röhl, Susanne (eds), *Language of Religion – Language of the People. Medieval Judaism, Christianity and Islam*. Munich: Fink, pp. 191–204.
- (2006c). 'Délimitation temporelle entre le latin et les langues romanes', in Ernst et al. (2003–8), pp. 544–554.

- 2008a. 'La latinophonie: caractères d'une dynamique vitale', in Greule, Albrecht, Herrmann, Hans-Walter, Ridder, Klaus and Schorr, Andreas (eds), *Studien zu Literatur, Sprache und Geschichte in Europa. Wolfgang Haubrichs zum 65. Geburtstag gewidmet*. St Ingbert: Roehrig, pp. 215–27.
- 2008b. 'Du latin des illettrés au roman des lettrés: la question des niveaux de langue en France (VIIIe-XIIe siècle)', in von Moos (2008), pp. 269–86.
- Barbieri, Edoardo 1992. *Le Bibbie italiane del Quattrocento e del Cinquecento. Storia e bibliografia ragionata delle edizioni in lingua italiana dal 1471 al 1600*. Milan: Editrice Bibliografica.
- Baretti, Giuseppe Marco Antonio 1768. *An Account of the Manners and Customs of Italy. With Observations on the Mistakes of Some Travellers, with Regard to that Country*. London: T. Davies.
- Barra Jover, Mario 2008. 'Tradición discursiva, creación y difusión de innovaciones sintácticas: la cohesión de los argumentos nominales a partir del siglo XIII', in Kabatek (2008), pp. 127–49.
- Bartoli, Matteo 1906. *Das Dalmatische. Altromanische Sprachreste von Veglia bis Ragusa und ihre Stellung in der Apennino-Balkanischen Romania*. Vienna: Hölder.
- 1945a. 'Sao ko kelle terre . . .', *LN* 6:1–6.
- 1945b. *Saggi di linguistica spaziale*. Turin: Rosenberg & Sellier.
2000. *Il dalmatico. Resti di un'antica lingua romanza parlata da Veglia a Ragusa e sua collocazione nella Romania appennino-balcanica*. Rome: Istituto della Enciclopedia Italiana fondata da Giovanni Treccani.
- Bartoli, Matteo and Vidossi, Giuseppe 1923. *Lineamenti di linguistica spaziale*. Milan: Le lingue estere.
- Battisti, Carlo 1927. 'Appunti sulla storia e sulla diffusione dell'ellenismo nell'Italia meridionale', *RLR* 3:1–91.
- Bauer, Brigitte 2001. 'Syntactic innovation in Latin poetry? The origins of the Romance adverbial formation in *-ment(e)*', in Orbán, Árpád and van der Poel, Marc (eds), *Ad Literas. Latin Studies in Honour of J. H. Brouwers*. Nijmegen: Nijmegen University Press, pp. 29–43.
2003. 'The adverbial formation in *mente* in Vulgar and Late Latin: a problem in grammaticalization', in Solin, Heikki, Leiwo, Martti and Halla-aho, Hilla (eds), *Latin tardif et latin vulgaire VI*. Hildesheim: Olms, pp. 439–57.
- Bauer, Roland, Goebel, Hans and Haimenl, Edgar (eds) 2005. *ALD. 1: Atlant linguistisch dl ladin dolomitich y di dialec vejins, 1: Sprechender Sprachatlas Karten 1–884 1 DVD*. Salzburg: Universität Salzburg.
- Bavoux, Claudine 2000. *Le Français de Madagascar*. Brussels: Duculot.
- Baxter, Alan 1987. 'Creole-like traces in rural Brazilian Portuguese dialects'. Department of Spanish, La Trobe University: unpublished ms.
- Bayer, Clemens 2007. 'Vita Eligii', in *Reallexikon der Germanischen Altertumskunde XXXV*. Berlin/New York: de Gruyter, pp. 461–524.
- Bec, Pierre 1986. *La Langue occitane*. Paris: PUF.
1991. 'Okzitanisch: Sprachnormierung und Standardsprache. Norme et standard', in *LRL* (V.2), pp. 45–58.
- Beccaria, Gian Luigi 1968. *Spagnolo e spagnoli in Italia. Riflessi iberici sulla lingua italiana del Cinque e Seicento*. Turin: Giappichelli.

- Beckmann, Gustav 1963. *Die Nachfolgekonstruktionen des instrumental Ablativs im Spätlatein und im Französischen*. Tübingen: Niemeyer.
- Bédarida, Henri and Hazard, Paul 1934. *L'Influence française en Italie au dix-huitième siècle*. Paris: Les Belles Lettres.
- Belardi, Walter 1991. *Storia sociolinguistica della lingua ladina*. Rome/Corvara Selva: Dipartimento di studi glottoantropologici, Università La Sapienza/Casse Raiffeisen della Val Badia e della Val Gardena.
2003. *Breve storia della lingua e della letteratura ladina*. San Martin de Tor: Istitut Ladin 'Micurà de Rù'.
- Bello Rivas, Maria Xesús 1998. 'O texto galego da Crónica General e o seu traductor', in Kremer, Dieter (ed.), *Homenaxe a Ramón Lorenzo*. Vigo: Galaxia, pp. 59–69.
2001. 'São Vicente de Fora (Lisboa) / Montederramo (Ourense): en torno a la tradición notarial gallego-portuguesa', in Jacob and Kabatek (2001), pp. 45–61.
- Beltrán, Vicenç 2005. *La corte de Babel. Lenguas, poética y política en la España del siglo XIII*. Madrid: Gredos.
- Bénabou, Marcel 1976. *La Résistance africaine à la romanisation*. Paris: Maspero.
- Benincà, Paola 1983–84. 'Un'ipotesi sulla sintassi delle lingue romanze medievali', *Quaderni patavini di linguistica* 4:3–19.
1988. *Piccola storia ragionata della dialettologia italiana*. Padua: Unipress.
1994. 'Linguistica e dialettologia italiana', in Lepschy (1994), pp. 525–644.
- Bentahila, Abdelali and Davies, Eirlys 1983. 'The syntax of Arabic-French code-switching', *Lingua* 59:301–30.
1998. 'Codeswitching: an unequal partnership', in Jacobson, Rodolfo (ed.), *Codeswitching Worldwide*. Berlin: Mouton de Gruyter, pp. 25–49.
- Bergé P. and Pomeau Y. (eds) 1995. *Le Chaos*. Dossier de *Pour la science* 1. Paris: Editions Belin.
- Bernstein, Judy 1991. 'DPs in French and Walloon: evidence for Parametric Variation in Nominal Head Movement', *Probus* 3:101–26.
- Berruto, Gaetano 1987. *Sociolinguistica dell'italiano contemporaneo*. Rome: La Nuova Italia Scientifica.
1995. *Fondamenti di sociolinguistica*. Rome/Bari: Laterza.
- Berschlin, Helmut 1994. 'Portugiesisch: Diglossie und Polyglossie', in *LRL* (VI.2), pp. 367–81.
2001. 'Interne und externe Sprachgeschichte. Histoire interne et externe des langues', in *LRL* (I.2), pp. 628–37.
- Beumann, Helmut 1964. 'Gregor von Tours und der sermo rusticus', in Repgen, Konrad and Skalweit, Stephan (eds), *Spiegel der Geschichte. Festgabe für Max Braubach*. Münster: Aschendorff, pp. 69–98.
- Bianchi, Patricia, De Blasi, Nicola and Librandi, Rita 1993. *I' te vurria parlà. Storia della lingua a Napoli e in Campania*. Naples: Tullio Pironti.
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan and Finegan, Edward 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Bickerton, Derek 1976. 'Pidgin and creole studies', *Annual Review of Anthropology* 5:169–93.
1981. *Roots of Language*. Ann Arbor: Karoma.
1984. 'The language bioprogram hypothesis', *Behavioral and Brain Sciences* 7:173–221.
- Bilola, Edmond 2003. *La Langue française au Cameroun*. Berne: Peter Lang.

- Bischoff, Bernhard 1990. *Latin Palaeography*. Cambridge: Cambridge University Press.
- Blake, Robert 1991. 'Syntactic aspects of Latinate texts of the Early Middle Ages', in Wright (1991), pp. 219–32.
- Blanc, Michel 1993. 'French in Canada', in Sanders, Carol (ed.), *French Today*. Cambridge: Cambridge University Press, pp. 239–56.
- Blanche-Benveniste, Claire 1997. *Approches de la langue parlée en français*. Paris: Ophrys.
- Blanchet, Philippe 1996. 'Pratiques linguistiques et sentiments d'appartenance dans le pays de Retz: résultats d'enquêtes', *Cahiers de sociolinguistique* 2:15–45.
2002. *Langues, cultures et identités régionales en Provence. La métaphore de l'aïoli*. Paris: L'Harmattan.
2004. 'Provençal as a distinct language?', *International Journal of Sociology of Language* 169:125–50.
- Blanchet, Philippe, Calvet, Louis-Jean, Hilléreau, Damien and Wilczyk, Ewen 2005. 'Le volet linguistique du recensement français de 1999. Résultats et analyse appliqués à la Provence plurilingue et au provençal', *Marges linguistiques* 10:65–87.
- Blas Arroyo, José 1991. 'Problemas teóricos en el estudio de la interferencia lingüística', *Revista española de lingüística* 21:265–89.
- Blasco Ferrer, Eduardo 1983. *Romania germanica. Spirito germanico in bocca latina?* Cagliari: Università di Cagliari.
1984. *Storia linguistica della Sardegna*. Tübingen: Niemeyer.
1985. *Grammatica del sardo logudorese e campidanese, con introduzione teorica sul problema della codificazione*. Cagliari: Della Torre.
1988. 'Sardisch: externe Sprachgeschichte. Storia del Sardo', in *LRL* (IV), pp. 884–97.
1993. 'Les plus anciens monuments de la langue sarde. Histoire, genre, description typologique et linguistique', in Selig, Frank and Hartmann (1993), pp. 109–49.
- 1995a. 'Il sardo', in *LRL* (IV), pp. 239–71.
- 1995b. 'Katalanische Koine', in *LRL* (IV), pp. 473–86.
- 1995c. 'Les scriptae catalanes', in *LRL* (IV), pp. 486–512.
- Bloomfield, Leonard [1933] 1956. *Language*. New York: Henry Holt.
- BNSRM 2004 = Biroul național de statistică al Republicii Moldova, *Recensământul 2004*, i. *Populația după naționalitate și limbă, în care vorbește de obicei*. Available at www.statistica.md (last accessed 15 August 2007).
- Bodvarsson, Örn and Van den Berg, Hendrik 2009. 'Hispanic immigration to the United States', in Bodvarsson and Van den Berg (eds), *The Economics of Immigration. Theory and Policy*. Berlin/Heidelberg: Springer, pp. 315–41.
- Boisvert, Georges 1983–85. "'Guerra às palavras afrancesadas!' Une polémique linguistique dans la presse lisboanaise en octobre 1812', *Bulletin des études portugaises et brésiliennes* 44/45:243–71.
- Bollée, Annegret 1993. 'Le français dans un contexte trilingue: le cas des Seychelles', in De Robillard and Beniamino (1993), pp. 119–27.
- Bonfante, Larissa (ed.) 1999. *Giuliano Bonfante. The Origin of the Romance Languages*. Heidelberg: Winter.
- Bonvini, Emilio 2008. 'Línguas africanas e português falado no Brasil', in Fiorin, José Luiz and Petter, Margarida (eds), *África no Brasil. A formação da língua portuguesa*. São Paulo: Contexto, pp. 15–62.
- Boretzky, Norbert 1983. *Kreolsprachen, Substrate und Sprachwandel*. Wiesbaden: Harrassowitz.

- Börner, Wolfgang 1976. *Schriftstruktur und Lautstruktur. Studien zur altgalicischen Skripta*. Tübingen: Niemeyer.
- Borst, Arno 1957. *Der Turmbau von Babel. Geschichte der Meinungen über Ursprung und Vielfalt der Sprachen und Völker*, vol. I. Stuttgart: Hiersemann.
1958. *Der Turmbau von Babel. Geschichte der Meinungen über Ursprung und Vielfalt der Sprachen und Völker*, vol. II. Stuttgart: Hiersemann.
- Bossong, Georg 1979. *Probleme der Übersetzung wissenschaftlicher Werke aus dem Arabischen in das Altspanische zur Zeit Alfons des Weisen*. Tübingen: Niemeyer.
1987. 'Sprachmischung und Sprachbau im Judenspanischen', *Ibero-romania* 25:1–22.
- Brâncuș, Grigore 1973. 'Originea structurii numeralului românesc', *SCL* 24:507–10.
1989. *Vocabularul autohton al limbii române*. Bucharest: Editura Științifică și Enciclopedică.
- Brasseur, Patrice (ed.) 1998a. *Français d'Amérique: variation, créolisation, normalisation*. Avignon: Centre d'études canadiennes.
- 1998b. 'Avant-propos', in Brasseur (1998a), pp. 7–15.
- Brasseur, Patrice and Falkert, Anika (eds) 2005. *Français d'Amérique. Approches morphosyntaxiques*. Paris: L'Harmattan.
- Brea, Mercedes 1994. 'Galegisch: externe Sprachgeschichte. Evolución lingüística externa', in *LRL* (VI.2), pp. 80–97.
- Brown, Peter 1981. *The Cult of the Saints. Its Rise and Function in Latin Christianity*. Chicago: University of Chicago Press.
1992. *Power and Persuasion in Late Antiquity*. Madison: University of Wisconsin Press.
2001. *La Vie de Saint Augustin*. Paris: Seuil.
- Brüch, Josef 1913. *Der Einfluss der germanischen Sprachen auf das Vulgärlatein*. Heidelberg: Winter.
- Brun, Auguste 1936. 'Linguistique et peuplement', *RLiR* 12:165–251.
- Brun-Trigaud, Guylaine 1990. *Le Croissant: le concept et le mot. Contribution à l'histoire de la dialectologie française au XIXe siècle*. Lyon: Centre d'études linguistiques Jacques Goudet.
- Brunel, Clovis (ed.) 1926. *Les Plus Anciennes Chartes en langue provençale. Recueil des pièces originales antérieures au XIIIe siècle*. Paris: Auguste Picard.
- Brunhölzl, Franz 1975. *Geschichte der lateinischen Literatur des Mittelalters*, vol. I. Munich: Fink.
- Bruni, Francesco 1984. *L'italiano. Elementi di storia della lingua e della cultura*. Turin: UTET.
- (ed.) 1990. *Il secondo Ottocento*. Bologna: Il Mulino.
- (ed.) 1992. *L'italiano nelle regioni. Lingua nazionale e identità regionali*. Turin: UTET.
- Bruni, Sandra (ed.) 1997. *Alcuino. De orthographia*. Florence: Sismel.
- Brunot, Ferdinand 1891. *La Doctrine de Malherbe d'après son commentaire sur Desportes*. Paris: G. Masson.
- 1905–. *Histoire de la langue française des origines à 1900*. Paris: Colin.
- 1966–79. *Histoire de la langue française*, 13 vols. Paris: Colin.
- Bulot, Thierry 2005. 'Que parle-t-on en Pays de Caux? Emergence et/ou continuité d'une communauté sociolinguistique régionale', *Marges linguistiques* 10:88–117.
- Buridant, Claude 2000. *Grammaire nouvelle de l'ancien français*. Paris: Sedes.
- Burnett, Charles 2001. 'The Coherence of the Arabic-Latin Translation Programme in Toledo in the Twelfth Century', *Science in Context* 14:249–88. [Reprinted with corrections 2009 as *Arabic into Latin in the Middle Ages. The Translators and their Intellectual and Social Context*. Farnham: Variorum Collected Studies Series, Article VII.]
- Busse, Winfred 1991. 'Zur Problematik des Judenspanischen', *Neue Romania* 21:37–84.

- Butler, Christopher 2003. *Structure and Function. A guide to three major structural-functional theories*. Amsterdam: Benjamins.
- Byrne, Francis and Holm, John 1993. *Atlantic Meets Pacific. A Global View of Pidginization and Creolization*. Amsterdam/Philadelphia: Benjamins.
- Caira Lumetti, Rossana (ed.) 1985. *Le umili operaie. Lettere di Luigia Codemo e Caterina Percoto*. Naples: Loffredo.
- Calabrese, Andrea 1993. 'The sentential complementation of Salentino: a study of a language without infinitival clauses', in Belletti, Adriana (ed.), *Syntactic Theory and the Dialects of Italy*. Turin: Rosenberg & Sellier, pp. 28–98.
- Calboli, Gualtiero 1992. 'Bemerkungen zu einigen besonderheiten des merowingisch-karolingischen Latein', in Iliescu, Maria and Maxgut, Werner (eds), *Latin vulgaire – latin tardif III*. Tübingen: Niemeyer, pp. 41–59.
2009. 'Latin syntax and Greek', in Baldi, Philip and Cuzzolin, Pierluigi (eds), *New Perspectives on Historical Latin Syntax*, Vol. 1: *Syntax of the Sentence*. Berlin/New York: Mouton de Gruyter, pp. 65–193.
- Calveras, Josep 1925. *La reconstrucció del llenguatge literari català*. Barcelona: Biblioteca Balmes.
- Calvet, Louis Jean 1999. *La Guerre des langues et les politiques linguistiques*. Paris: Hachette.
- Cameron, Averil 1991. *Christianity and the Rhetoric of Empire. The Development of Christian Discourse*. Berkeley/London: University of California Press.
- Campbell, Lyle and Muntzel, Martha 1989. 'The structural consequences of language death', in Dorian, Nancy (ed.), *Investigating Obsolescence*. Cambridge: Cambridge University Press, pp. 181–96.
- Camproux, Charles 1953. *Histoire de la littérature occitane*. Paris: Payot.
- Canfield, D. Lincoln 1981. *Spanish Pronunciation in the Americas*. Chicago/London: University of Chicago Press.
- Cano Aguilar, Rafael 1988. *El español a través de los tiempos*. Madrid: Arco/Libros.
- (ed.) 2005. *Historia de la lengua española*. Barcelona: Ariel.
- Cano González, Ana María 1992. 'Asturiano. Evolución lingüística interna', in *LRL* (VI.1), pp. 652–80.
- Canut, Cécile and Dumestre, Gérard 1993. 'Français, bambara et langues nationales au Mali', in Robillard and Beniamino (1993), pp. 219–28.
- Capidan, Theodor 1925. *Meglenoromânii. I. Istoria și graiul lor*. Bucharest: Cultura națională.
- Caracausi, Girolamo 1986. *Lingue in contatto nell'estremo Mezzogiorno d'Italia*. Palermo: Centro di Studi Filologici e Linguistici Siciliani.
- Carbonell, Jordi 1979. 'Elements d'història social i política de la llengua catalana', *Treballs de sociolingüística catalana* 2:87–102.
- Cardona, Giorgio Raimondo 1990. 'Il concetto di "koinè" linguistica', in Sanga, Glauco (ed.), *Koinè in Italia dalle origini al Cinquecento*. Bergamo: Lubrina, pp. 25–34.
1994. 'Portugiesisch in Asien: a língua portuguesa na Ásia. a) Ausdehnung und Verbreitung des Portugiesischen. Expansione e diffusione del portoghese', in *LRL* (VI.2), pp. 591–96.
- Carles, Hélène 2011. *L'Emergence de l'occitan prétextuel. Analyse linguistique d'un corpus auvergnat (IX^e-XI^e siècles)*. Strasbourg: EliPhi.
- Carolus-Barré, Louis 1964. *Les Plus Anciennes Chartes en langue française, 1, Problèmes généraux et recueil des pièces originales conservées aux archives de l'Oise 1241–1286*. Paris: Klincksieck.
- Carrié, Jean-Michel and Rousselle, Aline 1999. *L'Empire romain en mutation. Des Sévères à Constantin*. Paris: Seuil.

- Cartago, Gabriella 1990. *Ricordi d'italiano. Osservazioni intorno alla lingua e italianismi nelle relazioni di viaggio degli inglesi in Italia*. Bassano del Grappa: Ghedina e Tassotti.
- Carvalho Buescu, Maria Leonor 1984a. *Babel ou da ruptura do signo. A gramática e os gramáticos portugueses do século XVI*. Lisbon: Imprensa Nacional – Casa da Moeda.
- 1984b. *Historiografia da língua portuguesa. Século XVI*. Lisbon: Livraria Sá da Costa Editora.
- Cary, Max and Scullard, Howard 1975. *A History of Rome*. London: MacMillan.
- Casanova, Emili 1991. 'Katalanisch: Sprache und Literatur. Lengua y literatura', in *LRL* (V.2), pp. 218–31.
- Cassano, Paul 1976. 'Theories of language borrowing tested by American Spanish phonology', *RPh* 30:331–42.
- Castellani, Arrigo 1982a. 'Quanti erano gl'italofoni nel 1861?', *SLI* 8:3–26.
- 1982b. *La prosa italiana delle origini. Testi toscani di carattere pratico*. vol. I. Bologna: Pàtron.
- Castelvecchi, Alberto 1986. 'Introduzione', in Trissino, Giovan Giorgio (ed.), *Scritti linguistici*. Rome: Salerno Editrice, pp. xi–lxvi.
- Castillo Lluch, Mónica 2008. 'El castellano frente al latín: estudio léxico de las traducciones latinas de Alfonso X', in Elvira González, Javier et al. (eds), *Reinos, lenguas y dialectos en la Edad Media ibérica. La construcción de la identidad. (Homenaje a Juan Ramón Lodares)*. Madrid/Frankfurt: Iberoamericana/Vervuert, pp. 281–312.
- Catach, Nina 1968. *L'Orthographe française à l'époque de la Renaissance*. Geneva: Droz.
2001. *Histoire de l'orthographe française*. Paris: Champion.
- Caudmont, Jean 1994. 'Portugiesisch in Südostasien: le portugais dans le sud-est de l'Asie', in *LRL* (VI.2), pp. 597–609.
- Cavigelli, Pieder 1969. *Die Germanisierung von Bonaduz in geschichtlicher und sprachlicher Schau*. Frauenfeld: Huber.
- Cerquiglini, Bernard 1989. *Éloge de la variante. Histoire critique de la philologie*. Paris: Seuil.
- (ed.) 2003. *Les Langues de France*. Paris: PUF.
- Certeau, Michel de, Julia, Dominique and Revel, Jacques 1975. *Une Politique de la langue. La Révolution française et les patois: l'enquête de Grégoire*. Paris: Gallimard.
- Cesarotti, Melchiorre 1785. *Saggio sopra la lingua italiana*. Padua: Stamperia Penada.
- Changeux, Jean-Pierre 1983. *L'Homme neuronal*. Paris: Fayard.
- Charles-Picard, Gilbert 1959. *La Civilisation de l'Afrique romaine*. Paris: Plon.
- Charpentier, Jean-Michel 1993. 'La francophonie au Vanatu: concept juridique ou réalité sociologique?', in De Robillard and Beniamino (1993), pp. 303–14.
- Charpin, François 1977. *L'Idée de phrase grammaticale et son expression en latin*. University of Paris IV: unpublished doctoral thesis.
1989. 'Les finales homonymes dans les discours latins', *Revue informatique et statistique dans les sciences humaines* 25:65–108.
1990. 'La distribution de l'accent tonique dans la phrase latine', *Revue informatique et statistique dans les sciences humaines* 26:9–46.
- Chartier, Roger 1987. *The Cultural Uses of Print in Early Modern France*. Princeton, NJ: Princeton University Press.
- Chaudenson, Robert 1992. *Des Îles, des hommes, des langues. Essai sur la créolisation linguistique et culturelle*. Paris: L'Harmattan.
2001. *Creolization of Language and Culture* (revised in collaboration with Salikoko S. Mufwene). London/New York: Routledge.

- Chervel, André 1982. *Les Grammaires françaises 1800–1914. Répertoire chronologique*. Paris: Institut national de recherche pédagogique.
- Chia, Emmanuel 1990. 'The new speech forms of a rapidly growing city: Pidgin French and Camfranglais in Yaounde'. In *Annales de la Faculté des Lettres et Sciences Humaines*, Université de Yaounde 6:102–27.
- Chiappelli, Fredi (ed.) 1985. *The Fairest Flower. The Emergence of Linguistic National Consciousness in Renaissance Europe*. Florence: Accademia della Crusca.
- Chivu, Gheorge and Gheție, Ion (eds) 1997. *Istoria Limbii Române literare. Epoca veche (1532–1780)*. Bucharest: Editura Academiei Române.
- Choi, Jinny 2000. '[Person] direct object drop: the genetic cause of a syntactic feature in Paraguayan Spanish', *Hispania* 83:531–44.
- Chomarat, Jacques 1982. *Grammaire et rhétorique chez Erasme*. Paris: Les Belles Lettres.
- CIL = Corpus inscriptionum latinarum* 1882. Berlin: Reimer.
- Cintra, L. F. Lindley 1959. *A linguagem dos Foros de Castelo Rodrigo*. Lisbon: Publicações do Centro de Estudos Filológicos.
- Ciobanu, Anatol 2002. 'Limba romană și politică lingvistică în Republica Moldova', in Ichim, Ofelia and Olariu, Florin-Teodor (eds), *Identitatea limbii și literaturii române în perspectiva globalizării*. Iași: Trinitas, pp. 41–52.
- Ciolac, Marina 1989. *Rumänisch: Soziolinguistik; Sociolinguistique*, in *LRL* (III), pp. 185–97.
- Claudin, Anatole 1900–15. *Histoire de l'imprimerie en France au XVe et au XVIe siècle*. Paris: Imprimerie nationale.
- Close, Elizabeth 1974. *The Development of Modern Rumanian. Linguistic Theory and Practice in Muntenia 1821–1838*. Oxford: Oxford University Press.
- Clyne, Michael 1992. 'Pluricentric languages', in Clyne, Michael (ed.), *Pluricentric Languages*. Berlin: Mouton de Gruyter, pp. 1–9.
- Coelho, F. Adolpho 1880–86 = 1967. 'Os dialectos românicos ou neo-latinos na África, Ásia e América', in Morais-Barbosa, Jorge (ed.), *Estudos lingüísticos crioulos*. Lisbon: Academia Internacional da Cultura Portuguesa, pp. 1–108.
- Cohen Henriquez, Percy and Hesselning, Dirk Christiaan 1940. 'Papiamentse en Negerengelse spreekwoorden', *West-Indische Gids* 22:161–72.
- Collins, Roger 1990. 'Literacy and the laity in early medieval Spain', in McKitterick (1990), pp. 109–33.
1999. *The Arab Conquest of Spain, 710–797*. Oxford: Blackwell.
- Colombat, Bernard and Lazcano, Elisabeth (eds) 1998. *Corpus représentatif des grammaires et des traditions linguistiques*, vol. 1. Saint-Denis: Presses universitaires de Vincennes. (Special issue of *Histoire, Epistémologie, Langage*, 2.)
- Colon, Germà 1978. *La llengua catalana en els seus textos*, 1. Barcelona: Curial.
- Comas, Antoni 1972. *Història de la literatura catalana. IV. Part moderna*. Barcelona: Edicions Ariel.
1978. *La Decadència*. Barcelona: Dopesa.
- Contini, Michel 1987. *Etude de géographie phonétique et de phonétique instrumentale du sarde*. Alessandria: Edizioni dell'Orso.
- Corominas, Juan 1965. 'La survivance du basque jusqu'au bas moyen âge', in Corominas, Juan, *Estudis de toponímia catalana*, I. Barcelona: Editorial Barcino, pp. 93–151.
- Corrà, Loredana and Ursini, Flavia 1998. 'I migranti romanzi in età moderna', in *LRL* (VII), pp. 559–85.
- Cortelazzo, Manlio (1980). *I dialetti e la dialettologia in Italia (fino al 1800)*. Tübingen: Narr.

- Cortelazzo, Manlio, Marcato, Carla, De Blasi, Nicola and Clivio, Gianrenzo (eds) 1980. *I dialetti e la dialettologia in Italia (fino al 1800)*. Tübingen: Narr.
- Cortelazzo, Manlio, Marcato, Carla, De Blasi, Nicola and Clivio, Gianrenzo P. (eds) 2002. *I dialetti italiani. Storia, struttura, uso*. Turin: UTET.
- Cortelazzo, Michele and Paccagnella, Ivano 1992. 'Il Veneto', in Bruni (1992), pp. 220–81.
- Coseriu, Eugenio 1964. 'Pour une sémantique diachronique structurale', *Travaux de linguistique et de littérature* 2:139–86.
1973. 'Sistema, norma y habla', in *Teoría del lenguaje y lingüística general*. Madrid: Gredos, pp. 11–113.
- 1977a. '¿Arabismos o romanismos?', in Coseriu (1977d), pp. 40–69.
- 1977b. 'De semántica y etimología rumana', in Coseriu (1977d), pp. 184–202.
- 1977c. 'Sprachliche Interferenz bei Hochgebildeten', in Kolb, Herbert and Lauffer, Hartmut (eds), *Sprachliche Interferenz, Festschrift für Werner Betz*. Tübingen: Niemeyer, pp. 77–100.
- 1977d. *Estudios de lingüística románica*. Madrid: Gredos.
1980. "'Historische Sprache" und "Dialekt"', in Göschel, Joachim, Ivic, Pavle and Kehr, Kurt (eds), *Dialekt und Dialektologie. Ergebnisse des internationalen Symposiums 'Zur Theorie des Dialekts'*. Wiesbaden: Steiner, pp. 106–15.
1983. 'Linguistic change does not exist', *Linguistica nuova ed antica* 1:51–63.
1992. *Competencia lingüística. Elementos de la teoría del hablar*. Madrid: Gredos.
2000. *L'Homme et son langage*. Louvain: Peeters.
- Coseriu, Eugenio and Meisterfeld, Reinhard 2003. *Geschichte der romanischen Sprachwissenschaft, vol. I: Von den Anfängen bis zum Humanismus*. Tübingen: Narr.
- Coulmas, Florian 1985. *Sprache und Staat. Studien zu Sprachplanung und Sprachpolitik*. Berlin/New York: Walter de Gruyter.
- Coulon, Gérard 2009. *Les Voies romaines en Gaule*. Paris: Errance.
- Coveney, Aidan, Hintze, Marie-Anne and Sanders, Carol (eds) 2004. *Variation et francophonie*. Paris: L'Harmattan.
- Cravens, Thomas 1991. 'Phonology, phonetics and orthography in Late Latin and Romance: the evidence for early intervocalic sonorization', in Wright (1991), pp. 52–68.
2005. 'Microvariability in time and space: reconstructing the past from the present', in Cravens, Thomas (ed.), *Variation and Reconstruction*. Amsterdam: Benjamins, pp. 17–36.
- Cremona, Joseph 1970. 'L'axe nord-sud de la Romania et la position du toscan', in Rosetti, Alexandru (ed.), *Actele celui de-al XII-lea Congres internațional de lingvistică și filologie romanică*. Bucharest: Editura Academiei, pp. 155–59.
2002. 'Italian-based lingua francas around the Mediterranean', in Lepschy and Tosi (2002), pp. 24–30.
2003. 'Histoire linguistique externe de l'Italien au Maghreb', in Ernst et al. (2003), pp. 961–66.
- Crevatin, Franco 1989. 'Istroromanisch. a) Interne Sprachgeschichte', in *LRL* (III), pp. 537–54.
- Croce, Benedetto 1917. *La Spagna nella vita italiana durante la Rinascenza*. Bari: Laterza.
1927. 'Letteratura dialettale riflessa, la sua origine nel Seicento e il suo ufficio storico', in Croce, *Uomini e cose della vecchia Italia*. Bari: Laterza, pp. 222–34.
- CROSTAT 2006 = Republic of Croatia – Central Bureau of Statistics (CROSTAT), *Census of Population, Households and Dwellings, 31st March 2001*. Available at www.dzs.hr/default_e.htm.

- Cugno, Federica and Massobrio, Lorenzo 2010. *Gli atlanti linguistici della România*. Alessandria: dell'Orso.
- Cyrino, Sonia Maria Lazzarini 1996². 'Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos', in Roberts and Kato (1996), pp. 163–81.
- D'Agostino, Alfonso 1994. 'L'apporto spagnolo, portoghese e catalano', in Serianni and Trifone (1993–94) (vol. III), pp. 791–824.
- D'Anglejan, Alison and Tucker, G. Richard 1973. 'Sociolinguistic correlates of speech style in Quebec', in Shuy, Roger and Fasold, Ralph (eds), *Language Attitudes. Current Trends and Prospects*. Washington: Georgetown University Press, pp. 1–27.
- D'Hervé, Gildas 2005. 'Le Gallo dans l'enseignement, l'enseignement du gallo', *Marges linguistiques* 10:262–81.
- Dahmen, Wolfgang 1989. 'Rumänisch: Arealinguistik III. Meglenorumänisch', in *LRL* (III), pp. 436–47.
2005. 'Die Aromunen heute – eine Volksgruppe in der Identitätskrise?', *Südosteuropa Mitteilungen* 02/2005:66–77.
- Dahmen, Wolfgang and Kramer, Johannes 1985–94. *Aromunischer Sprachatlas-Atlasul lingvistic aromân (ALiA)*, vol. I, Hamburg: Buske; vol. II, Veitshöchheim: Lehmann.
2001. 'Sprache und Politik: b) Sprachgrenzen', in *LRL* (I.2), pp. 500–12.
- Dalbera, Jean-Philippe 2003. 'Les îlots liguriens en France', in Cerquiglini (2003), pp. 125–36.
- Dardel, Robert de 1989. 'L'hypothèse d'une base OVS en protoroman', *Probus* 1:121–43.
- 1996a. 'Roman commun – protoroman', in *LRL* (II,1), pp. 90–100.
- 1996b. *A La Recherche du protoroman*. Tübingen: Niemeyer.
2007. 'Une mise au point et une autocritique relatives au protoroman', *RLiR* 71:329–58.
- Dardi, Andrea 1984. 'Uso e diffusione del francese', in Formigari, Lia (ed.), *Teorie e pratiche nell'Italia del Settecento*. Bologna: Il Mulino, pp. 347–72.
1992. *Dalla provincia all'Europa. L'influsso francese sull'italiano tra il 1650 e il 1715*. Florence: Le Lettere.
2008. 'Un equivoco duro a morire: il "fiorentino delle persone colte"', *LN* 69:93–97.
2009. 'Ancora sul "fiorentino delle persone colte"', *LN* 70:74.
- Dauzat, Albert 1946. *Les Patois. Evolution-Classification-Etude*. Paris: Delagrave.
- Dazzi, Anna-Alice and Gross, Manfred 1989. 'Bündnerromanisch: Grammatikographie und Lexikographie. Grammaticographie et lexicographie. b) Lexikographie. Lexicographie', in *LRL* (III), pp. 897–912.
- DBI* = *Dizionario biografico degli italiani*, 1960–. Rome: Istituto della Enciclopedia italiana.
- de Bartholomaeis, Vincenzo 1899–1901. 'Contributo alla conoscenza dei dialetti dell'Italia meridionale ne' secoli anteriori al XIII. 1, Spoglio del "Codex diplomaticus cavensis"', *AGI* 15:247–74, 327–62.
- 1902–05. 'Contributo alla conoscenza dei dialetti dell'Italia meridionale ne' secoli anteriori al XIII. 2, Spoglio del "Codex diplomaticus cajetanus"', *AGI* 16:9–27.
- De Fort, Ester 1996. *La scuola elementare dall'Unità alla caduta del Fascismo*. Bologna: Il Mulino.
- de Granda, Germán 1974. 'El repertorio lingüístico de los sefarditas de Curaçao durante los siglos XVII y XVIII y el problema del origen del papiamentu', *RPh* 28:1–16.
1988. *Sociedad, historia y lengua en el Paraguay*. Bogotá: Instituto Caro y Cuervo.
- De Mauro, Tullio 1963. *Storia linguistica dell'Italia unita*. Rome/Bari: Laterza.
- De Rooij, Vincent 1995. 'Variation', in Arends, Muysken and Smith (1995), pp. 53–64.

- Débax, Hélène 2002. 'Vers une nouvelle *scripta* juridique occitane: la langue des serments languedociens du XI^e siècle', in Banniard, Michel (ed.), *Langages et peuples de l'Europe. Cristallisation des identités romanes et germaniques (VIIe–XIIe siècles)*. Toulouse: CNRS and Université de Toulouse, pp. 66–77.
- DeCamp, David 1971a. 'Introduction: the study of pidgin and creole languages', in Hymes (1971), pp. 13–39.
- 1971b. 'Towards a generative analysis of a post-creole speech continuum', in Hymes (1971), pp. 349–70.
- 1971c. 'Implication scales and sociolinguistic linearity', *Linguistics* 73:30–43.
- Dees, Anthonij 1980. *Atlas des formes et des constructions des chartes du 13e siècle*. Tübingen: Niemeyer.
1985. 'Dialectes et scriptae à l'époque de l'ancien français', *RLiR* 49:87–117.
1987. *Atlas des formes linguistiques des textes littéraires de l'ancien français*. Tübingen: Niemeyer.
- DeGraff, Michel 2003. 'Against creole exceptionalism', *Language* 79:391–410.
- 2005a. 'Morphology and word order in "creolization" and beyond', in Cinque, Guglielmo and Kayne, Richard (eds), *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press, pp. 293–372.
- 2005b. 'Linguistics' most dangerous myth: the fallacy of creole exceptionalism', *Language in Society* 34:533–91.
- Deguillaume, Colette and Amrane, Eric 2002. 'Langues parlées en Aquitaine: la pratique héritée', in *Le Quatre Pages* 110. Bordeaux: INSEE-Aquitaine. Available at www.insee.fr/fr/themes/document.asp?ref_id=5515.
- Deletant, Dennis 1996. 'Language policy and linguistic trends in the Republic of Moldavia, 1924–1992', in Dyer, Donald (ed.), *Studies in Moldovan. The history, culture, language and contemporary politics of the people of Moldova*. New York: Columbia University Press, pp. 53–87.
- Della Valle, Valeria 1993. 'La lessicografia', in Serianni and Trifone (1993–94), 1, pp. 29–91.
- Delton, Paola 1999. 'Partendo da boumbaro: un'indagine sulla situazione linguistica di Dignano d'Istria', in Marcato, Gianna (ed.), *Dialecti oggi*. Padua: Unipress, pp. 379–87.
2000. 'Gli italiani di Croazia e Slovenia: una minoranza divisa tra realtà europea ed extraeuropea', in Marcato, Gianna (ed.), *Isole linguistiche?*. Padua: Unipress, pp. 119–23.
2001. 'Abitudini linguistiche dei giovani italiani di Dignano d'Istria', in Marcato, Gianna (ed.), *I confini del dialetto*. Padua: Unipress, pp. 335–40.
- Demirtaş-Coşkun, Birgül 2001. *The Vlachs. A Forgotten Minority in the Balkans*. London: Portland.
- Densuşianu, Ovide 1938. *Histoire de la langue roumaine. Tome II*. Paris: Librairie Ernest Leroux.
- de Oliveira Marques, António Henrique 1972. *History of Portugal. I. From Lusitania to Empire*. New York/London: Columbia University Press.
1999. 'As realidades culturais. 1. As bases da cultura. 2. Meios e agentes da cultura. 3. As instituições de cultura', in Alves Dias, João José (ed.), *Nova história de Portugal. V. Portugal do Renascimento à crise dinástica*. Lisbon: Editorial Presença, pp. 447–83.
- Derrer, Felix 1974. *Lo Codi. Eine Summa Codicis in provenzalischer Sprache aus dem XII. Jahrhundert. Die provenzalische Fassung der Handschrift A (Sorbonne 632). Vorarbeiten zu einer kritischen Textausgabe*. Zurich: Juris.

- Désirat, Claude, Hordé, Tristan and Auroux, Sylvain (eds) 1982. *Les Idéologues et les sciences du langage*. Lille: Presses Universitaires de Lille.
- Dessi Schmid, Sarah, Hafner, Jochen and Heinemann, Sabine (eds) 2011. *Koinisierung und Standardisierung in der Romania*. Heidelberg: Winter.
- Detges, Ulrich 2000. 'Two types of restructuring in French creoles: a cognitive approach to the genesis of tense markers', in Neumann-Holzschuh and Schneider (2000), pp. 135–62.
- Dettori, Antonietta 1988. 'Sardisch: Grammatikographie und Lexicographie. Grammaticografia e lessicografia', in *LRL* (IV), pp. 913–15.
- Devoto, Giacomo 1953. *Profilo di storia linguistica italiana*. Florence: La Nuova Italia.
- Di Giovine, Paolo 2003. 'Sostrati, adstrati e superstrati e i loro effetti sulle lingue romanze: Italo-romania e Alpi orientali', in Ernst et al. (2003–8), pp. 578–93.
- Diallo, Alpha Mamadou 1993. 'Le français en Guinée: une situation en plein changement', in De Robillard and Beniamino (1993), pp. 229–42.
- Díaz y Díaz, Manuel 1978. *Las primeras glosas hispánicas*. Barcelona: Universidad Autónoma.
1992. 'El latín de España en el siglo VII: lengua y escritura según los textos documentales', in Fontaine, Jacques and Hillgarth, Jocelyn (eds), *L'Europe au septième siècle. Changement et continuité*. London: Studies of the Warburg Institute, pp. 25–40.
- 1996., 'Las glosas protohispánicas', in *Actas del III Congreso Internacional de Historia de la lengua española*. Salamanca: Arco Libros, pp. 653–66.
1998. 'La transición del latín al romance en perspectiva hispana', in Herman (1998a), pp. 155–72.
- Diez, Friedrich 1836. *Grammatik der Romanischen Sprachen* (I). Bonn: Weber.
- Dima, Nicholas 2001. *Moldova and the Transdnestr Republic*. Boulder: Columbia University Press.
- Dionisotti, Carlo 1970. 'Per una breve storia delle dottrine linguistiche del Rinascimento', *Filosofia* 21:15–24.
- Dolbeau, François 1996. *Vingt-six sermons au peuple d'Afrique*. Paris: Institut d'Etudes augustiniennes.
- Donner, Herbert and Rolig, Wolfgang 1966 (2nd edn) / 1968 (2nd edn). *Kanaanäische und aramäische Inschriften. I / II*. Wiesbaden: Harrassowitz.
- Donzé, Roland André 1967. *La Grammaire générale et raisonnée de Port-Royal*. Berne: Francke.
- Doria, Mario 1989. 'Dalmatico: Storia linguistica interna', in *LRL* (III), pp. 522–30.
- Dossena, Marina and Lass, Roger (eds) 2004. *Methods and Data in Historical Dialectology*. New York: Peter Lang.
- Drăganu, Nicolae 1938. Review of Pușcariu 1937, *Dacoromania* 9:287–304.
- Drescher, Martina and Neumann-Holzschuh, Ingrid (eds) 2010a. *La Syntaxe de l'oral dans les variétés non-hexagonales du français*. Tübingen: Stauffenburg.
- 2010b. 'Les variétés non-hexagonales du français et la syntaxe de l'oral: premier approche', in Drescher and Neumann-Holzschuh (2010a), pp. 9–35.
- Drusi, Riccardo 1995. *La 'lingua cortigiana romana'. Note su un aspetto della questione cinquecentesca della lingua*. Venice: il Cardo.
- Duarte, Maria Eugênia Lamoglia 1989. 'Clítico acusativo, pronome lexical e categoria vazia no português do Brasil', in Tarallo, Fernando (ed.), *Fotografias sociolinguísticas*. Campinas: Pontes, pp. 19–34.
1996. 'Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil', in Roberts and Kato (1996), pp. 107–28.

- Duarte i Montserrat, Carlos 1996. 'La llengua del manuscrit Consuetudines Dertuse', in Massip i Fonollosa, Jesús (ed.), *Costums de Tortosa*. Barcelona: Fundació Noguera, pp. xlix–lxxiii.
- Dubarry, Bernard and Dupouts, David (eds) 1995. *Pratique, présence et représentations de l'Occitan dans les Hautes-Pyrénées. Enquête sociolinguistique*. Tarbes: Conseil Général des Hautes-Pyrénées, Mission 'Langue et Culture Occitanes'.
- Dumézil, Bruno 2005. *Les Racines chrétiennes de l'Europe. Conversion et liberté dans les royaumes barbares (Ve-VIIIe s.)*. Paris: Fayard.
- Dumoulin, Olivier and Valéry, Raphael (eds) 1991. *Périodes. La construction du temps historique*. Paris: Éditions Universitaires.
- Duponchel, Laurent 1975. *Dictionnaire du français de Côte d'Ivoire*. Côte d'Ivoire: Ministère de l'éducation.
- Durante, Marcello 1981. *Dal latino all'italiano moderno. Saggio di storia linguistica e culturale*. Bologna: Il Mulino.
- Duval, Yvette 1988. *Auprès des saints corps et âmes. L'inhumation ad 'sanctos' dans la chrétienté d'Orient et d'Occident du IIIe au VIIe siècle*. Paris: Etudes augustinienes.
- Dyer, Donald 1996. 'The making of the Moldavian language', in Dyer, Donald (ed.), *Studies in Moldovan. The history, culture, language and contemporary politics of the people of Moldova*. New York: Columbia University Press. pp. 89–109.
1999. *The Romanian Dialect of Moldova. A study in language and politics*. Lewiston: The Edwin Mellen Press.
- Eberenz, Rolf 1989. 'Conciencia lingüística y pre-nacionalismo en los reinos de la España medieval', in Strosetzki, Christoph and Tietz, Manfred (eds), *Einheit und Vielfalt der Iberoromania. Geschichte und Gegenwart*. Hamburg: Buske, pp. 201–10.
1991. 'Castellano antiguo y español moderno: reflexiones sobre la periodización en la historia de la lengua', *Revista de filología española* 71:80–106.
- Eckert, Gabriele 1990. 'Französisch: Periodisierung', in *LRL* (V.1), pp. 816–29.
- Edwards, Cyril 1994. 'German vernacular literature: a survey', in McKitterick (1994), pp. 140–70.
- Efoua-Zengue, Rachel 1999. 'L'emprunt: figure néologique récurrente du camfranglais', in Mendo Zé, Gervais (ed.), *Le Français langue africaine. Enjeux et atouts pour la francophonie*. Paris: PubliSud, pp. 168–77.
- Egerland, Verner 2010. 'On Old Italian *uomo* and the classification of indefinite expressions', in D'Alessandro, Roberta, Ledgeway, Adam and Roberts, Ian (eds), *Syntactic Variation. The Dialects of Italy*. Cambridge: Cambridge University Press, pp. 71–101.
- Eisenstein, Elizabeth 1979. *The Printing Press as an Agent of Change. Communications and Cultural Transformations in Early-Modern Europe*. Cambridge: Cambridge University Press.
- Elcock, W. Dennis 1975. *The Romance Languages. Revised with a New Introduction by J. N. Green*. London: Faber.
- Elia, Silvio 1994. 'Portugiesisch: Brasilianisch. O português do Brasil', in *LRL* (VI.2), pp. 559–75.
- Elizaincín, Adolfo 1973. *Algunos aspectos de la sociolingüística del dialecto fronterizo*. Montevideo: Universidad de la República.
1976. 'Dialectología en los contactos: un ensayo metodológico', *Anuario de Letras* 26:5–29.
1981. 'The emergence of bilingual dialects on the Brazilian–Uruguayan border', in Cvitanović, Dinko and Fontanella de Weinberg, M. B. (eds), *Simposio internacional de lengua y literatura hispánicas*. Bahía Blanca: Universidad del Sur, pp. 193–201.

1988. 'The emergence of bilingual dialects on the Brazilian–Uruguayan border', *Linguistics* 177:124–34.
1992. *Dialectos en contacto. Español y portugués en España y América*. Montevideo: Arca.
- Eltis, David 2001. 'The volume and structure of the transatlantic slave trade: a reassessment', *William and Mary Quarterly* (3rd series) 58:17–46.
- Elwert, W. Theodor 1960. *Das zweisprachige Individuum*. Wiesbaden: Steiner.
- Emiliano, António 1999. 'O mais antigo documento latino-português (882 A.D.) – edição e estudo grafémico', *Verba* 26:7–42.
- 2003a. 'The textualization of Portuguese in the late 12th and early 13th centuries', *Medioevo Romano* 27:275–310.
- 2003b. *Latim e Romance na segunda metade do Século XI*. Lisbon: Fundação Gulbenkian.
- Emiliano, António and Pedro, Susana 2004. 'Da Notícia do Torto', *ZRPh* 120:1–82.
- Englebert, Annick, Pierrard, Michel, Rosier, Laurence and Van Raemdonck, Dan (eds) 1998. 'Table ronde: Atlas linguistiques et variabilité', in *Actes du XXII^e Congrès International de Linguistique et de Philologie romanes*, vol. III. Tübingen: Niemeyer, 405–33.
- Ernst, Gerhard 1970. *Die Toskanisierung des römischen Dialekts im 15. und 16. Jahrhundert*. Tübingen: Niemeyer.
- Ernst, Gerhard, Gleßgen, Martin-Dietrich, Schmitt, Christian and Schweickard, Wolfgang (eds) 2003–8. *Romanische Sprachgeschichte. Ein internationales Handbuch zur Geschichte der romanischen Sprachen / Histoire linguistique de la Romania. Manuel international d'histoire linguistique de la Romania*, 3 vols. Berlin/New York: De Gruyter.
- EULAl 2004 = Generalitat de Catalunya – Secretaria de Política Lingüística, *Enquesta d'usos lingüístics a l'Alguer 2004. Dades sintètiques*.
- EULAR 2004 = Generalitat de Catalunya – Secretaria de Política Lingüística, *Estadística d'usos lingüístics a la Franja d'Aragó 2004. Dades sintètiques*.
- EULBal 2004 = Generalitat de Catalunya – Secretaria de Política Lingüística. Govern des les Illes Balears, *Enquesta d'usos lingüístics a la Illes Balears 2004. Dades sintètiques*. Available at www6.gencat.net/llengcat/socio/docs/EULB2004.pdf.
- EULCat 2003 = *Estadística d'usos lingüístics a Catalunya 2003*. Realitzada conjuntament per la Secretaria de Política Lingüística del Departament de la Presidència i l'Idescat, Generalitat de Catalunya – Institut d'Estadística de Catalunya, 2004. Available at www6.gencat.net/llengcat/socio/docs/EULC2003.pdf; www.idescat.net/dequavi/?TC=444&Vo=15&Vi=2.
- EULCatNord 2004 = Generalitat de Catalunya – Secretaria de Política Lingüística, *Enquesta d'usos lingüístics a la Catalunya Nord 2004. Dades sintètiques*.
- Ewald, Konrad 1964. 'Formelhafte Wendungen in den Straßburger Eiden', *VR* 23:35–55.
- Ewert, Alfred 1963. *The French Language*. London: Faber.
- Falkowski, Rudolf 1971. 'Studien zur Sprache der Merowingerdiplome', *Archiv für Diplomatik* 17:1–125.
- Fanciullo, Franco 1993. 'Particolarismo siciliano e dialetti del sud-continentale: episodi storici ed episodi lessicali', in Trovato, Paolo (ed.), *Lingue e culture dell'Italia meridionale, 1200–1600*. Rome: Bonacci, pp. 345–63.
- Fauconnier, Jean-Luc 1998. 'Les données de Belgique romane', in Eloy, Jean-Michel (ed.), *Evaluer la vitalité: variétés d'oïl et autres langues. Actes du colloque 'Evaluer la vitalité des variétés régionales du domaine d'oïl'*. Amiens: Université de Picardie-Jules Verne, pp. 91–100.

- Féral, Carole de 2010. 'Pourquoi on doit seulement speak comme les white? Appropriation vernaculaire du français chez les jeunes au Cameroun', in Drescher and Neumann-Holzschuh (2010a), pp. 53–64.
- Ferguson, Charles 1959. 'Diglossia', *Word* 15:325–40.
- Fernández Rei, Francisco 1999. 'A situación do galego en Galicia e no Occidente de Asturias, de León e de Zamora', in Fernández Rei, Francisco and Santamarina Fernández, Antón (eds), *Estudios de sociolingüística románica. Linguas e variedades minorizadas*. Santiago de Compostela: Universidade de Santiago de Compostela, pp. 27–81.
2003. 'A situación do galego nas comarcas occidentais de Asturias, León e Zamora', in Monteagudo Romero, Henrique and Bouzada Fernández, Xan M. (eds), *O proceso de normalización do idioma galego (1980–2000) (III). Elaboración e difusión da lingua*. Santiago de Compostela: Consello da Cultura Galega, pp. 363–434.
2007. 'Plurilingüismo y contacto de lenguas en la Romania Europea', in Gargallo Gil, José and Reina Bastardas, María (eds), *Manual de lingüística románica*. Barcelona: Ariel, pp. 477–516.
- Fernández-Ordóñez, Inés 2011. *La lengua de Castilla y la formación del español*. Madrid: Real Academia Española.
- Ferreira, Manuel 1988. *Que futuro para a língua portuguesa em África? Uma perspectiva socio-cultural*. Linda-a-Velha: ALAC.
- Ferrer i Gironès, Francesc 1986. *La persecució política de la llengua catalana. Història de les mesures preses contra el seu ús de la Nova Planta fins avui*. Barcelona: Edicions 62.
- Ferro, Teresa 1992. *Latino e lingue balcaniche nella formazione del romeno*. Catania: CUECM.
- Feugère, Michel (eds) 2004. 'L'écriture dans la société galloromaine', *Gallia* 61:1–192.
- FEW = von Wartburg, Walter 1922–2003. *Französisches Etymologisches Wörterbuch*. Basel: Zbinden.
- Filipi, Goran 2003. 'Câteva observații asupra istroromânei actuale', *Tabula. Journal of the Faculty of Letters in Pula* 6:83–94.
- Finotti, Irene and Minerva, Nadia (eds) 2011–12. *Voix féminines. Ève et les langues dans l'Europe moderne, Actes du colloque organisé par la SIHFLES à Gargnano, les 6–8 juin 2011. Documents pour l'histoire du français langue étrangère ou seconde* vols 47–48. Paris: SIHFLES.
- Fiorelli, Piero 1950. 'Pour l'interprétation de l'ordonnance de Villers-Cotterêts', *Le Français Moderne* 18:277–88.
- Fischer, Iancu 1975. 'Originea latină a neutrului românesc', *SCL* 25:569–75.
1985. *Latina dunăreană. Introducere în istoria limbii române*. Bucharest: Editura Științifică și Enciclopedică.
- Fishman, Joshua 1965. 'Who speaks what language to whom and when?', *La Linguistique* 1:67–88.
1966. *Language Loyalty in the United States*. The Hague: Mouton.
1971. *Sociolinguistique*. Paris: Nathan.
- Fleischman, Suzanne 1982. *The Future in Thought and Language*. Cambridge: Cambridge University Press.
2000. 'Methodologies and ideologies in historical linguistics: working with older languages', in Herring, Susan, van Reenen, Pieter and Schøsler, Lene (eds), *Textual Parameters in Older Languages*. Amsterdam: Benjamins, pp. 33–58.
- Fleischman, Ulrich 1986. *Das Französisch-Kreolische in der Karibik. Zur Funktion von Sprache im sozialen und geographischen Raum*. Tübingen: Narr.

- Flobert, Pierre 1999. 'Le mythe du latin dit "vulgaire"', in Bureau, B. and Nicolas, C. (eds), *Moussylanea. Mélanges de linguistique et de littérature anciennes offerts à Claude Moussy*. Louvain: Peeters, pp. 401–10.
- Florescu, Radu 1999. *Essays on Romanian History*. Iași/Oxford/Portland: Center for Romanian Studies.
- Flydal, Leiv 1951. 'Remarques sur certains rapports entre le style et l'état de langue', *Norsk Tidsskrift for Sprogvidenskap* 16:240–57.
- Fohlen, Claude 1990. 'Perspectives historiques sur l'immigration française aux Etats-Unis', *Revue européenne de migrations internationales* 6:29–43.
- Folena, Gianfranco 1983. *L'italiano in Europa. Esperienze linguistiche del Settecento*. Turin: Einaudi.
- Fondazione, Emile Chanoux 2003. *Une Vallée d'Aoste bilingue dans une Europe plurilingue – Una Valle d'Aosta bilingue in un'Europa plurilingue*. Aosta: Tipografia Valdostana (see also www.fondchanoux.org/sondagelinguistiqueq.aspx).
- Fontaine, Jacques 1983. *Isidore de Séville et la culture classique dans l'Espagne wisigothique* (II). Paris: Etudes augustinienes.
1987. 'Une révolution littéraire dans l'Occident latin: les Confessions de saint Augustin', *Bulletin de littérature ecclésiastique* 88:173–93.
- Fontanella de Weinberg, María Beatriz 1992. *El español de América*. Madrid: Mapfre.
- Formentin, Vittorio 2007. *Poesia italiana delle origini*. Rome: Carocci.
- Fosso 1999. 'Le camfranglais: une praxéogénie complexe et iconoclaste', in Mendo Zé, Gervais (ed.), *Le Français langue africaine. Enjeux et atouts pour la francophonie*. Paris: PubliSud, pp. 178–94.
- Fouché, Pierre 1958. *Phonétique historique du français, vol. II, Les voyelles*. Paris: Klincksieck.
1961. *Phonétique historique du français, vol. III, Les consonnes*. Paris: Klincksieck.
- Fourquet, Jean 1980a. 'La "grande lacune". Comment la réduire?', in Buschinger, Danielle and Vernon, Jean-Paul (eds), *Jean Fourquet, Recueil d'études. Linguistique allemande et philologie germanique, littérature médiévale*, vol. II. Amiens: Centre d'Etudes médiévales, pp. 365–72.
- 1980b. 'Pour une reconstruction structurale des enseignements de phonétique historique', in Buschinger, Danielle and Vernon, Jean-Paul (eds), *Jean Fourquet, Recueil d'études. Linguistique allemande et philologie germanique, littérature médiévale*, vol. II. Amiens: Centre d'Etudes médiévales, pp. 382–94.
- Fragnito, Gigliola 1997. *La Bibbia al rogo. La censura ecclesiastica e i volgarizzamenti della Scrittura: 1471–1605*. Bologna: Il Mulino.
- Francescato, Giuseppe 1966. *Dialettologia friulana*. Udine: Società filologica friulana.
1989. 'Friaulisch: Soziolinguistik; Sociolinguistica', in LRL (III), pp. 601–10.
- Frank, Barbara 1994. *Die Textgestalt als Zeichen. Lateinische Handschriftentradition und die Verschriftlichung der romanischen Sprachen*. Tübingen: Narr.
- Frank, Barbara and Hartmann, Jörg (eds) 1997. *Inventaire systématique des premiers documents des langues romanes*. Tübingen: Narr.
- Frank, Barbara, Hartmann, Jörg and Selig, Maria (eds) 1993. *Le Passage à l'écrit des langues romanes*. Tübingen: Narr.
- Freire, Paulo and Donaldo, Macedo 1987. *Literacy. Reading the Word and the World*. Westport, CT: Bergin and Garvey.
- Fuchs, August 1849. *Die Romanischen Sprachen in ihrem Verhältnis zum Lateinischen*. Halle: Schmidt.
- Fuchs, Catherine (ed.) 2004. *La Linguistique cognitive*. Paris/Gap: Ophrys.

- Fusina, Jacques 2003. 'Le corse', in Cerquiglini (2003), pp. 95–106.
- Gadet, Françoise 2003. *La Variation sociale en français*. Paris/Gap: Ophrys.
- Gaeng, Paul 1977. *A Study of Nominal Inflections in Latin Inscriptions*. Chapel Hill: University of North Carolina Press.
1984. *Collapse and Reorganization of the Latin Nominal Flexion as Reflected in Epigraphic Sources*. Potomac: Scripta Humanistica.
1992. 'La morphologie nominale des inscriptions chrétiennes de l'Afrique', in Iliescu, Maria and Maxgut, Werner (eds), *Latin vulgaire – latin tardif III*. Tübingen: Niemeyer, pp. 115–31.
- Gáldi, László 1939. *Les Mots d'origine néo-grecque en roumain*. Budapest: Kir. M. Pázmány Péter Tudományegyetemi Görög Filológiai intézet.
- Gallina, Annamaria 1959. *Contributi alla storia della lessicografia italo-spagnola dei secoli XVI e XVII*. Florence: Olshcki.
- Galmés de Fuentes, Álvaro 1983. *Dialectología mozárabe*. Madrid: Gredos.
- Gamillscheg, Ernst 1934–36. *Romania Germanica*. Berlin: de Gruyter.
1936. *Zum romanischen Artikel und Possessivpronomen*. Berlin: Akademie der Wissenschaften.
1948. 'Germanische Wörter im Vulgärlatein', *RF* 61:212–24.
1957. *Historische französische Syntax*. Tübingen: Niemeyer.
- García, Constantino 1976. 'Interferencias lingüísticas entre gallego y castellano', *Revista Española de Lingüística* 6:327–43.
- García Mouton, Pilar (ed.) 1994. *Geolingüística. Trabajos europeos*. Madrid: Consejo Superior de Filología hispánica.
- García Turza, Claudio, González Bachiller, Fabián and Mangado Martínez, Javier (eds) 1999. *Actas del IV Congreso Internacional de Historia de la Lengua Española*. Logroño: Universidad de la Rioja.
- Gärtner, Kurt, Holtus, Günter, Rapp, Andrea and Völker, Harald (eds) 2001. *Skripta, Schreiblandschaften und Standardisierungstendenzen. Urkundensprachen im Grenzbereich von Germania und Romania im 13. und 14. Jahrhundert*. Trier: KlioMedia.
- Gartner, Theodor 1879. *Die Gredner Mundart*. Linz: Wimmer.
1883. *Raetoromanische Grammatik*. Heilbronn: Sammlung Romanischen Grammatiken.
1910. *Handbuch der rätoromanischen Sprache und Literatur*. Halle: Niemeyer.
1923. *Ladinische Wörter aus den Dolomitentälern*. Halle: Niemeyer.
- Geary, Patrick 1988. *Before France and Germany. The creation and transformation of the Merovingian world*. Oxford: Oxford University Press.
- Gérin, Pierre 1984. 'La création d'un troisième code comme mode d'adaptation à une situation où deux langues sont en contact, le chiac', in Zobl, Helmut (ed.), *Papers from the Seventh Annual Meeting of the Atlantic Provinces Linguistic Association*. Moncton: Université de Moncton, pp. 31–38.
- Gerov, Boris 1948–53. 'La romanisation entre le Danube et les Balkans d'Auguste à Adrien', in *Godišnik na Sofijskija Universitet*, 45:4–92; 47:17–122; 48:307–415.
1980. 'Die lateinisch-griechische Sprachgrenze auf der Balkanhalbinsel', in Bauchhenss, Gerhard, Neumann, Günter and Untermann, Jürgen (eds), *Die Sprachen im römischen Reich der Kaiserzeit*. Cologne: Rheinland, pp. 147–66.
- Gheție, Ion 1997a. 'Chestiuni generale (1532–1640)', in Chivu et al. (1997), pp. 59–85.
- 1997b. 'Chestiuni generale (1640–1780)', in Chivu et al. (1997), pp. 271–92.

- Giannelli, Luciano 1983. 'Aspirate etrusche e gorgia toscana: valenza delle condizioni fonetiche dell'area toscana', in Agostiniani, Luciano and Giannelli, Luciano (eds), *Fonologia etrusca e fonetica toscana. Il problema del sostrato*. Florence: Olschki, pp. 61–102.
- Giannelli, Luciano and Savoia, Leonardo 1978. 'L'indebolimento consonantico in Toscana. I'. *Rivista italiana di dialettologia* 2:23–58.
- 1979–80. 'L'indebolimento consonantico in Toscana. II'. *Rivista italiana di dialettologia* 4:38–101.
- Gifford, Douglas and Hodcroft, Frederick 1966. *Textos lingüísticos del medioevo español*. Oxford: Dolphin.
- Gili Gaya, Samuel 1963. *La lexicografía académica del siglo XVIII*. Oviedo: Universidad.
- Gilliéron, Jules 1918. *Généalogie des mots qui désignent l'abeille d'après l'Atlas linguistique de la France*. Paris: Champion.
- Gilliéron, Jules and Mongin, Jean 1905. *Scier dans la Gaule romane à Sud et à l'Est*. Paris: Champion.
- Gilliéron, Jules and Roques, Mario 1912. *Etudes de géographie linguistique d'après l'Atlas linguistique de la France*. Paris: Champion.
- Gimeno Menéndez, Francisco 1995. *Sociolingüística histórica (siglos X–XII)*. Madrid: Visor.
- Gingrich, Kurt 2009. "'That will make Carolina powerful and flourishing": Scots and Huguenots in Carolina in the 1680s', *South Carolina Historical Magazine* 110:6–34.
- Giovanardi, Claudio 1998. *La teoria cortigiana e il dibattito linguistico del primo Cinquecento*. Rome: Bulzoni.
- Gleick, James 1989. *La Théorie du chaos. Vers une nouvelle science*. Paris: Flammarion.
- Gleßgen, Martin-Dietrich 2007. *Linguistique romane. Domaine et méthodes en linguistique française et romane*. Paris: Colin.
- Gleßgen, Martin-Dietrich and Lebsanft, Franz (eds) 1997. *Alte und neue Philologie*. Tübingen: Niemeyer.
- Gleßgen, Martin-Dietrich and Pfister, Max 1995. 'Okzitanische Koine', in *LRL* (II.2), pp. 406–12.
- Godding, Robert 2001. *Prêtres en Gaule mérovingienne*. Brussels: Société des Bollandistes.
- Goebl, Hans 1970. *Die normandische Urkundensprache. Ein Beitrag zur Kenntnis der nordfranzösischen Urkundensprache des Mittelalters*. Vienna: Böhlau.
1975. '«Le Rey est mort, vive le Roy». Nouveaux regards sur la scriptologie', *Travaux de linguistique et de littérature* 13:145–210.
1988. 'Korsisch, Italienisch und Französisch auf Korsika', in *LRL* (IV), pp. 829–35.
1995. 'Les scriptae françaises III. Normandie', in *LRL* (II.2), pp. 314–37.
- (ed.) 1998. *Atlant linguistisch dl ladin dolomitich y di dialec vejins, 1. pert / Sprachatlas des Dolomitenladinischen und angrenzender Dialekte, 1. Teil. / Atlante linguistico del ladino dolomitico e dei dialetti limitrofi, 1. parte*. Wiesbaden: Reichert.
2000. 'La dialectométrisation de l'ALF, présentation des premiers résultats', *Linguistica* 40:209–36.
2002. 'Analyse dialectométrique des structures de profondeur de l'ALF', *RLiR* 66:5–63.
2003. 'Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): relations quantitatives et structures de profondeur', *Estudis romànics* 25:59–96.
2006. 'Sur le changement macrolinguistique survenu entre 1300 et 1900 dans le domaine d'oïl. Une étude diachronique d'inspiration dialectométrique', *Linguistica* 46:3–43.
2007. 'A bunch of dialectometric flowers: a brief introduction to Dialectometry', in Smit, Ute, Dollinger, Stefan, Hüttner, Julia, Kaltenböck, Gunter and Lutzky, Ursula (eds), *Tracing English Through Time*. Vienna: Braumüller, pp. 133–71.

2008. 'Sur le changement macrolinguistique survenu entre 1300 et 1900 dans le domaine d'Oïl. une étude diachronique d'inspiration dialectométrique', *Dialectologia* 1:3–43.
- Goldiș-Poalelungi, Ana 1973. *L'Influence du français sur le roumain (vocabulaire et syntaxe)*. Paris: Société Les Belles lettres.
- Golembeski, Daniel and Rottet, Kevin 2004. 'Régularisations de l'imparfait dans certaines variétés de français parlées aux Amériques', in Coveney et al. (2004), pp. 131–54.
- Gómez Redondo, Fernando 1998. *Historia de la prosa medieval castellana*, vol. I: *La creación del discurso prosístico. El entramado cortesano*. Madrid: Cátedra.
- Gonçalves, Maria Filomena 2003. *As ideias ortográficas em Portugal. De Madureira Feijó a Gonçalves Viana (1734–1911)*. Lisbon: Fundação Calouste Gulbenkian / Fundação para a Ciência e a Tecnologia.
- Gonçalves, Perpétua, 1996. *Português de Moçambique. Uma variedade em formação*. Maputo: Universidade Eduardo Mondlane, Faculdade de Letras, Livraria Universitária.
- González González, Manuel 1994. 'Galegisch: Soziolinguistik; Sociolinguística', in *LRL* (VI.2), pp. 46–66.
- González Ollé, Fernando 1996. *El habla toledana, modelo de la lengua española*. Toledo: Instituto Provincial de Investigaciones y Estudios Toledanos.
- Goodman, Morris 1987. 'The Portuguese element in the American creoles', in Gilbert, Glenn (ed.), *Pidgin and Creole Languages. Essays in the Memory of John E. Reinecke*. Honolulu: University of Hawaii Press, pp. 361–405.
- Goody, Jack 1977. *The Domestication of the Savage Mind*. Cambridge: Cambridge University Press.
- Gossen, Carl Theodor 1967. *Französische Skriptstudien. Untersuchungen zu den nordfranzösischen Urkundensprachen des Mittelalters*. Vienna: Hermann Böhlau.
1968. 'Graphème et Phonème: le problème central de l'étude des langues écrites du Moyen Age', *RLiR* 32:1–16.
1979. 'Méditations scriptologiques', *Cahiers de Civilisation Médiévale* 22:263–83.
1982. Review of Dees 1980, *VR* 41:273–76.
- Gouron, André 1978. *La Science juridique française aux XIe et XIIe siècles. Diffusion du Droit de Justinien et influences canoniques jusqu'à Gratien*. Milan: Giuffrè.
1985. 'Le manuscrit 632 de la Sorbonne: a la convergence des droits savants en pays d'oc', *Mélanges de la Bibliothèque de la Sorbonne* 6:6–20.
- Grafström, Ake 1958. *Étude sur la graphie des plus anciennes chartes languedociennes avec un essai d'interprétation phonétique*. Uppsala: Almqvist & Wiksell.
- Grassi, Corrado 2001. 'Die Sprachgeographie / La geografia linguistica', in *LRL* (I.1), pp. 207–35.
2001. 'Die Sprachgeographie / La geografia linguistica', in *LRL* (I.1), pp. 207–35.
- Graur, Alexandru 1929. 'A propos de l'article postposé en roumain'. *Romania* 15:475–80.
1937. 'Autour de l'article postposé', *Bulletin linguistique* 5:204–18.
1954. 'Genul neutru în românește', *Limba română* 3:30–44.
1963. *A Bird's-Eye View of the Evolution of the Rumanian Language*. Bucharest: Meridiane.
1967. 'De nouveau sur l'article postposé en roumain', *Revue roumaine de linguistique* 12:3–18.
- Graus, František 1965. *Volk, Herrscher und Heiliger im Reich der Merowinger*. Prague: Nakladatelství Československé Akademie Věd.
- Green, John 1991. 'The collapse and replacement of verbal inflection in Late Latin / early Romance: how would one know?', in Wright (1991), pp. 83–99.
2006. 'The north–south axis of Romance: contact reinforcing typology', in Lepschy and Tosi (2006), pp. 73–86.

- Greenberg, Joseph 1966. 'Some universals of grammar with particular reference to the order of meaningful elements', in Greenberg, J. (ed.), *Universals of Language*. Cambridge, MA: MIT Press, pp. 73–113.
- Grignani, Maria Antonietta 1990. 'Koinè nell'Italia settentrionale: note sui volgari scritti settentrionali', in Sanga, Glauco (ed.), *Koiné in Italia dalle origini al Cinquecento*. Bergamo: Lubrina, pp. 35–53.
- Gröber, Gustav 1884. 'Vulgarlateinische Substrate romanischer Wörter', *Archiv für lateinische Lexikographie und Grammatik* 1:204–32.
- (ed.) 1888. *Grundriss der romanischen Philologie*, vol. I. Strasbourg: Trübner.
- (ed.) 1904 (2nd edn). *Grundriss der romanischen Philologie*, vol. I. Strasbourg: Trübner.
- Gruber, Jörn 1983. *Die Dialektik des Trobar, Untersuchungen zur Struktur und Entwicklung des occitanischen und französischen Minnesangs des 12. Jahrhunderts*. Tübingen: Niemeyer.
- Grübl, Klaus 2010. 'Les multiples origines du standard: à propos du concept de *koinéisation* en linguistique diachronique', in Iliescu, Maria, Siller-Runggaldier, Heidi and Danler, Paul (eds), *Actes du XXVe Congrès International de Linguistique et de Philologie Romanes*, III. Berlin/New York: Mouton De Gruyter, pp. 573–81.
2011. 'Zum Begriff der Koine(isierung) in der historischen Sprachwissenschaft', in Dessi Schmid *et al.* (2011), pp. 37–64.
- Guey, Julien and Audin, Amable 1958. 'L'inscription récemment découverte à l'amphithéâtre des Trois-Gaules', *Comptes-rendus de l'Académie des Inscriptions et Belles-Lettres* 102:106–10.
- Guțu Romalo, Valeria 1980. 'Studiu', in Guțu Romalo, Valeria (ed.), *Gramatica Românească de Ion Heliade Rădulescu*. Bucharest: Editura Eminescu, pp. 405–536.
- Guyon, Jean 1987. *Le Cimetière aux deux lauriers. Recherches sur les catacombes romaines*. Rome: Ecole française de Rome.
- Hafner, Jochen and Oesterreicher, Wulf (eds) 2007. *Mit Clio im Gespräch. Romanische Sprachgeschichten und Sprachgeschichtsschreibung*. Tübingen: Narr.
- Haiman, John and Benincà, Paola 1992. *The Rhaeto-Romance Languages*. London/New York: Routledge.
- Hall, Robert A. Jr. 1942. *The Italian Questione della Lingua. An Interpretative Essay*. Chapel Hill: University of North Carolina.
1949. 'A note on the *gorgia toscana*', *Italica* 26:64–71.
1952. 'Bilingualism and applied linguistics', *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft* 6:13–30.
1958. 'Creolized languages and genetic relationships', *Word* 14:367–73.
1962. 'The life cycle of pidgin languages', *Lingua* 11:151–56.
1966. *Pidgin and Creole Languages*. Ithaca: Cornell University Press.
1974. *External History of the Romance Languages*. New York/London/Amsterdam: American Elsevier Publishing Company.
1976. *Proto-Romance Phonology*. New York: Elsevier.
1983. *Proto-Romance Morphology*. Amsterdam: Benjamins.
- Hannick, Christian 1999. 'Le développement des langues régionales et l'introduction d'alphabets dans les communautés d'illettrées', in Chrysos, Euangelos and Wood, Ian (eds), *East and West. Modes of Communication*. Leiden: Brill, pp. 205–21.
- Harris, Martin 1978. *The Evolution of French Syntax. A Comparative Approach*. London: Longman.
1988. 'The Romance languages', in Harris and Vincent (1988), pp. 1–25.

- Harris, Martin and Vincent, Nigel (eds) 1988. *The Romance Languages*. London/Sydney: Croom Helm.
- Harris, Tracy 1994. *Death of a Language. The History of Judeo-Spanish*. Newark: University of Delaware.
- Harris-Northall, Ray, 2005. 'Reduction of variation as a feature of the standardization of Castilian Spanish around 1500', in Cravens, Thomas (ed.), *Variation and Reconstruction*. Amsterdam: Benjamins, pp. 89–101.
- Hartmann, Jörg 1992. 'Fallstudien zum seneser Schriftgebrauch im XIII. Jahrhundert'. Tübingen: unpublished dissertation.
- Hasdeu, Bogdan Petriceicu 1883. *Din istoria limbii române*. Bucharest: Tipografia Academiei Române.
- Haskins, Charles Homer 1927. *The Renaissance of the Twelfth Century*. Cambridge, MA / London: Harvard University Press.
- Haubrichs, Wolfgang 1995. *Geschichte der deutschen Literatur von den Anfängen bis zum Beginn der Neuzeit. I. Die anfangen: Versuche volkssprachiger Schriftlichkeit im frühen Mittelalter*. Tübingen: Niemeyer.
- Haubrichs, Wolfgang and Pfister, Max 1989. 'In Francia fui'. *Studien zu den romanisch-germanischen Interferenzen und zur Grundsprache der althochdeutschen 'Pariser (Aldeutschen) Gespräche'*. Stuttgart: Akademie der Wissenschaft und der Literatur, Stuttgart.
- Haug, Walter 1997. *Vernacular Literary Theory in the Middle Ages. The German Tradition, 800–1300, in its European Context*. Cambridge: Cambridge University Press.
- Haugen, Einar 1953. *The Norwegian Language in America. A Study of Bilingual Behavior*. Philadelphia: University of Pennsylvania Press.
1966. *Language Conflict and Language Planning. The Case of Modern Norwegian*. Cambridge, MA: Harvard University Press.
- Hausmann, Frank-Rutger 1999. "'Anschrirren, Furchen, Pflügen, Säen" – Eine neue Deutung des "Indovinello Veronese"', in Nolte, Theodor and Keck, Anna (eds), *Ze hove und an der strazen. Die deutsche Literatur des Mittelalters und ihr 'Sitz im Leben'*. Stuttgart: Hirzel, pp. 153–59.
- HCSO 2004 = Hungarian Central Statistical Office, *Population Census 2001. 1.28 Population by mother tongue, nationality and sex, 1900–2001*. Available at www.nepszamlalas.hu/eng/volumes/18/tables/prntr_28.html (last accessed 13 March 2008).
- Heilmann, Luigi and Plangg, Guntram 1989. 'Ladinisch: externe Sprachgeschichte. Storia linguistica esterna', in *LRL* (III), pp. 720–33.
- Heinzelmann, Martin 1976. *Bischofsherrschaft in Gallien*. Munich: Artemis.
- Heitmann, Klaus 1989. 'Rumänisch: Moldauisch', in *LRL* (III), pp. 508–21.
- Heliade, Pompiliu 1898. *De l'influence française sur l'esprit public en Roumanie. Les origines. Étude sur l'état de la société roumaine à l'époque des règnes phanariotes*. Paris: Ernest Leroux.
- Heller, Monica 1988. 'Strategic ambiguity: codeswitching in the management of conflict', in Heller, Monica (ed.), *Codeswitching*. Berlin/New York/Amsterdam: Mouton de Gruyter, pp. 77–96.
1989. 'Communicative resources and local configurations: an explanation of language contact processes', *Multilingua* 8:357–95.
- Hennemann, Julia 2007. 'Remarques à propos du système prépositionnel de l'acadien en Nouvelle-Ecosse', *Linx* 57:79–90.
- Henry, Albert 1953. *Chrestomathie de la littérature en ancien français*. Berne: Franke.

- Héran, François, Filhon, Alexandra and Deprez, Christine 2002. 'La dynamique des langues en France au fil du XX^e siècle', *Population et sociétés* 376:1–4.
- Herman, József 1954. 'Recherches sur l'ordre des mots dans les plus anciens textes français en prose', *Acta linguistica academiae scientiarum Hungariae* 4:69–93, 351–79.
1967. *Le Latin vulgaire*. Paris: PUF.
1983. 'La langue latine dans la Gaule romaine', *Aufstieg und Niedergang der römischen Welt* 29:1045–60.
1988. 'La situation linguistique en Italie au VI^e siècle', *RLiR* 52:55–67.
1989. 'Accusativus cum infinitivo et subordonné à quod, quia en latin tardif. – Nouvelles remarques sur un vieux problème', in Calboli, Gualtiero (ed.), *Subordination and Other Topics in Latin. Proceedings of the Third Colloquium on Latin Linguistics, Bologna, April 1985*. Amsterdam: Benjamins, pp. 133–52.
1990. *Du Latin aux langues romanes. Études de linguistique historique*. Tübingen: Niemeyer.
1991. 'Spoken and written Latin in the last centuries of the Roman Empire: a contribution to the linguistic history of the Western provinces', in Wright (1991), pp. 29–43.
1992. 'Sur quelques aspects du latin mérovingien: langue écrite et langue parlée', in Iliescu, Maria and Maxgut, Werner (eds), *Latin vulgaire – latin tardif III*. Tübingen: Niemeyer, pp. 173–85.
1995. 'Les ardoises wisigothiques et le problème de la différenciation territoriale du latin', in Callebat, Louis (ed.), *Latin vulgaire–latin tardif IV*. Hildesheim: Olms-Weidmann, pp. 63–76.
- 1996a. 'Les variétés du latin', in *LRL* (II.1), pp. 44–61.
- 1996b. 'The end of the history of Latin', *RPh* 49:364–82.
- (ed.) 1998a. *La transizione dal latino alle lingue romanze. Atti della Tavola Rotonda di Linguistica Storica. Università Ca'Foscari di Venezia, 14–15 giugno 1996*. Tübingen: Niemeyer.
- 1998b. 'La chronologie de la transition: un essai', in Herman (1998a), pp. 5–26.
1999. 'Le sentiment linguistique de Grégoire de Tours', in Petersmann, Hubert and Kettemann, Rudolf (eds), *Latin vulgaire, latin tardif. V. Actes du Ve Colloque international sur le latin vulgaire et tardif*. Heidelberg: Winter, pp. 31–39.
2000. *Vulgar Latin*. Philadelphia: Pennsylvania State University Press.
2002. 'La disparition du passif synthétique latin: nouvel essai sur l'écrit et le parlé en latin mérovingien', *Estudios románicos* 24:31–46.
- Hernández, César 1992. 'Spanisch: Sprachnormierung und Standardsprache. Norma y lengua estándar', in *LRL* (VI.1), pp. 354–68.
- Hernández González, Carmen 1992. 'Spanisch: externe Sprachgeschichte. Historia lingüística externa', in *LRL* (VI.1), pp. 406–14.
- Hilty, Gerold 1968. 'Westfränkische Superstrateinflüsse auf die galloromanische Syntax', in Baldinger, Kurt (ed.), *Festschrift für W. von Wartburg I*. Tübingen: Niemeyer, pp. 493–517.
1973. 'Les origines de la langue littéraire française', *VR* 32:254–71.
1975. 'Westfränkische Superstrateinflüsse auf die galloromanische Syntax', *RF* 87:413–26.
- Hirsch, Rudolf 1967. *Printing, Selling and Reading 1450–1550*. Wiesbaden: Harrassowitz.
- Hjelmslev, Louis 1938. 'Études sur la notion de parenté linguistique: relations de parenté de langues créoles', *Revue des études indo-européennes* 2:271–86.
- Hochberg, Judith 1986. 'Functional compensation for /s/ deletion in Puerto Rican Spanish', *Language* 62:609–21.
- Hofmann, Johannes 1959. *Lateinische Syntax und Stilistik*. Munich: Beck.

- Hohl, Ernst Heinrich, Samberger, Ch. [sic] and Seyfarth, Wolfgang (eds) 1965. *Scriptores Historiae Augustae, Volume 1*. Leipzig: Teubner.
- Holm, John 1988–89. *Pidgins and Creoles*. Cambridge/New York: Cambridge University Press.
2004. *Languages in Contact: The Partial Restructuring of Vernaculars*. Cambridge/New York: Cambridge University Press.
- Holm, John and Michaelis, Susanne (eds) 2008. *Contact Languages. Critical Concepts in Language Studies*, vol. I. London/New York: Routledge.
- Holtus, Günter 1989. 'Bündnerromanisch: externe Sprachgeschichte. Histoire linguistique externe', in *LRL* (III), pp. 854–71.
1994. 'Linguística variacional do português', in *LRL* (VI.2), pp. 623–49.
- Holtus, Günter, Körner, Anja and Völker, Harald 2001. "'Endogene" und "exogene" Analogien. Hyperkorrektismen und andere Analogiebildungen in den altfranzösischen Urkunden der Grafen von Luxemburg (1237–1281)', in Henrard, Nadine, Moreno, Paola and Thiry Stassin, Martine (eds), *Convergences médiévales. Épopée, lyrique, roman. Mélanges offerts à Madeleine Tyssens*. Brussels: De Boeck, pp. 257–70.
- Holtus, Günter, Metzeltin, Michael and Schmitt, Christian (eds) 1988. *Lexikon der Romanistischen Linguistik. Band IV. Italienisch, Korsisch, Sardisch*. Tübingen: Niemeyer.
- (eds) 1989. *Lexikon der Romanistischen Linguistik. Band III. Die einzelnen romanischen Sprachen und Sprachgebiete von der Renaissance bis zur Gegenwart. Rumänisch, Dalmatisch/Istoromanisch, Friaulisch, Ladinisch, Bündnerromanisch*. Tübingen: Niemeyer.
- (eds) 1990a. *Lexikon der Romanistischen Linguistik. Band V.1. Französisch*. Tübingen: Niemeyer.
- (eds) 1990b. *Lexikon der Romanistischen Linguistik. Band V.2. Okzitanisch, Katalanisch*. Tübingen: Niemeyer.
- (eds) 1992. *Lexikon der Romanistischen Linguistik. Band VI.1. Aragonesisch-Navarresisch, Spanisch, Asturianisch-Leonesisch*. Tübingen: Niemeyer.
- (eds) 1994. *Lexikon der Romanistischen Linguistik. Band VI.2. Galegisch-Portugiesisch*. Tübingen: Niemeyer.
- (eds) 1995. *Lexikon der Romanistischen Linguistik. Band II.2. Die einzelnen romanischen Sprachen und Sprachgebiete vom Mittelalter bis zur Renaissance / Les différentes langues romanes et leurs régions d'implantation du Moyen Age à la Renaissance*. Tübingen: Niemeyer.
- (eds) 1996. *Lexikon der Romanistischen Linguistik. Band II.1. Latein und Romanisch. Historische-vergleichende Grammatik der romansichen Sprachen. Le latin et le roman. Grammaire historico-comparative des langues romanes*. Tübingen: Niemeyer.
- (eds) 1998. *Lexikon der Romanistischen Linguistik. Band VII. Kontakt, Migration und Kunstsprachen. Kontrastivität, Klassifikation und Typologie*. Tübingen: Niemeyer.
- (eds) 2001a. *Lexikon der Romanistischen Linguistik. Band I.1. Geschichte des Faches Romanistik. Methodologie (Das Sprachsystem). Histoire de la philologie romane. Méthodologie (Langue et système)*. Tübingen: Niemeyer.
- (eds) 2001b. *Lexikon der Romanistischen Linguistik. Band I.2. Methodologie (Sprache in der Gesellschaft / Sprache und Klassifikation / Datensammlung und -verarbeitung)*. Tübingen: Niemeyer.
- Hope, Thomas 1962–63. 'Loanwords as cultural and lexical symbols', *Archivum Linguisticum* 14:111–21; 15:29–42.
1964. 'The process of neologism reconsidered with reference to lexical borrowing in Romance', *Transactions of the Philological Society* 63:46–84.

1965. 'L'interprétation des mots d'emprunt et la structure lexicale', in Straka, Georges (ed.), *Actes du Xe Congrès International de Linguistique et de Philologie Romanes*, vol. 1. Paris: Klincksieck, pp. 149–55.
1971. *Lexical Borrowing in the Romance Languages. A Critical Study of Italianisms in French and Gallicisms in Italian from 1100 to 1900*. Oxford: Blackwell.
- Horrocks, Geoffrey 1997. *Greek. A History of the Language and its Speakers*. London: Longman.
- Hubschmid, Johannes 1978. 'Romania méditerranéenne et Méditerranée prélatine'. *Atti del XIV Congresso di Linguistica e Filologia Romanza I*. Naples: Macchiaroli, pp. 351–64.
- Hull, Alexander 1955. *The Franco-Canadian Dialect of Windsor, Ontario. A Preliminary Study*. University of Washington: unpublished dissertation.
- Humboldt, Wilhelm von [1827–29] 1996 (8th edn). 'Ueber die Verschiedenheiten des menschlichen Sprachbaues', in Flitner, Andreas and Giel, Klaus (eds), *Werke in fünf Bänden. III. Schriften zur Sprachphilosophie*. Stuttgart: J. G. Cotta'sche Buchhandlung, pp. 144–367.
- Hunnius, Klaus 1975. 'Zur Frage der syntaktischen Entlehnung', *RF* 87:64–81.
1988. 'Français parlé – ein problematisches Konzept', *ZRPh* 104:336–46.
2003. 'Vulgärlatein und gesprochenes Französisch: zur Entstehung des Konzepts des français avancé', *ZRPh* 119:510–19.
- Hupka, Werner 2001. 'Periodisierung von Sprache. Périodisation', in *LRL* (1.2), pp. 875–83.
- Hurren, Anthony 1969. 'Verbal aspect and archi-aspect in Istro-Romanian', *La Linguistique* 2:59–90.
- Hymes, Dell (ed.) 1971. *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press.
- Iliescu, Maria 1972. *Le Frioulan (à partir des dialectes parlés en Roumanie)*. The Hague: Mouton.
- INSSE 2002 = INSSE – Institutul Național de Statistică, *Recensământul populației și locuințelor 2002*, vol. IV: *Structura etnică și confesională*. Available at www.insse.ro/cms/files/RPL2002INS/vol4/tabele/t3.pdf (last accessed 10 March 2008).
- Ionescu, Christian 1999. *O romanés. Aspectos de historia externa con especial atención a sua situación fõra das fronteiras de Romania*, in Fernández Rei, Francisco and Santamarina Fernández, Antón (eds), *Estudios de sociolingüística románica. Linguas e variedades minorizadas*. Santiago de Compostela: Universidade de Santiago de Compostela, pp. 445–501.
- Iordan, Iorgu 1976. 'Du bilinguisme dans le domaine roman', in Boudreault, Marcel and Möhren, Frankwald (eds), *Actes du XIIIe Congrès International de Linguistique et Philologie Romanes*, vol. I. Québec: Presses de l'Université Laval, pp. 5–18.
- ISTAT 2007 = 'La lingua italiana, i dialetti e le lingue straniere. Anno 2006', in *Statistiche in breve*, Istituto Nazionale di Statistica. Available at www.istat.it/salastampa/comunicati/non_calendario/20070420_00/.
- Ivănescu, Gheorghe 1957. 'Soarta neutrului latin clasic în latina populară și în limbile romanice', *SCL* 8:299–313.
2000. *Istoria limbii române*. Iași: Junimea.
- Ive, Antonio, 1886. 'L'antico dialetto di Veglia', *AGI* 9:115–87.
- Izzo, Herbert 1972. *Tuscan and Etruscan*. Toronto: University of Toronto Press.
- Jaberg, Karl 1908. *Sprachgeographie. Beitrag zum Verständnis des 'Atlas linguistique de la France'*. Aarau: Sauerländer.
1959. *Geografía lingüística*. Granada: Universidad de Granada.

- Jaberg, Karl and Jud, Jakob 1928–40. *Sprach- und Sachatlas Italiens und der Südschweiz*. Zofingen: Ringier.
- Jackendoff, Ray 2002. *Foundations of Language. Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jacob, Daniel and Kabatek, Johannes (eds) 2001. *Lengua medieval y tradiciones discursivas en la Península Ibérica: descripción gramatical – pragmática histórica – metodología*. Frankfurt/Madrid: Vervuert.
- Jakobson, Roman 1938. 'Sur la théorie des affinités phonologiques des langues', in *Actes du IVe Congrès international de Linguistes*. Copenhagen: Munksgaard, pp. 48–58.
1963. *Essais de linguistique générale*. Paris: Editions de Minuit.
- James, Edward 2001. *Britain in the First Millennium*. London: Arnold.
- Janik, Roman 2004. 'Langue corse: réception, transmission et pratique'. *Insee-Société* 105. Available at www.insee.fr/fr/regions/corse/default.asp?page=themes/etudes/eco/corse/ecoc105/ecoc1053.htm (last accessed 1 June 2013).
- Janson, Tore 1991. 'Language change and metalinguistic change: Latin to Romance and other cases', in Wright (1991), pp. 19–28.
2008. 'The vowel systems of Late Latin', in Wright, Roger (ed.), *Latin vulgaire ~ latin tardif VIII*. Hildesheim: Olms-Weidmann, pp. 345–50.
- Jeanroy, Alfred 1934. *La Poésie lyrique des Troubadours*. Toulouse: Édouard Privat / Paris: Henri Didier.
- Jireček, Constantin 1893. 'Über die Vlach von Moglena', *Archiv für slavische Philologie* 15:98.
1901. *Die Romanen in den Städten Dalmatiens während des Mittelalters*, vol. I. Vienna: Denkschriften der Kaiserlichen Akademie der Wissenschaften in Wien (phil.-hist. Klasse).
1911. *Geschichte der Serben*, vol. I. Gotha: Friedrich-Andreas Perthes.
- Joan, Bernat, Pazos, Maria-Lluïsa and Sabater, Ernest 1994. *Història de la llengua catalana*. Vilassar de Mar (Maresme): Oikos-Tau.
- Jodl, Frank 2003. *Francia, Langobardia und Ascolis Ladinia. Die Bedeutung außersprachlicher Faktoren im Zusammenhang mit innersprachlichen Entwicklungen in drei Teilgebieten der Romania*. Frankfurt: Lang.
- Jolivet, A. and Mossé, Fernand 1941. *Manuel de l'allemand du Moyen Age des origines au 14^e siècle. Grammaire, textes, glossaire*. Paris: Aubier-Montaigne.
- Joly, André and Stéfani, Jean (eds) 1977. *La Grammaire générale des modistes aux idéologues*. Villeneuve-d'Ascq: Publications de l'Université de Lille III, Centre interdisciplinaire de recherches en linguistique.
- Jones, Mari 2001. *Jersey Norman French. A Linguistic Study of an Obsolescent Dialect*. Oxford: Philological Society.
2008. *The Guernsey Norman French Translations of Thomas Martin. A Linguistic Study of an Unpublished Archive*. Leuven: Peeters.
- Jones, Verina 1990. 'Dialect literature and popular literature', *Italian Studies* 45:103–17.
- Jordanova, Ludmilla 2000. *History in Practice*. London: Arnold.
- Joseph, Brian 1990. 'Greek', in Comrie, Bernard (ed.), *The Major Languages of Eastern Europe*. London: Routledge, pp.144–73.
- Joseph, John 1987. *Eloquence and Power. The Rise of Language Standards and Standard Languages*. London: Frances Pinter.

- Joubert, Sidney 2002. 'La estandarización de una lengua criolla: el papiamentu', in Perl, Matthias and Pörtl, Klaus (eds), *Estudios de lingüística hispanoamericana, brasileña y criolla*. Frankfurt am Main: Peter Lang, pp. 149–68.
- Jourdain, Elodie 1956. *Du Français aux parlers creoles*. Paris: Klincksieck.
- Jouveau, René 1970–84. *Histoire du Félibrige*. Aix: privately published.
- Juneau, Marcel 1987. 'La prononciation française au Québec et les parlers de France', in Niederehe, Hans-Josef and Wolf, Lothar (eds), *Français du Canada – Français de France*. Tübingen: Niemeyer, pp. 307–18.
- Jungmann, Joseph 1951–54. *Missarum solemnina. Explication génétique de la messe romaine*. Paris: Aubier.
- Kabatek, Johannes 1996. *Die Sprecher als Linguisten. Interferenz- und Sprachwandelphänomene dargestellt am Galicischen der Gegenwart*. Tübingen: Niemeyer.
- 1999a. 'Von Burgos nach Toledo, altkastilischer Normenkonflikt und Probleme der Rekonstruktion', in Wesch, Andreas and Brumme, Jenny (eds), *Normen und Subnormen in Geschichte und Gegenwart – Methoden ihrer Rekonstruktion und Beschreibung*. Vienna: Praesens, pp. 115–30.
- 1999b. 'Sobre el nacimiento del castellano desde el espíritu de la oralidad (apuntes acerca de los textos jurídicos castellanos de los siglos XII y XIII)', in Company, Concepción, González, Aurelio and von der Walde Moheno, Lilian (eds), *Discursos y representaciones en la Edad Media. Actas de las VI Jornadas Medievales*, México D.F.: UNAM – El Colegio de México, pp. 169–87.
- 2000a. 'Lo Codi und die okzitanischen Texttraditionen im 12. und 13. Jahrhundert', in Rieger, Angelica (ed.), *Okzitanistik, Altokzitanistik und Provenzalistik. Geschichte und Auftrag einer europäischen Philologie. Akten der gleichnamigen Sektion des Deutschen Romanistentages in Osnabrück 1999*. Frankfurt: Peter Lang, pp. 147–63.
- 2000b. 'L'oral et l'écrit – quelques aspects théoriques d'un «nouveau» paradigme dans le canon de la linguistique romane', in Dahmen, Wolfgang, Holtus, Günter, Kramer, Johannes, Metzeltin, Michael, Schweickard, Wolfgang and Winkelmann, Otto (eds), *Kanonbildung in der Romanistik und in den Nachbardisziplinen. Romanistisches Kolloquium XIV*. Tübingen: Narr, pp. 305–20.
2001. '¿Cómo investigar las tradiciones discursivas medievales? El ejemplo de los textos jurídicos castellanos', in Jacob and Kabatek (2001), pp. 97–132.
- 2005a. *Die bolognesische Renaissance und der Ausbau romanischer Sprachen. Juristische Diskurstraditionen und Sprachentwicklung in Südfrankreich und Spanien im 12. und 13. Jahrhundert*. Tübingen: Niemeyer.
- 2005b. 'Tradiciones discursivas y cambio lingüístico', *Lexis* 29:151–77.
- 2005c. 'Las tradiciones discursivas del español medieval: historia de textos e historia de la lengua', *Iberoromania* 62:28–43.
2007. 'Otra historia de las lenguas ibero-románicas: en torno a la actualidad de una vieja idea', in Hafner and Oesterreicher (2007), pp. 173–94.
- (ed.) 2008. *Sintaxis histórica del español y cambio lingüístico. Nuevas perspectivas desde las Tradiciones Discursivas*. Frankfurt/Madrid: Vervuert/Iberoamericana.
- Kabatek, Johannes and Schlieben-Lange, Brigitte 2000. 'Zu Notwendigkeit und theoretischem Status der Sprachkategorisierungsforschung', *Sociolinguistica* 14:115–20.
- Kahane, Henry and Kahane, Renée 1970–76. 'Abendland und Byzanz: Sprache', in Wirth, Peter (ed.), *Reallexikon der Byzantinistik*, vol. I. Amsterdam: Adolf M. Hakkert, cols. 345–640.

- Kahl, Thede 1999. *Ethnizität und räumliche Verbreitung der Aromunen in Südosteuropa*. Münster: Institut für Geographie der Westfälischen Wilhelms-Universität Münster.
2006. *Istoria aromânilor*. Bucharest: Tritonic.
- Kany, Charles 1951 (2nd edn). *American-Spanish Syntax*. Chicago: University of Chicago Press.
- Kato, Mary 1996. 'Recontando a história das relativas em uma perspectiva paramétrica', in Roberts and Kato (1996), pp. 223–61.
2000. *Brazilian Portuguese and the Null-Subject Parameter*. Madrid/Frankfurt: Iberoamericana/Vervuert.
- Kattenbusch, Dieter 1991. 'L'unificaziun dla grafia dl ladin dolomitan: l'prom vare por ne lingaz de scritöra', in *Actes du XVIIIe Congrès International de Linguistique et de Philologie Romanes, III*. Tübingen: Niemeyer, pp. 232–42.
- Kay, Sarah and Gaunt, Simon (eds) 1999. *The Troubadours. An Introduction*. Cambridge: Cambridge University Press.
- Keller, Rudi 1990. *Sprachwandel. Von der unsichtbaren Hand in der Sprache*. Tübingen: Francke.
- Kiesler, Reinhard 2006. *Einführung in die Problematik des Vulgärlateins*. Tübingen: Niemeyer.
- Kihm, Alain 1994. *Kriyol Syntax. The Portuguese-based Creole Language of Guinea-Bissau*. Amsterdam/Philadelphia: John Benjamins.
- King, Charles 2000. *The Moldovans. Romania, Russia, and the Politics of Culture*. Stanford, CA: Hoover Institution Press.
- King, Ruth 2000. *The Lexical Basis of Gramatical Borrowing*. Amsterdam/Philadelphia: Benjamins.
2008. 'Chiac in context: overview and evaluation of Acadie's Joul', in Sankoff, Gillian, Meyerhoff, Myriam and Nagy, Naomi (eds), *Social Lives in Language. Sociolinguistics and Multilingual Speech Communities*. Amsterdam: Benjamins, pp. 137–178.
- Kitts, Sally-Ann 1995. *The Debate on the Nature, Role and Influence of Woman in Eighteenth-Century Spain*. Lewiston/Queenston/Lampeter: Edwin Mellen Press.
- Kiwitt, Marc 2007. 'Des noms de contenants et d'unités de poids pharmaceutiques en judéo-français', *Cultures et civilisations médiévales* 40:45–57.
2010. 'Les glossaires hébreu-français du XIII^e siècle et la culture juive en France du nord', in Heinz, Michaela (ed.), *Cultures et lexicographie. Actes des «Troisièmes Journées allemandes des dictionnaires» en l'honneur d'Alain Rey, Klingenberg, 4–6 juillet 2008*. Berlin: Frank & Timme, pp. 113–25.
- Klausenburger, Jürgen 2000. *Grammaticalization. Studies in Latin and Romance Morphosyntax*. Amsterdam/Philadelphia: Benjamins.
- Klee, Carol and Lynch, Andrew 2002. *El español en contacto con otras lenguas*. Washington, DC: Georgetown University Press.
- Kloss, Heinz 1987. 'Abstandsprache und Ausbausprache', in Ammon, Ulrich, Dittmar, Norbert and Mattheier, Klaus (eds), *Sociolinguistics/Soziolinguistik*, vol. I. Berlin/New York: De Gruyter, pp. 302–8.
- Koch, Peter 1987. *Distanz im Dictamen. Zur Schriftlichkeit und Pragmatik mittelalterlicher Brief- und Redemodelle in Italien*. Freiburg im Breisgau: Habilitationsschrift.
1988. 'Italienisch: Externe Sprachgeschichte I', in *LRL* (IV), pp. 343–60.
1993. 'Pour une typologie conceptionnelle et médiale des plus anciens documents/monuments des langues romanes', in Selig, Frank and Hartmann (1993), pp. 39–81.

1995. 'Une langue comme toutes les autres: latin vulgaire et traits universels de l'oral', in Callebat, Louis (ed.), *Latin vulgaire–latin tardif IV*. Hildesheim: Olms-Weidmann, pp. 125–44.
1997. 'Diskurstraditionen: zu ihrem sprachtheoretischen Status und ihrer Dynamik', in Frank, Barbara, Haye, Thomas and Tophinke, Doris (eds), *Gattungen mittelalterlicher Schriftlichkeit*. Tübingen: Narr, pp. 43–79.
2004. 'Sprachwandel, Mündlichkeit und Schriftlichkeit', *ZRPh* 120:605–30.
2006. 'Romanische Sprachgeschichte und Varietätenlinguistik', in Ernst, Gleßgen, Schmitt and Schweickardt (2003–8), pp. 102–23.
2008. 'Le latin – langue diglossique?', in von Moos (2008), pp. 287–316.
- Koch, Peter and Oesterreicher, Wulf 1985. 'Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte', *Romanistisches Jahrbuch* 36:15–43.
2001. 'Langage parlé et langage écrit', in *LRL* (I.2), pp. 584–627.
2008. 'Comparaison historique de l'architecture des langues romanes', in Ernst et al. (2003–8), pp. 2575–610.
2011. *Gesprochene Sprache in der Romania. Französisch, Italienisch, Spanisch*. Berlin/New York: De Gruyter.
- Kontzi, Reinhold (ed.) 1982. *Substrate und Superstrate in den romanischen Sprachen*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Kopitar, Jernej 1829. 'Albanische, walachische und bulgarische Sprache', *Jahrbücher der Literatur* 46:59–106.
- Kouwenberg, Silvia and Singler, John Victor (eds) 2008. *The Handbook of Pidgin and Creole Studies*. Oxford: Blackwell.
- Kovačec, August 1963. 'Notes de lexicologie istro-roumaine: sur la disparition des mots anciens et leur remplacement par des mots croates', *Studia Romanica et Anglica Zagabriensia* 15/16:3–39.
1966. 'Quelques influences croates dans la morphosyntaxe istroroumaine', *Studia Romanica et Anglica Zagabriensia* 21/22:57–76.
1968. 'Observations sur les influences croates dans la grammaire istroroumaine', *La Linguistique* 1:79–115.
1971. *Descrierea istroromânei actuale*. Bucharest: Editura Academiei.
1984. 'Istroromâna', in Rusu (1984), pp. 550–91.
1998. *Istrorumunjski-hrvatski rječnik (s gramatikom i tekstovima)*. Pula: Mediteran.
- Kramer, Johannes 1989. 'Ladinisch: Grammatikographie und Lexicographie: Grammaticografia e lessicografia', in *LRL* (III), pp. 757–63.
1998. 'Warum die Alltagssprache des 6 Jh. nicht zur Literatursprache wurde', in Herman (1998a), pp. 27–40.
1999. 'Sind die romanischen Sprachen kriolisiertes Latein?', *ZRPh* 115:1–19.
- Krefeld, Thomas 1988. 'Italienisch: Periodisierung', in *LRL* (IV), pp. 748–62.
- Krefeld, Thomas and Schmitt, Christian 1989. 'Rumänisch: Diglossie und Polyglossie', in *LRL* (III), pp. 229–39.
- Kremnitz, Georg 1991. 'Okzitanisch: externe Sprachgeschichte', in *LRL* (V.2), pp. 69–80.
- Kristol, Andres 1989. 'Bündnerromanisch: Soziolinguistik', in *LRL* (III), pp. 813–26.
1998. 'Que reste-t-il des dialectes gallo-romans de Suisse romande?', in Eloy, Jean-Michel (ed.), *Évaluer la vitalité. Variétés d'oil et autres langues. Actes du Colloque international*

- «Évaluer la vitalité des variétés régionales du domaine d'oil». Amiens: Université de Picardie – Jules Verne, Centre d'Etudes Picardes, pp. 101–14.
- Krusch, Bruno and Levison, Wilhelm 1951. *Gregorii turonensis opera. Libri historiarum X. Monumenta Germaniae historica. Scriptores rerum Merovingicarum*, 1,1. Hannover: Hahn.
- Kukenheim, Louis 1932. *Contributions à l'histoire de la grammaire italienne, espagnole et française à l'époque de la Renaissance*. Amsterdam: Noord-Holland.
- Labov, William 1972a. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- 1972b. *Language in the Inner City. Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
1975. 'The use of the present to explain the past', in Heilmann, Luigi (ed.), *Proceedings of the 11th International Congress of Linguistics*. Bologna: Il Mulino, pp. 825–51.
1994. *Principles of Linguistic Change, volume 1: Internal Factors*. Oxford: Blackwell.
- Lafage, Suzanne 1993. 'French in Africa', in Sanders, Carol (ed.), *French Today*. Cambridge: Cambridge University Press, pp. 215–38.
- Lafont, Robert 1974. *La Revendication occitane*. Paris: Flammarion.
- Lagueux, Paul-André 2005. 'Caractéristiques morpho-syntaxiques du français québécois', in Brasseur and Falkert (2005), pp. 57–69.
- Laliena, Carlos and Sénac, Philippe 1991. *Musulmans et chrétiens dans le haut Moyen Age. Aux origines de la reconquête aragonaise*. Paris: Minerve.
- Lapesa, Rafael 1980⁸/1981⁹. *Historia de la lengua española*. Madrid: Gredos.
- Lass, Roger 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
2004. 'Ut custodiant litteras: editions, corpora and witnesshood', in Dossena and Lass (2004), pp. 21–48.
- Lathuillère, Roger 1966. *La Préciosité: étude historique et linguistique. Tome 1, Position du problème, les origines*. Geneva: Droz.
- Lauer, Philippe 1926. *Nithard. Histoire des fils de Louis le Pieux*. Paris: Champion.
- Lausberg, Heinrich 1956–62. *Romanische Sprachwissenschaft, I. Einleitung und Vokalismus, II. Konsonantismus, III. Formenlehre*. Berlin: de Gruyter.
- Law, Vivien 1982. *The Insular Latin Grammarians*. Woodbridge: Boydell.
- Lazard, Sylviane 2007. 'La crise de la scripta au XIe siècle: ce que nous apprend le plaid de Lusignano de 915', in Trotter, David (ed.), *Actes du XXIXe Congrès de linguistique et de philologie romanes, II*, pp. 511–26.
- Lázaro Carreter, Fernando 1972². *Crónica del Diccionario de Autoridades (1726–1740)*. Madrid: Real Academia Española.
1985. *Las ideas lingüísticas en España durante el siglo XVIII*. Barcelona: Editorial Crítica.
- Le Goff, Jacques 1977. *Pour un autre Moyen Age*. Paris: Gallimard.
- Le Jan, Régine 2006. *Les Mérovingiens*. Paris: PUF.
- Le Page, Robert (ed.) 1961. *Proceedings of the Conference on Creole Language Studies Held at the University College of the West Indies March 28–April 4, 1959*. London: Macmillan.
1988. 'Some premises concerning the standardization of languages, with special reference to Caribbean Creole English', *International Journal of the Sociology of Language* 71:25–36.
- Le Page, Robert and Tabouret-Keller, Andrée 1985. *Acts of Identity. Creole-based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.

- Ledgeway, Adam 2002. 'Linguistic theory and the mysteries of Italian dialects', in Lepschy and Tosi (2002), pp. 108–40.
2006. 'The dual complementiser system in southern Italy: *spirito greco, materia romanza*?', in Lepschy and Tosi (2006), pp. 112–26.
2011. 'Adverb agreement and split intransitivity: evidence from southern Italy', *AGI* 96:31–66.
- Lefebvre, Claire 1985. 'Grammaires en contact: définition et perspectives de recherche', *Revue québécoise de linguistique* 14:11–47.
1998. *Creole Genesis and the Acquisition of Grammar. The Case of Haitian Creole*. Cambridge: Cambridge University Press.
- Lentner, Leopold 1963. *Volkssprache und Sakralsprache. Geschichte einer Lebensfrage bis zum Ende des Konzils von Trient*. Vienna: Herder.
- Lenz, Rodolfo 1928. *El Papiamento, la lengua criolla de Curazao. La gramática más sencilla*. Santiago de Chile: Anales de la Universidad.
- Lepschy, Anna Laura and Tosi, Arturo (eds) 2002. *Multilingualism in Italy. Past and Present*. Oxford: European Humanities Research Centre / University of Oxford, Modern Humanities Research Association, Legenda.
- (eds) 2006. *Rethinking Languages in Contact. The Case of Italian*. Oxford: Legenda.
- Lepschy, Anna Laura, Lepschy, Giulio and Voghera, Miriam 1996. 'Linguistic variety in Italy', in Levy, Carl (ed.), *Italian Regionalism. History, Identity and Politics*. Oxford/Washington: Berg, pp. 69–80.
- Lepschy, Giulio (ed.) 1994. *Storia della linguistica*, vol. III. Bologna: Il Mulino.
1996. 'Dialect and language: a brief introduction', *Journal of the Institute of Romance Studies*, suppl. 1:1–6.
- (ed.) 1998. *History of Linguistics. III. Renaissance and Early Modern Linguistics*. London and New York: Longman.
2002. 'What is the standard?', in Lepschy and Tosi (2002), pp. 74–81.
- Lévi-Strauss, Claude 1962. *La Pensée sauvage*. Paris: Plon.
- Lewitt-Gibbon, Tamara 2001. 'Changing concepts of Late Antiquity: the decline and fall of Gibbonism', *Bulletin de l'Antiquité Tardive* 10:33–37.
- Librandi, Rita 1993. 'L'italiano nella comunicazione della Chiesa e nella diffusione della cultura religiosa', in Serianni and Trifone (1993–94), vol. I, pp. 335–81.
- Lidzbarski, Mark 1907. *Kanaanäische Inschriften*. Giessen: Töpelmann.
- LIE 1990 = Asor Rosa, Alberto (ed.), *Letteratura italiana. Gli autori. Dizionario bio-bibliografico e indice*. Turin: Einaudi.
- Lightfoot, David 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Lipski, John 1985. *The Spanish of Equatorial Guinea. The Dialect of Malabo and its Implications for Spanish Dialectology*. Tübingen: Niemeyer.
1990. *The Language of the Isleños*. Baton Rouge/London: Louisiana State University Press.
1994. *Latin American Spanish*. London: Longman.
2005. *A History of Afro-Hispanic Language. Five Centuries, Five Continents*. Cambridge: Cambridge University Press.
- Liver, Ricarda 1993. 'Le démonstratif dans la version interlinéaire d'Einsiedeln', in Selig et al. (1993), pp. 181–91.

1995. 'Bündnerromanisch', in *LRL* (III), pp. 68–81.
1999. *Rätoromanisch: eine Einführung in das Bündnerromanische*. Tübingen: Narr.
- Lleal, Coloma 1990. *La formación de las lenguas romances peninsulares*. Barcelona: Barcanova.
- Llera Ramo, Francisco, Iraola Alkorta, Yosune and San Martín Antuña, Pablo (eds) 2001. *Estudio sociolingüístico de las hablas del Alto Aragón*. [unpublished]
- Llera Ramo, Francisco and San Martín Antuña, Pablo. 2003. *Estudio sociolingüístico de Asturias*. 2002. Oviedo: Academia de la Llingua Asturiana.
- Llibre blanc 2005 = Llibre blanc de l'ús del valencià – I. Enquesta sobre la situació social del valencià*. 2004. Valencia: Acadèmia Valenciana de la Llengua.
- LN = *Lingua nostra*
- Lodge, Anthony 1993. *French. From Dialect to Standard*. London/New York: Routledge.
2011. 'Standardisation and Koinéisation: Deux approches contraires à l'historiographie d'une langue', in Dessì Schmid et al. (eds) (2011), pp. 65–79.
- Löfstedt, Bengt 1961. *Studien über die Sprache der langobardischen Gesetze*. Stockholm: Almqvist & Wiksell.
- Löfstedt, Einar 1911. *Philologischer Kommentar zur Peregrinatio Aetheriae*. Uppsala: Almqvist & Wiksell.
1933. *Syntactica: Studien und Beiträge zur historischen Syntax des Lateins. Zweiter Teil: Syntaktisch-stilistische Gesichtspunkte und Probleme*. Lund: Gleerup.
1942. *Syntactica: Studien und Beiträge zur historischen Syntax des Lateins. Erster Teil: Über einige Grundfragen der lateinischen Nominalsyntax*. Lund: Gleerup.
- Lope Blanch, Juan Miguel 1987. *Estudios sobre el español de Yucatán*. México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas, Centro de Lingüística Hispánica.
- López García, Ángel 2000. *Cómo surgió el español*. Madrid: Gredos.
- López Morales, Humberto 2005. 'La actuación de las academias en la historia del idioma', in Cano Aguilar (2005), pp. 919–40.
- Lorenzino, Gerardo 1993. 'African vs. Austronesian substrate influence on the Spanish-based creoles', in Byrne and Holm (1993), pp. 399–408.
- Lorenzo, Ramón 1975. *La traducción gallega de la Crónica General y de la Crónica de Castilla*. Orense: Instituto Padre Feijoo.
- Lot, Ferdinand 1931. 'A quelle époque a-t-on cessé de parler latin?', *Archivum Latinitatis Medii Aevi (Bulletin Du Cange)* 6:97–159.
- LRL* = Holtus, Günter, Metzeltin, Michael and Schmitt, Christian (eds) 1988–2004. *Lexikon der romanistischen Linguistik*. Tübingen: Niemeyer.
- Lucchi, Piero 1985. 'La prima istruzione. Idee, metodi, libri', in Brizzi, Gian Paolo (ed.), *Il catechismo e la grammatica. I. Istruzione e controllo sociale nell'area emiliana e romagnola nel '700*. Bologna: Il Mulino, pp. 25–81.
- Lüdi, Georges 1990. 'Diglossie et polyglossie', in *LRL* (V.1), pp. 307–34.
1992. 'French as a pluricentric language', in Clyne, Michael (ed.), *Pluricentric Languages*. Berlin: Mouton de Gruyter, pp. 149–78.
- Lüdi, Georges and Werlen, Iwar 2005. *Recensement fédéral de la population 2000. Le Paysage linguistique en Suisse*. Neuchâtel: Office fédéral de la statistique.
- Lüdtke, Helmut 1957. 'Inchiesta sul confine dialettale fra il veneto e il friulano', *Orbis* 6:118–21.
1964. 'Die Entstehung romanischer Schriftsprachen', *VR* 23:3–21.

- 1993a. 'La reconstruction de la prononciation latine d'époque républicaine et de son évolution ultérieure', in Lorenzo, Ramón (ed.), *Actas do XIX Congreso Internacional de Lingüística e Filoloxía Románicas, Universidade de Santiago de Compostela, 1989* (V). A Coruna: Fundación Pedro Barrié de la Maza, Conde de Fenosa, pp. 43–45.
- 1993b. 'Theoretische und historische Bemerkungen zum Verhältnis von Mündlichkeit und Schiffllichkeit in der mittelalterlichen Romania', in Werner, Otmar (ed.), *Probleme der Graphie*. Tübingen: Narr, pp. 53–61.
2000. 'Latino vivo – vernacolo – Lingua "morta". I precedenti dell'assetto linguistico culturale nell'Occidente (Post)Carolingio', in Fusco, Fabiana, Orioles, Vincenzo and Parmeggiani, Alice (eds), *Processi di convergenza e differenziazione nelle lingue dell'Europa medievale e moderna*. Udine: Forum, pp. 15–44.
2005. *Der Ursprung der romanischen Sprachen. Eine Geschichte der Sprachlichen Kommunikation*. Kiel: Westensee.
- Lüdtke, Jens 1991. 'Katalanisch: externe Sprachgeschichte', in *LRL* (V.2), pp. 232–42.
2001. 'Romanische Philologie von Dante bis Raynouard: Diachrone romanische Sprachwissenschaft und Sprachgeschichtsschreibung', in *LRL* (I.1), pp. 1–35.
- Lurçat, François 1999. *Le Chaos*. Paris: PUF.
- Lusignan, Serge 1986. *Parler vulgairement*. Paris: Vrin.
- Lutz, Florentin 1989. 'Bündnerromanisch: Grammatikographie und Lexikographie. Grammaticographie et lexicographie. a) Grammatikographie. Grammaticographie', in *LRL* (III), pp. 886–97.
- Machado, Álvaro Manuel 1984. *O 'francesismo' na literatura portuguesa*. Lisbon: Biblioteca Breve.
- Mackey, William 1970. *Interference, Integration and the Synchronic Fallacy*. Québec: Centre international sur le bilinguisme.
1976. *Bilinguisme et contact des langues*. Paris: Klincksieck.
- Magie, David 1967. *The Scriptores historiae augustae*. I. Cambridge, MA: Harvard University Press.
- Maia, Clarinda de Azevedo 1986. *História do galego-português. Estado linguístico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI (com referência à situação do galego moderno)*. Coimbra: Instituto nacional de Investigação Científica.
- Maiden, Martin 1995. *A Linguistic History of Italian*. London/New York: Longman.
1996. 'On the Romance inflectional endings -i and -e', *RPh* 50:147–82.
1999. 'Il ruolo dell' "idoneità" nella morfologia diacronica. I suffissi romeni -ea, -icã ed -oi', *RLiR* 63:321–45.
2000. 'Il sistema desinenziale del sostantivo nell'italo-romanzo preletterario. Ricostruzione parziale a partire dai dialetti moderni (il significato storico di plurali del tipo amici)', in Herman, József, Marinetti, Anna and Mondin, Luca (eds), *La preistoria dell'italiano*. Niemeyer: Tübingen, pp. 167–79.
2006. 'Accommodating synonymy: how some Italo-Romance verbs react to lexical and morphological borrowings', in Lepschy and Tosi (2006), pp. 87–98.
- Mair, Walter 1989. 'Ladinisch: Soziolinguistik', in *LRL* (III), pp. 697–704.
- Malato, Enrico (ed.) 1997. *Storia della letteratura italiana. V. La fine del Cinquecento e il Seicento*. Rome: Salerno editrice.
- Malkiel, Yakov 1978. 'Critères pour l'étude de la fragmentation du latin', in *Atti XIV Congresso internazionale di linguistica e filologia romanza*. I. Naples/Amsterdam: Macchiaroli/Benjamins, pp. 27–47.

- Malmberg, Bertil 1990. Review of Sala 1988, *Studia Linguistica* 44:82–85.
- Mandouze, André 1968. *Saint Augustin. L'Aventure de la raison et de la grâce*. Paris: Etudes augustiniennes.
- Manessy, Gabriel 1981. *Les Langues de l'Afrique subsaharienne*. Volume 1 of Perrot, Jean (ed.), *Les Langues dans le monde ancien et moderne*. Paris: Éditions du Centre national de la recherche scientifique.
- Maraschio, Nicoletta 1977. 'Il parlato nella speculazione linguistica del Cinquecento', *Studi di grammatica italiana* 6:207–26.
- Marazzini, Claudio 1992. 'Il Piemonte e la Valle d'Aosta', in Bruni (1992), pp. 1–44.
- 1993a. *Il secondo Cinquecento e il Seicento*. Bologna: il Mulino.
- 1993b. 'Le teorie', in Serianni and Trifone (1993–94), 1, pp. 231–329.
1994. *La lingua italiana. Profilo storico*. Bologna: Il Mulino.
- Marcato, Carla 1989a. 'Friaulisch: externe Sprachgeschichte. Storia linguistica esterna', in *LRL* (III), pp. 617–27.
- 1989b. 'Friaulisch: Grammatikographie und Lexikographie. Grammaticografia e lessicografia', in *LRL* (III), pp. 637–45.
2002. 'Friuli-Venezia Giulia', in Cortelazzo *et al.* (2002), pp. 329–56.
- Marcellesi, Jean Baptiste and Thiers, Ghjacumu 1988. 'Korsisch: Soziolinguistik; Sociolinguistique', in *LRL* (IV), pp. 809–20.
- Marchello-Nizia, Christiane 1995. *L'Évolution du français. Ordre des mots, démonstratifs, accent tonique*. Paris: Armand Colin.
- Marcos Marín, Francisco 1992. 'Spanisch: Periodisierung', in *LRL* (VI.1), pp. 602–7.
- Mariño, Ramón 2008. *Historia de la lengua gallega*. Munich: Lincom.
- Márquez-Sterling, Manuel 1980. *Fernán González, First Count of Castile. The man and the legend*. Mississippi: University of Mississippi.
- Marrou, Henri Irénée 1958. *Saint Augustin et la fin de la culture antique*. Paris: E. de Boccard.
1971. *Les Troubadours*. Paris: Edition du Seuil.
- Martí i Castell, Joan and Moran, Josep 1986. *Documents d'història de la llengua catalana. Dels orígens a Fabra*. Barcelona: Editorial d'Empúries.
- Martin, Georges 1997. *Histoires de l'Espagne médiévale. Historiographie, geste, romancero*. Paris: Klincksieck.
- Martin, Jean-Baptiste 1990. 'Französisch: Frankoprovençalisch. Francoprovençal', in *LRL* (V.1), pp. 671–85.
- Martín Abad, Julián 2003. *Los primeros tiempos de la imprenta en España (c. 1471–1520)*. Madrid: Ediciones del Laberinto.
- Martín Zorraquino, Maria Antonia and Enguita Utrilla, José Maria 2000. *Las lenguas de Aragón*. Zaragoza: Caja de Ahorros de la Inmaculada.
- Martinell, Emma 1984. 'Posturas adoptadas ante los galicismos introducidos en el castellano en el siglo XVIII', *Revista de filología* 3:102–28.
- Martinet, André 1945 [=1939]. 'La description phonologique du parler franco-provençal d'Hauteville (Savoie)', *RLiR* 15:1–86.
1953. Preface to Weinreich (1953).
1956. *La Description phonologique avec application au parler franco-provençal d'Hauteville (Savoie)*. Geneva and Paris: Droz and Minard.
- Martino, Paolo 1991. *L'area Lausberg'. Isolamento e arcaicità*. Rome: Università di Roma La Sapienza.

- Martins, Ana Maria 1994. *Clíticos na história do português*. Faculdade de Letras da Universidade de Lisboa: unpublished doctoral thesis.
- Martinus, Efraim Frank 1990. 'Papiamentu: the road to emancipation', in Fodor, István and Hagège, Claude (eds), *Language Reform. History and Future*. V. Hamburg: Buske, pp. 127–49.
1997. *The Kiss of a Slave. Papiamentu's West-African Connections*. Curaçao: De Curaçaosche Courant.
- Massa, Jean-Michel 1994. 'Portugiesisch in Afrika. La langue portugaise en Afrique. a) Ausdehnung und Verbreitung des Portugiesischen. Expansion et diffusion du portugais', in *LRL* (VI.2), pp. 575–84.
- Massignon, Geneviève 1962. *Les Parlers français d'Acadie. enquête linguistique*. Paris: Klincksieck.
- Matarrese, Tina 1993. *Il Settecento*. Bologna: Il Mulino.
- Matthews, David 2000. *The Invention of Middle English*. Philadelphia: Penn State University Press.
- Mattoso Câmara Jr, Joaquim 1972. *The Portuguese Language*. Chicago and London: Chicago University Press.
- Maurer, Philippe 1988. *Les Modifications temporelles et modales du verbe dans le Papiamentu de Curaçao (Antilles Néerlandaises). Avec une anthologie et un vocabulaire Papiamentu-français*. Hamburg: Buske.
1998. 'El papiamentu de Curazao', in Perl and Schwegler (1998), pp. 139–217.
- McKitterick, Rosamond 1989. *The Carolingians and the Written Word*. Cambridge: Cambridge University Press.
- (ed.) 1990. *The Uses of Literacy in Early Mediaeval Europe*. Cambridge: Cambridge University Press.
1991. 'Latin and Romance, an historian's perspective', in Wright (1991), pp. 130–45.
- (ed.) 1994. *Carolingian Culture. Emulation and Innovation*. Cambridge: Cambridge University Press.
1999. 'Die karolingische Renovatio. Eine Einführung', in Stiegemann, Christoph and Wemhoff, Matthias (eds), *799: Kunst und Kultur der Karolingerzeit. Karl der Grosse und Papst Leo III. in Paderborn*, vol. II. Mainz: von Zabern, pp. 668–85.
- McLaughlin, Mairi 2011. *Syntactic Borrowing in Contemporary French. A Linguistic Analysis of News Translation*. Oxford: Legenda.
- McWhorter, John 1998. 'Identifying the creole prototype: vindicating a typological class', *Language* 74:788–818.
2001. 'The world's simplest grammars are creole grammars', *Linguistic Typology* 5:125–66.
- Meier, Harri 1930. *Beiträge zur sprachlichen Gliederung der Pyrenäenhalbinsel*. Hamburg: Romanisches Seminar.
- Meillet, Antoine 1951. *Linguistique historique et linguistique générale*. I. Paris: Champion.
- Meillet, Antoine and Sauvageot, Aurélien 1934. 'Le bilinguisme des hommes cultivés', *Conférences de l'Institut de Linguistique de l'Université de Paris* 2:5–14.
- Melazzo, Lucio 1984. 'Latino e greco in Sicilia', in Quattordio Moreschini, Adriana (ed.), *Tre millenni di storia linguistica della Sicilia: atti del Convegno della Società Italiana di Glottologia: Palermo, 25–27 marzo 1983*. Pisa: Giardini, pp. 37–54.
- Melillo, Michele 1959. 'Intorno alle probabili sedi originarie delle colonie francoprovenzali di Celle e Faeto', *RLiR* 23:1–34.

- Mendes Drumond Braga, Isabel 2001. 'As realidades culturais', in de Freitas de Meneses, Avelino (ed.), *Nova história de Portugal. VII. Portugal da paz da Restauração ao ouro do Brasil*. Lisbon: Editorial Presença, pp. 465–565.
- Menéndez Pidal, Ramón (ed.) 1919. *Documentos lingüísticos de España, I, Reino de Castilla*. Madrid: Centro de Estudios Históricos.
1926. *Orígenes del español. Estado lingüístico de la Península Ibérica hasta el siglo XI*. Madrid: Espasa Calpe.
- 1956 (4th edn). *Orígenes del español. Estado lingüístico de la Península Ibérica hasta el siglo XI*. Madrid: Espasa Calpe.
1962. *El dialecto leonés*. Oviedo: Instituto de Estudios Asturianos.
- Menéndez Pidal, Ramón 1964 (5th edn). *Orígenes del español. Estado lingüístico de la Península Ibérica hasta el siglo XI*. Madrid: Espasa Calpe.
2005. *Historia de la lengua española*. Madrid: Fundación Menéndez Pidal.
- Meo Zilio, Giovanni 1989. *Estudios hispanoamericanos. Temas lingüísticos*. Rome: Bulzoni.
- Merlo, Clemente 1927. 'Lazio sannita e Etruria latina', *L'Italia dialettale* 3:84–93.
1941. *La Francia linguistica odierna e la Gallia di Giulio Cesare*. Rome: Reale Accademia d'Italia.
- Messner, Dieter 1994. 'Portugiesisch: Periodisierung', in *LRL* (VI.2), pp. 618–23.
- Metzeltin, Michele 1988. 'Italianisch: Externe Sprachgeschichte. Storia della lingua. b) Externe Sprachgeschichte II. Gesellschaft und Sprache im 19. und 20. Jahrhundert. Storia della lingua II. Società e lingua nell'800 e '900', in *LRL* (IV), pp. 343–79.
- Meyer-Lübke, Wilhelm 1890. *Grammatik der romanischen Sprachen. I. Lautlehre*. Leipzig: Reisland.
1895. *Grammatik der romanischen Sprachen. II: Formenlehre*. Leipzig: Reisland.
1898. 'Rum. spre', *ZRPh* 22:492–96.
1899. *Grammatik der Romanischen Sprachen. III: Syntax*. Leipzig: Reisland.
1902. *Grammatik der romanischen Sprachen. IV. Register*. Leipzig: Reisland.
- MGH = *Monumenta Germaniae Historica* 1819–. Hannover: Hahn.
- Michaelis, Susanne 2000. 'The fate of subject pronouns: evidence from creole and non-creole languages', in Neumann-Holzschuh and Schneider (2000), pp. 163–83.
- Miestamo, Matti, Sinnemäki, Kaius and Karlsson, Fred 2009. *Language Complexity. Typology, Contact, Change*. Amsterdam/Philadelphia: Benjamins.
- Migliorini, Bruno 1961. *Storia della lingua italiana*. Florence: Sansoni.
- Migliorini, Bruno and Griffith, Gwynfor 1984. *The Italian Language*. London: Faber and Faber.
- Mihăescu, Haralambie 1966. *Influența grecească asupra limbii române pînă în secolul al XV-lea*. Bucharest: Editura Academiei Republicii Socialiste România.
1978. *La Langue latine dans le sud-est de l'Europe*. Bucharest/Paris: Editura Academiei / Les Belles Lettres.
- Millar, Fergus 1984. 'Epigrafia', in Crawford, Michael, Gabba, Emilio, Millar, Fergus and Snodgrass, Anthony, *Le basi documentarie della storia antica*. Bologna: Il Mulino, pp. 85–137.
- Miller, Jim and Fernandez-Vest, Jocelyne 2006. 'Spoken and written language', in Bernini, Giuliano and Schwartz, Marcia (eds), *Pragmatic Organisation of Discourse in the Languages of Europe*. Berlin/New York: de Gruyter, pp. 9–64.
- Millet-Gérard, Dominique 1984. *Chrétiens mozarabes et culture islamique dans l'Espagne des VIIIe–IXe siècles*. Paris: Institut des études augustiniennes.

- Milroy, James 2001. 'Language ideologies and the consequences of standardization', *Journal of Sociolinguistics* 5:530–55.
- Milroy, James and Milroy, Lesley 1985. 'Linguistic change, social network and speaker innovation', *Journal of Linguistics* 21:339–84.
- Minerva, Nadia 2000. 'Le donne e la grammatica. Su alcune "Grammaires des Dames" tra Sette e Ottocento', in Minerva, Nadia (ed.), *Dames, demoiselles, honnêtes femmes. Studi di lingua e letteratura francese offerti a Carla Pellandra*. Bologna: CLUEB, pp. 73–105.
- Minerva, Nadia, and Pellandra, Carla (eds) 1997. *Insegnare il francese in Italia. Repertorio di manuali pubblicati dal 1625 al 1860*. Bologna: CLUEB.
- Mingas, Amélia 2000. *Interferência do kimbundu no português falado em Lwanda*. Porto: Campo das Letras.
- Mioni, Alberto 1983. 'Italiano tendenziale. Osservazioni su alcuni aspetti della standardizzazione', in Benincà, Paola, Cortelazzo, Michele, Prosdocimi, Aldo, Vanelli, Laura and Zamboni, Alberto (eds), *Scritti linguistici in onore di Giovan Battista Pellegrini*. Pisa: Pacini, pp. 495–517.
- Miracle, Josep 1960. *La Restauració dels Jocs Florals*. Barcelona: Aymà.
- Miret i Sans, Joaquin 1915. *Antics documents de llengua catalana*. Barcelona: Casa Provincial de Caritat.
- Miron, Paul 1978. *Der Wortschatz Dimitrie Cantemirs. Eine lexikalische Untersuchung von Divanul und Istoria Ieroglifică in Vergleich zu Texten aus dem XVI., XIX. und XX. Jahrhundert*. Frankfurt/Bern/Las Vegas: Peter Lang.
- Monaci, Ernesto 1955. *Crestomazia italiana dei primi secoli*. Rome/Naples/Città di Castello: Società Editrice Dante Alighieri.
- Monfrin Jacques 1968. 'Le mode de tradition des actes écrits et les études de dialectologie', *RLiR* 32:17–47.
- Monjour, Alf 1995. 'Galegische und portugiesische Skriptae', in *LRL* (II.2), pp. 692–720.
- Montanyés Gómez, Ruben. 2005. 'El judeospanyol: concepte i orígens, història, situació actual', in Burdeus, María Dolores and Verdegel Cerezo, Joan Manuel (eds). *Europa Parla (I). Anuari de l'agrupació borrianenca de cultura* 16:97–123.
- Montoliu, Manuel de 1962. *La Renaixença i els Jocs Florals. Verdaguer*. Barcelona: Editorial Alpha.
- Moos, Peter von (ed.) 2008. *Zwischen Babel und Pfingsten / Entre Babel et Pentecôte, Sprachdifferenzen und Gesprächsverständigung in der Vormoderne (8.–16. Jahrhundert). Différences linguistiques et communication orale avant la modernité (VIII^e–XVI^e siècle)*. Zurich/Berlin: LIT-Verlag.
- Morala, José 2002. 'Originales y copias. El proceso de castellanización en el área leonesa', in Echenique Elizondo, María Teresa and Sánchez Méndez, Juan (eds), *Actas del Quinto Congreso Internacional de Historia de la Lengua Española*, I. Madrid: Gredos, pp. 1335–45.
2005. 'Del leonés al castellano', in Cano Aguilar (2005), pp. 555–69.
- Moran, Josep 2004. 'El proceso de creación del catalán escrito', *Aemilianense* 1:431–55.
- Moretti, Bruno and Spiess, Federico 2002. 'La Svizzera italiana', in Cortelazzo et al. (2002), pp. 261–75.
- Moreux, Bernard 2004. 'Béarnais and Gascon today: language behavior and perception', *International Journal of the Sociology of Language* 169:25–62.
- Morf, Heinrich 1911. 'Zur sprachliche Gliederung Frankreichs', in *Abhandlungen der Königlich Preussischen Akademie der Wissenschaften zu Berlin*. Berlin: Phil.-hist. Classe, Mem. II, pp. 1–37.

- Morgana, Silvia 1994. 'L'influsso francese', in Serianni and Trifone (1993–94): III, pp. 671–719.
- Mørland, Henning 1932. *Die lateinischen Oribasiusübersetzungen*. Oslo: Brøgger.
- Morosi, Giacomo 1890–92. 'Il dialetto franco-provenzale di Faeto e Celle, nell'Italia meridionale', *AGI* 12:33–75.
- Morpurgo Davies, Anna 1998. 'Nineteenth-Century Linguistics' (*History of Linguistics, Volume IV*). London/New York: Longman.
- Mortara Garavelli, Bice 2008. *Storia della punteggiatura in Europa*. Bari/Rome: Laterza.
- Mossé, Fernand 1945. *Manuel de l'anglais du Moyen Age, 1, Vieil-anglais*. Paris: Aubier, Éditions Montaigne.
- Mougeon, Raymond 2004. 'Perspective sociolinguistique sur le français en Ontario', in Coveney *et al.* (2004), pp. 155–90.
- Mougeon, Raymond and Beniak, Edouard 1989. 'Language contraction and linguistic change: the case of Welland French', in Dorian, Nancy (ed.), *Investigating Obsolescence*. Cambridge: Cambridge University Press, pp. 287–312.
1991. *Linguistic Consequences of Language Contact and Restriction. The Case of French in Ontario, Canada*. Oxford: Clarendon Press.
- MPM 1997. *Pratiques et représentations de la langue occitane en Aquitaine. Décembre 1997*. Sondage du Conseil Régional d'Aquitaine, enquête réalisée par Média Pluriel Méditerranée. http://crdp.ac-bordeaux.fr/langues/sondage_aquitaine.htm (last accessed 1 October 2007).
- MPM 1998. *Pratiques et représentations de l'occitan. Région Languedoc-Roussillon, janvier 1998*, enquête réalisée par Média Pluriel Méditerranée. Castelnau-le-Lez: Média Pluriel Méditerranée.
- MSG 2004 = González González, Manuel *et al.* (eds) 2007. *Mapa sociolingüístico de Galicia 2004. Vol. I: Lingua inicial e competencia lingüística en Galicia*. Coruña: Real Academia Galega, Seminario de sociolingüística.
- Mufwene, Salikoko 1986. 'The universalist and substrate hypothesis complement one another', in Muysken and Smith (1986), pp. 129–62.
1996. 'The founder principle in creole genesis', *Diachronica* 13:83–134.
1997. 'Jargons, pidgins, creoles, and koines: what are they?', in Spears and Winford (1997), pp. 35–70.
2000. 'Creolization is a social, not a structural, process', in Neumann-Holzschuh and Schneider (2000), pp. 65–84.
2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Mühlhäusler, Peter [1986] 1997 (2nd edn). *Pidgin and Creole Linguistics*. London: University of Westminster Press.
- Muljačić, Žarko 1978. 'Lingue romanze e lingue slave', in *Atti XIV Congresso internazionale di linguistica e filologia romanza*. I. Naples/Amsterdam: Macchiaroli/Benjamins, pp. 407–21.
1993. 'Standardization in Romance', in Posner and Green (1980–93), pp. 77–114.
1995. 'Dalmatisch', in *LRL* (III.2), pp. 32–42.
2000. *Das Dalmatische. Studien zu einer untergegangenen Sprache*. Cologne/Weimar/Vienna: Böhlau.
- Müller, Bodo 1985. *Le Français d'aujourd'hui*. Paris: Klincksieck.
1996. 'Bezeichnungen für die Sprachen, Sprechen und Länder der Romania', in *LRL* (II.1), pp. 134–51.

- Müller, Roman 2001. *Sprachbewusstsein und Sprachvariation im lateinischen Schrifttum der Antike*. Munich: Beck.
- Munteanu, Dan 1996. *El papiamento, lengua criolla hispánica*. Madrid: Gredos.
- Muysken, Pieter 1977. 'Media lengua', in Thomason, Sarah (ed.), *Contact Languages. A Wider Perspective*. Amsterdam: Benjamins, pp. 365–426.
1988. 'Are creoles a special type of language?', in Newmeyer, Fritz (ed.), *Linguistics. The Cambridge Survey II*. Cambridge: Cambridge University Press, pp. 285–301.
- Muysken, Pieter and Smith, Norval (eds) 1986. *Substrata versus Universals in Creole Genesis. Papers from the Amsterdam Creole Workshop, April 1985*. Amsterdam/Philadelphia: Benjamins.
- Myers-Scotton, Carol 1993a. *Social Motivations for Codeswitching*. Oxford: Clarendon Press.
- 1993b. *Duelling Languages. Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.
2002. *Contact Linguistics. Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, Carol and Okeju, John 1973. 'Neighbours and lexical borrowings', *Language* 49:871–89.
- Nadal, Josep and Prats, Modest 1996. *Història de la llengua catalana. II. Segle XV*. Barcelona: Edicions 62.
- Nagore Laín, Francho 2001. *Os territorios lingüísticos en Aragón*. Zaragoza: Rolde de Estudios Aragoneses.
- Nandriş, Octave 1961. 'Le genre, ses réalisations et le genre personnel en roumain', *RLiR* 25:47–74.
- Naro, Anthony 1978. 'A study on the origins of pidginization', *Language* 54:314–47.
- Naro, Anthony and Scherre, Maria Marta Pereira 1999. 'Variable concord in Portuguese: the situation in Brazil and Portugal', in McWhorter, John (ed.), *Language Change and Language Contact in Pidgins and Creoles*. Amsterdam/Philadelphia: Benjamins, pp. 235–55.
- Neagoe, Victorela 1984. 'Subdialectul bănăţean', in Rusu (1984), pp. 240–84.
- Nebrija, Antonio de 1492a. *Grammatica Antonii Nebrissensis*. Salamanca.
- 1492b. *Lexicon hoc est Dictionarium ex sermone latino in hispaniensem*. Salamanca.
1495. *Dictionarium ex hispaniensi in latinum sermonem*. Salamanca.
- Nehama, Joseph 1977. *Dictionnaire du judéo-espagnol*. Madrid: CSIC.
- Neumann-Holzschuh, Ingrid 1992. 'Spanisch: Grammatikographie. Gramaticografía', in *LRL* (VI.1), pp. 616–35.
2000. 'Restructuration dans un créole "conservateur": le cas du créole louisianais', in Neumann-Holzschuh and Schneider (2000), pp. 383–407.
- Neumann-Holzschuh, Ingrid and Schneider, Edgar (eds) 2000. *Degrees of Restructuring in Creole Languages*. Amsterdam/Philadelphia: Benjamins.
- Niculescu, Alexandru 1965. *Individualitatea limbii române între limbile romanice. Contribuţii gramaticale*. Bucharest: Editura ştiinţifică.
1978. *Individualitatea limbii române între limbile romanice. Contribuţii socioculturale*. Bucharest: Editura ştiinţifică şi enciclopedică.
1981. *Outline History of the Romanian Language*. Bucharest: Editura Ştiinţifică şi enciclopedică.
1992. 'Le daco-roumain. Romania antiqua, Romania nova et la continuité mobile. Une synthèse', in *Actes du XVIII^e Congrès International de Linguistique et Philologie Romanes I*. Tübingen: Niemeyer, pp. 86–104.

1999. *Individualitatea limbii române între limbile romanice 3. Noi contribuții*, Cluj: Clusium.
2003. *Individualitatea limbii române între limbile romanice. Elemente de istorie culturală*. Cluj: Clusium.
2007. *L'altra romanità. Storia linguistica del romeno tra Oriente e Occidente*. Verona: Fiorini.
- Niederehe, Hans-Josef 1975. *Die Sprachauffassung Alfons des Weisen: Studien zur Sprach- und Wissenschaftsgeschichte*. Tübingen: Niemeyer.
1987. 'Les dictionnaires bilingues français–espagnols jusqu'en 1800', *Histoire Epistémologie Langage* 9:13–26.
1988. 'Les dictionnaires bilingues français–espagnols et espagnols–français au XVIII siècle', *Travaux de Linguistique et de Philologie* 26:33–47.
1994. *Bibliografía cronológica de la lingüística, la gramática y la lexicografía del español. Desde los comienzos hasta el año 1600 (BICRES I)*. Amsterdam/Philadelphia: Benjamins.
1999. *Bibliografía cronológica de la lingüística, la gramática y la lexicografía del español. Desde el año 1601 hasta el año 1700 (BICRES II)*. Amsterdam/Philadelphia: Benjamins.
2001. 'La gramaticografía del siglo XVIII entre tradición y reorientación', in Koerner, Konrad and Niederehe, Hans-Josef (eds), *History of Linguistics in Spain II*. Philadelphia/Amsterdam: Benjamins, pp. 181–94.
2005. *Bibliografía cronológica de la lingüística, la gramática y la lexicografía del español. Desde el año 1701 hasta el año 1800 (BICRES III)*. Amsterdam/Philadelphia: Benjamins.
- Nielsen, Hans and Schosler, Lene (eds) 1996. *The Origins and Development of Emigrant Languages*. Odense: Odense University Press.
- Niño-Murcia, Mercedes 1995. 'The gerund in the Spanish of the North Andean Region', in Silva-Corvalán, Carmen (ed.), *Spanish in Four Continents. Studies in Language Contact and Bilingualism*. Washington, DC: Georgetown University Press, pp. 83–100.
- Norberg, Dag 1966. 'A quelle époque a-t-on cessé de parler latin en Gaule?', *Annales. Economies, Sociétés, Civilisations* 21:346–56.
1968. *Manuel pratique de latin médiéval*. Paris: Picard.
1999. *Manuale di latino medievale*. Cava dei Tirreni: Avagliano.
- Norden, Eduard 1898. *Die antike Kunstprosa. Vom VI Jahrhundert vor Chr. bis in die Zeit der Renaissance*. Leipzig: Teubner.
- Norton, Frederick John 1966. *Printing in Spain, 1501–1520, with a Note on the Early Editions of the Celestina*. Cambridge/New York: Cambridge University Press.
- Novati, Francesco 1926. *Storia letteraria d'Italia: le origini*. Milan: Vallardi.
- NSIRB 2001 = [Republic Of Bulgaria – National Statistical Institute] Република България Национален Статистически Институт, *Преброяване 2001. Население към 01.03.2001 г. По области и етническа група*. Available at www.nsi.bg/Census/Ethnos.htm (last accessed 17 August 2007).
- Oesterreicher, Wulf 1993. 'Verschriftung und Verschriftlichung im Kontext medialer und konzeptioneller Schriftlichkeit', in Schäfer, Ursula (ed.), *Schriftlichkeit im frühen Mittelalter*. Tübingen: Narr, pp. 267–92.
1995. 'L'oral dans l'écrit. Essai d'une typologie à partir des sources du latin vulgaire', in Callebat, Louis (ed.), *Latin vulgaire–latin tardif IV*. Hildesheim: Olms-Weidmann, pp. 145–57.
1997. 'Zur Fundierung von Diskurstraditionen', in Frank, Barbara, Haye, Thomas and Tophinke, Doris (eds), *Gattungen mittelalterlicher Schriftlichkeit*. Tübingen: Narr, pp. 19–41.
2000. 'L'étude des langues romanes', in Auroux (2000), pp. 9–22.

- 2001a. 'La "recontextualización" de los géneros medievales como tarea hermenéutica', in Jacob and Kabatek (2001), pp. 199–232.
- 2001b. 'Sprachwandel, Varietätenwandel, Sprachgeschichte', in Bäumel, Franz, Schaefer, Ursula and Spiedelmann, Edda (eds), *Folgen und Folgen von Schriftlichkeit und Mündlichkeit*. Tübingen: Narr, pp. 217–48.
2005. 'Textos entre inmediatez y distancia comunicativas. El problema de lo hablado escrito en el Siglo de Oro', in Cano Aguilar (2005), pp. 729–69.
- Oliva, Francesco 1989. *Grammatica della lingua napoletana* (ed. Enrico Malato). Naples: Rome/Naples: Bulzoni.
- Olivar, Alexandre 1991. *La predicación cristiana antigua*. Barcelona: Herder.
- OnCat. = Coromines, Juan 1989–97. *Onomasticon Cataloniae*. Barcelona: Curial.
- Ong, Walter 1982. *Orality and Literacy. The Technologizing of the Word*. London: Methuen.
- Oppo, Anna (ed.) 2007. *Le lingue dei sardi. Una ricerca sociolinguistica*. Cagliari: Regione Autonoma della Sardegna / Università degli Studi di Cagliari.
- Orbanić, Srda 1995. 'Status attuale delle comunità istro-romene', *Annales* (Koper) 6:57–64.
- Otheguy, Ricardo and Zentella, Ana Celia 2007. 'Apuntes preliminares sobre el contacto lingüístico y dialectal en el uso pronominal del español en Nueva York', in Potowski and Cameron (2007), pp. 275–95.
- Overbeke, Maurits van 1968. 'La description phonétique et phonologique d'une situation bilingue', *La Linguistique* 4:93–109.
- Paccagnella, Ivano 1994. 'Uso letterario dei dialetti', in Serianni and Trifone (1993–94) (III), pp. 495–539.
- Paden William 1998. *An Introduction to Old Occitan*. New York: Modern Language Association.
- Padley, George Arthur 1976. *Grammatical Theory in Western Europe 1500–1700. The Latin Tradition*. Cambridge: Cambridge University Press.
1985. *Grammatical Theory in Western Europe 1500–1700: Trends in Vernacular Grammar I*. Cambridge: Cambridge University Press.
1988. *Grammatical Theory in Western Europe 1500–1700. Trends in Vernacular Grammar II*. Cambridge: Cambridge University Press.
- Papen, Robert 2004. 'Sur quelques aspects structuraux du français des Métis de l'Ouest canadien', in Coveney et al. (2004), pp. 105–29.
- Pariente, Jean-Claude 1985. *L'Analyse du langage à Port-Royal*. Paris: Éditions de Minuit.
- Parisot, Jean-Cristophe 1998. 'Pratique linguistique et représentations politiques des élus locaux', in Eloy, Jean-Michel (ed.), *Evaluer la vitalité: variétés d'oil et autres langues. Actes du colloque 'Evaluer la vitalité des variétés régionales du domaine d'oil'*. Amiens: Université de Picardie–Jules Verne, pp. 245–60.
- Parisot, Jean-Cristophe, Moro, Ahmed and Eloy, Jean-Michel 1998. 'L'enquête quantitative de 1994 en Picardie. Notes de lecture', in Eloy, Jean-Michel (ed.), *Evaluer la vitalité: variétés d'oil et autres langues. Actes du colloque 'Evaluer la vitalité des variétés régionales du domaine d'oil'*. Amiens: Université de Picardie–Jules Verne, pp. 115–29.
- Parodi, Severina (ed.) 1974. *Gli atti del primo Vocabolario*. Florence: Sansoni.
- Patota, Giuseppe 1993. 'I percorsi grammaticali', in Serianni and Trifone (1993–94) I, pp. 93–137.
- Paul, Hermann 1880. *Principien der Sprachgeschichte*. Halle: Niemeyer.
1888. *Principles of the History of Language*. London: Swan Sonnenschein, Lowrey & Co.

- Pei, Mario 1932. *The Language of the Eighth Century Texts in Northern France. A Study of the Original Documents in the Collection of Tardif and other Sources*. New York: Harper.
- Peitgen, Heinz-Otto and Richter, Peter 1986. *The Beauty of Fractals. Images of Complex Dynamical Systems*. Berlin: Springer.
- Pellegrini, Giovan Battista 1971. 'Delle varie accezioni ed estensione di *ladino*', in *Atti del convegno sui dialetti del Trentino (17–18–19 ottobre 1969)*. Trento: Centro culturale Fratelli Bronzetti, pp. 1–20.
1972. *Gli arabismi nelle lingue neolatine con speciale riguardo all'Italia*. Brescia: Paideia.
1977. *Carta dei dialetti d'Italia*. Pisa: Pacini.
- Penny, Ralph 1996. 'Judeo-Spanish varieties before and after the expulsion', *Donaire* 6:54–58.
2000. *Variation and Change in Spanish*. Cambridge: Cambridge University Press.
2002. *A History of the Spanish Language*. Cambridge: Cambridge University Press.
- Percoto, Caterina 1988. *Scritti friulani, tradotti, commentati e restaurati da Amedeo Giacomini*. Udine: Società Filologica Friulana.
- Pérez González, Maurilio (ed.) 1993. *Actas I congreso nacional de latín medieval*. León: Universidad de León.
- Périn, Patrick 1992. 'La part du Haut Moyen Age dans la genèse des terroirs de la France médiévale', in Parisse, Michel and Barral i Altet, Xavier (eds), *Le Roi de France et son royaume autour de l'an Mil*. Paris: Picard, pp. 225–34.
- Perl, Matthias 1994. *Portugiesisch und Crioulo in Afrika. Geschichte, Grammatik, Lexik, Sprachentwicklung*. Bochum: Brockmeyer.
- Perl, Matthias and Schwegler, Armin (eds) 1998. *América Negra. Panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas*. Frankfurt am Main: Vervuert.
- Pertusi, Agostino 1962. 'ΕΡΩΤΗΜΑΤΑ. Per la storia e le fonti delle prime grammatiche greche a stampa', *Italia Medioevale e Umanistica* 5:321–51.
- Petersen, Leiva (ed.) 1970. *Prosopographia Imperii Romani Saec. I. II. III.*, pars V, fasc. 1. Berlin: de Gruyter.
- Petrovici, Emil and Neiescu, Petru 1964. 'Persistența insulelor lingvistice. Constatări făcute cu prilejul unor noi anchete dialectale la istroromâni, meglenoromâni și aromâni', *Cercetări de lingvistică* 9:187–214.
- Petrucchi, Armando and Romeo, Carlo 1992. *Scriptores in urbibus*. Bologna: Il Mulino.
- Petrucchi, Livio 2010. *Alle origini dell'epigrafia volgare. Iscrizioni italiane e romanze fino al 1275*. Pisa: Edizioni Plus.
- Petrucchi, Peter 1999. *Slavic Features in the History of Rumanian*. Munich: Lincom.
- Pfänder, Stefan 2000. *Aspekt und Tempus im Frankokreol*. Tübingen: Narr.
- Pfister, Max 1970. 'Die Anfänge der altprovenzalischen Schriftsprache', *ZRPh* 86:305–23.
1973. 'Die sprachliche Bedeutung von Paris und der Ile-de-France vor dem 13. Jahrhundert', *VR* 32:217–53.
1976. 'La langue de Guilhem IX, comte de Poitiers', *Cahiers de Civilisation Médiévale* 19:91–113.
1978. 'Le superstrat germanique dans les langues romanes', in *Atti del XIV Congresso internazionale di linguistica e filologia romanza*. I. Naples/Amsterdam: Macchiaroli/Benjamins, pp. 49–97.
- Philipp-Sattel, Sabine 1996. *Parlar bellament en vulgar. Die Anfänge der katalanischen Schriftkultur im Mittelalter*. Tübingen: Narr.
- Philippide, Alexandru 1925. *Originea românilor*, I. Iași: Viața românească.

- Pi de Cabanyes, Oriol 1979. *La Renaixença*. Barcelona: Dopesa.
- Picard, Christophe 2000. *Le Portugal musulman (VIII–XIIIe siècle). L'Occident d'al-Andalus sous domination islamique*. Paris: Maisonneuve et Larose.
- Picco, Linda 2001. *Ricerche su la condizion sociolenghistiche dal furlan / Ricerca sulla condizione sociolinguistica del friulano*. Udine: Forum.
- Piètri, Charles (ed.) 1993. *Evêques, moines et empereurs, 610–1054*. Paris: Desclée.
- (ed.) 1998. *Les Deux Eglises (430–610)*. Paris: Desclée.
- Piètri, Charles, Pietri, Luce and Biane, Jacques (eds) 1995. *Naissance d'une chrétienté 250–430*. Paris: Desclée De Brouwer.
- Pinker, Stephen 1994. *The Language Instinct*. New York: Morrow.
- Pinkster, Harm 1991. 'Evidence for SVO in Latin?', in Wright (1991), pp. 69–82.
1995. 'Word order in the Late Latin, Gesta conlationis carthaginiensis', in Callebat, Louis (ed.), *Latin vulgaire–latin tardif IV*. Hildesheim: Olms-Weidmann, pp. 549–60.
- Pitz, Martina 1997. *Siedlungsnamen auf -villare (-weiler, -villers) zwischen Mosel, Hunsrück und Vogesen*. Saarbrücken: Saarbrücker Druckerei und Verlag.
2000. 'Le superstrat francique dans le Nord-Est de la Gaule. Vers une nouvelle approche philologique et toponymique'. *Nouvelle Revue d'Onomastique* 35/36:69–85.
- Ploog, Katja 2010. 'L'ambiguïté constructionnelle dans la dynamique langagière', in Drescher and Neumann-Holzschuh (2010a), pp. 81–94.
- Poggi Salani, Teresa 1988. 'Italienisch: Grammatikographie. Storia delle grammatiche', in *LRL* (IV), pp. 774–86.
- Poghirc, Cicerone 1969. 'Influența autohtonă', in Coteanu, Ion (ed.), *Istoria limbii române II*, Bucharest: Editura Academiei, pp. 313–65.
- Pohl, Walter (ed.), 1999. *Kingdoms of the Empire. The Integration of Barbarians in Late Antiquity*. Leiden: Brill.
- Pons Rodríguez, Lola 2010. *La lengua de ayer. Manual práctico de historia del español*. Madrid: Arco.
- Pontes, Eunice 1987. *O tópico no português do Brasil*. Campinas: Editora Pontes.
- POPIS 2002 = Statistični Urad Republike Slovenije / Statistical Office of the Republic of Slovenia, *Popis 2002 / Census 2002 results*, in *Statistične informacije / Rapid Reports*, 23/2003. See also www.stat.si/popis2002/en/.
- Poplack, Shana 1987. 'Contrasting patterns of codeswitching in two communities', in Wande, Erling, Anward, Jan, Nordberg, Bengt, Steenslad, Lars and Thelanders, Mats (eds), *Aspects of Multilingualism*. Uppsala: Borgströms, pp. 51–77.
- Posner, Rebecca 1985. 'Creolization as typological change: some examples from Romance syntax', *Diachronica* 2:167–88.
1996. *The Romance Languages*. Cambridge: Cambridge University Press.
1997. *Linguistic Change in French*. Oxford: Clarendon Press.
- Posner, Rebecca and Green, John (eds) 1980–93. *Trends in Romance Linguistics and Philology*, 5 vols. Berlin/New York: Mouton de Gruyter.
- Potowski, Kim and Cameron, Richard (eds) 2007. *Spanish in Contact. Policy, Social and Linguistic Inquiries*. Amsterdam: Benjamins.
- Pottier, Bernard (ed.) 1983. *América Latina en sus lenguas indígenas*. Caracas: Monte Ávila.
- Pountain, Christopher 2000. 'Capitalization', in Smith, John Charles and Bentley, Delia (eds), *Historical Linguistics 1995, Volume 1: General Issues and non-Germanic Languages*. Amsterdam: Benjamins, pp. 295–309.

2003. *Exploring the Spanish Language*. London: Arnold.
2006. 'Syntactic borrowing as a function of register', in Lepschy and Tosi (2006), pp. 99–111.
2007. "'Spanglish': myths and realities', *Vida Hispánica* 35:14–18.
- Pozzi, Mario (ed.) 1978. *Trattatisti del Cinquecento*. Milan: Ricciardi.
- (ed.) 1988. *Discussioni linguistiche del Cinquecento*. Turin: UTET.
- Prince, Ellen 1988. 'On pragmatic change: the borrowing of discourse functions'. *Journal of Pragmatics* 12:508–18.
- Principato, Aurelio 2000. *Breve storia della lingua francese. Dal Cinquecento ai giorni nostri*. Rome: Carocci.
- Prosdocimi, Aldo (ed.) 1978. *Lingue e dialetti dell'Italia antica*. Rome: Biblioteca di Storia Patria.
- Prüßmann-Zemper, Helga 1990. 'Französisch: Varietätenlinguistik des Französischen', in *LRL* (V.1), pp. 830–43.
- Puppo, Mario 1966. *Discussioni linguistiche del Settecento*. Turin: UTET.
1975. *Critica e linguistica del Settecento*. Verona: Fiorini.
- Puşcariu, Sextil 1926. *Studii istroromâne*. Bucharest: Cultura naţională.
1937. 'Zur Nachstellung des rumänischen Artikels', *ZRPh* 57:240–74.
1940. *Limba română I. Privire generală*, Bucharest: Fundația pentru literatură și artă „Regele Carol II”.
- Putte, Florimon van 1999. *Dede pikiña ku su bisina. Papiament-Nederlands en de onverwerkt verleden tijd*. Zutphen: Walburg.
- Quilis Merín, Mercedes 1999. *Orígenes históricos de la lengua española*. Valencia: Universitat de València.
- Quondam, Amedeo 1978. 'Nascita della grammatica, appunti e materiali per una descrizione', *Quaderni storici* 38:555–92.
- Raible, Wolfgang 1993. 'Die Anfänge der italienischen Schriftkultur', *RF* 105:231–55.
1996. 'Relatinisierungstendenzen', *LRL* (II.1), pp. 120–34.
- Ramajo Caño, Antonio 1987. *Las gramáticas de la lengua castellana desde Nebrija a Correas*. Salamanca: Universidad de Salamanca.
- Raposo, Eduardo 1986. 'On the null object in European Portuguese', in Jaeggli, Osvaldo and Silva-Corvalán, Carmen (eds), *Studies in Romance Linguistics*. Dordrecht; Foris, pp. 373–90.
- Rastier, François 2001a. *Sémantique et recherches cognitives*. Paris: PUF.
- 2001b. *Arts et sciences du texte*. Paris: PUF.
- Ravier, Xavier 2005. 'Registres linguistiques et types d'actes', in Ravier, Xavier and Cursente, Benoît (eds), *Le Cartulaire de Bigorre (XIe–XIIIe s.)*. Paris: Editions du Comité des travaux historiques et scientifiques, pp. liii–lxvii.
- Raynaud, Gaston 1876. *Etude sur le dialecte picard dans le Ponthieu d'après les chartes des XIIIe et XIVe siècles 1254–1333*. Paris: Vieweg.
- Real Academia Española 2009. *Nueva gramática de la lengua española. Morfología / Sintaxis I, Sintaxis II*. Madrid: Espasa Libros.
- Récatas, Basile 1934. *L'Etat actuel du bilinguisme chez les Macédo-romains du Pinde et le rôle de la femme dans le langage*. Paris: Droz.
- Reichenkron, Günter 1963. 'Der Typus des Balkansprachen', *Zeitschrift für Balkanologie* 1:91–122.

1965. *Historische Latein-Altromanische Grammatik, I Teil. Das sogenannte Vulgärlatein und das Wesen der Romanisierung*. Wiesbaden: Harassowitz.
- Reinecke, John 1938. 'Trade jargons and creole dialects as marginal languages', *Social Forces* 17:107–18.
- Remacle, Louis 1948. *Le Problème de l'ancien wallon*. Liège: Faculté de Philosophie et de Lettres.
1992. *La Différenciation dialectale en Belgique romane avant 1600*. Geneva: Droz.
- Renzi, Lorenzo 1976. *Introduzione alla filologia romanza*. Bologna: Il Mulino.
1981. *La politica linguistica della Rivoluzione francese. Studio sulle origini e la natura del giacobinismo linguistico*. Naples: Liguori.
1985. *Nuova introduzione alla filologia romanza*. Bologna: Il Mulino.
1998. *ItalAnt: per una Grammatica dell'Italiano Antico*. Padua: Centrostamp.
- Renzi, Lorenzo and Andreose, Alvise 2006. *Manuale di linguistica e filologia romanza*. Bologna: Il Mulino.
- Reuillon-Blanquet, Madeleine 1994. 'Les grammaires des Dames en France et l'apprentissage des langues à la fin du XVIIIe siècle', *Histoire Epistémologie Langage* 16:55–76.
1995. 'Grammaires des Dames, ou comment rendre "instructive et amusante" une étude généralement considérée comme rébarbative', in *Langages de la Révolution (1770–1815), Actes du 4ème Colloque international de lexicologie politique (1991)*. Paris: Publications de l'INALF, Klincksieck, pp. 163–75.
- REW = Meyer-Lübke, Wilhelm 1935 (3rd edn). *Romanisches Etymologisches Wörterbuch*. Heidelberg: Winter.
- RF = *Romanische Forschungen*.
- Rézeau, Pierre 1987. 'Le français du Québec à travers la presse écrite', in Niederehe, Hans-Josef and Wolf, Lothar (eds), *Français du Canada – Français de France*. Tübingen: Niemeyer, pp. 201–75.
- Richardson, Brian (ed.) 1984. *Trattati sull'ortografia del volgare 1524–1526*. Exeter: University of Exeter.
1987. 'Gli italiani e il toscano parlato nel Cinquecento', *LN* 48:97–107.
1994. *Print Culture in Renaissance Italy. The Editor and the Vernacular Text, 1470–1600*. Cambridge: Cambridge University Press.
1999. *Printing, Writers and Readers in Renaissance Italy*. Cambridge: Cambridge University Press.
2002. *The Italian of Renaissance Elites in Italy and Europe*, in Lepschy and Tosi (2002), pp. 5–23.
2009. *Manuscript Culture in Renaissance Italy*. Cambridge: Cambridge University Press.
- Riché, Pierre 1973. *Education et culture en Occident barbare, VIe–VIIIe siècles*. Paris: Seuil.
1993. *La Christianisation des pays entre Loire et Rhin, IVe–VIIe siècles*. Paris: Cerf.
- Richter, Elise 1903. *Zur Entwicklung der Romanischen Wortstellung aus der Lateinischen*. Halle: Niemeyer.
1934. *Beiträge zur Geschichte der romanismen*. I. Halle: Niemeyer.
- Richter, Michael 1976. 'Kommunikationsprobleme im lateinischen Mittelalter', *Historische Zeitschrift* 222:43–80.
1983. 'A quelle époque a-t-on cessé de parler latin en Gaule? A propos d'une question mal posée', *Annales. Economies, Sociétés, Civilisations* 38:439–48.
1985. 'Towards a methodology of historical sociolinguistics', *Folia Linguistica Historica* 6:41–61.

- 1994a. *Oral Traditions in the Early Middle Ages*. Turnhout: Brepols.
- 1994b. *The Formation of the Medieval West. Studies in the Oral Culture of the Barbarians*. Dublin: Four Courts Press.
2005. 'Oral communication', in *Settimana di Studio del Centro Italiano di Studi sull'Alto Medioevo LII. Comunicare e significare nell'alto medioevo*. Spoleto: CISAM, pp. 447–64.
- Rickard, Peter 1989². *A History of the French Language*. London/Boston/Sydney/Wellington: Unwin Hyman.
- Rickford, John 1987. *Dimensions of a Creole Continuum*. Stanford: Stanford University Press.
- Rico, Alberto and Solà, Joan 1991. 'Katalanisch: Grammatikographie und Lexikographie. b) Lexicografia', in *LRL* (V.2), pp. 281–310.
- Riiho, Timo 1994. 'Portugiesische: Interne Sprachgeschichte und Entwicklungslehre', in *LRL* (VI.2), pp. 498–511.
- Ripert, Émile 1918³. *La Renaissance provençale (1800–1860)*. Paris: Champion.
1948. *Le Félibrige*. Paris: Colin.
- Rivarola, José Luis 1990. *La formación lingüística de Hispanoamérica*. Lima: Pontificia Universidad Católica del Perú.
2001. *El español de América en su historia*. Valladolid: Universidad de Valladolid.
2005. 'La difusión del español en el Nuevo Mundo', in Cano Aguilar (2005), pp. 799–823.
- RLiR* = *Revue de linguistique romane*.
- Roberts, Ian 1993. *Verbs and Diachronic Syntax. A Comparative History of English and French*. Dordrecht: Kluwer.
- Roberts, Ian and Kato, Mary (eds) 1996. *Português brasileiro. Uma viagem diacrônica. Homenagem a Fernando Tarallo*. Campinas: Editora da UNICAMP.
- Robillard, Didier de and Beniamino, Michel (eds) 1993. *Le Français dans l'espace francophone*. 1. Paris: Champion.
- Rogge, Waltraud and Beinke, Christiane 1991. 'Katalanisch: Sprachnormierung und Standardsprache. Norme et standard', in *LRL* (V.2), pp. 192–218.
- Rohlf, Gerhard 1924. *Griechen und Romanen in Unteritalien*. Geneva: Olschki.
1937. *La struttura linguistica dell'Italia*. Lepizig: Heinrich Keller.
1954. *Die lexikalische Differenzierung der romanischen Sprachen*. Munich: Beck.
1966. *Grammatica storica della lingua italiana e dei suoi dialetti. Fonetica*. Turin: Einaudi.
1968. *Grammatica storica della lingua italiana e dei suoi dialetti. Morfologia*. Turin: Einaudi.
1971. *Romanische Sprachgeschichte. Geschichte und Grundlagen, Aspekte und Probleme, mit dem Versuch eines Sprachatlas der romanischen Sprachen*. Munich: Beck.
- 1972a. 'Su alcuni calchi sintattici dal greco nell'Italia meridionale', in *Studi e ricerche su lingua e dialetti d'Italia*. Florence: Sansoni, pp. 306–17.
- 1972b. 'La perdita dell'infinito nelle lingue balcaniche e nell'Italia meridionale', in *Studi e ricerche su lingua e dialetti d'Italia*. Florence: Sansoni, pp. 318–32.
- 1972c. 'La congiunzione *mi* (in sostituzione dell'infinito) in Sicilia', in *Studi e ricerche su lingua e dialetti d'Italia*. Florence: Sansoni, pp. 333–38.
1975. *Rätoromanisch. Die Sonderstellung des Rätoromanischen zwischen Italienisch und Französisch: eine kulturgeschichtliche und linguistische Einführung*. Munich: Beck.
1977. *Grammatica storica dei dialetti italogreci (Calabria, Salento)*. Munich: Beck.
1979. 'Romanische Lehnübersetzungen aus germanischer Grundlage', in Löffler, Manfred, Vernay, Henri and Lothar, Wolf (eds), *Festschrift für Kurt Baldinger zum 60. Geburtstag II*. Tübingen: Niemeyer, pp. 805–812.

1982. 'Germanische Inversion in der Romania?', in Stimm, Helmut, Heinz, Sieglinde and Wandruszka, Ulrich (eds), *Fakten und Theorien. Beiträge zur romanischen und allgemeinen Sprachwissenschaft. Festschrift für Helmut Stimm zum 65. Geburtstag*. Tübingen: Narr, pp. 241–44.
1986. *Panorama delle lingue neolatine. Piccolo atlante linguistico panromanzo*. Tübingen: Narr.
- Röllig, Wolfgang 1980. 'Das Punische im römischen Reich', in Bauchhenss, Gerhard, Neumann, Günter and Untermann, Jürgen (eds), *Die Sprachen im römischen Reich der Kaiserzeit*. Cologne: Rheinland, pp. 285–99.
- Romaine, Suzanne 1982. *Socio-Historical Linguistics. Its Status and Methodology*. Cambridge: Cambridge University Press.
- Rona, José Pedro 1959. *El dialecto 'fronterizo' del Norte del Uruguay*. Montevideo: Universidad de la República.
- Rosetti, Alexandru 1938–41. *Istoria limbii române*. Bucharest: Fundația pentru literatură și artă Regele Carol II.
1946. *Istoria limbii române. VI. Din secolul al XIII-lea până în secolul al XVII-lea*. Bucharest: Fundația regală pentru literatură și artă.
1957. 'Despre genul neutru și genul personal în limba română', *SCL* 8:407–13.
1968. *Istoria limbii române de la origini până în secolul al XVII-lea*, București; Editura pentru literatură.
1973. *Brève histoire de la langue roumaine des origines à nos jours*. The Hague/Paris: Mouton.
- Rosetti, Alexandru and Cazacu, Boris 1961. *Istoria limbii române literare. I. De la origini până la începutul secolului al XIX-lea*. Bucharest: Editura Științifică.
- Rosetti, Alexandru, Cazacu, Boris and Onu, Liviu 1971. *Istoria limbii române literare. De la origini până la începutul secolului al XIX-lea*. Bucharest: Minerva.
- Rossi, Mario 1999. *L'Intonation. Le Système du français. Description et modélisation*. Gap: Ophrys.
- Rottet, Kevin 1998. 'Clause subordination structures in language decline', *Journal of French Language Studies* 8:63–95.
2005. 'Attestation et disparition du type *j'avons* en français cadien', in Brasseur and Falkert (2005), pp. 213–27.
- Rozzo, Ugo 2000. 'Il Nuovo Testamento a stampa nei secoli XV e XVI', in D'Aiuto, Francesco, Morello, Giovanni and Piazzoni, Ambrogio (eds), *Vangeli dei popoli. Le parole e l'immagine del Cristo nelle culture e nella storia*. Vatican City: Biblioteca Apostolica Vaticana, pp. 107–14.
- RPh = Romance Philology.
- Rubin, Joan 1968. *National Bilingualism in Paraguay*. The Hague/Paris: Mouton.
- Rubio, Gonzalo 2009. 'Semitic influence in the history of Latin syntax', in Baldi, Philip and Cuzzolin, Pierluigi (eds), *New Perspectives on Historical Latin Syntax, volume 1: Syntax of the Sentence*. Berlin/New York: Mouton de Gruyter, pp. 195–239.
- Ruffino, Giovanni (ed.) 1992. *Atlanti linguistici italiani e romanzi. Esperienze a confronto, Atti del Congresso internazionale Palermo, 3–7 ottobre 1990*. Palermo: Centro di Studi filologici e linguistici siciliani.
1998. 'Tavola rotonda. Principi e metodi della geografia linguistica. Conservazione, rinnovamento o rilancio?', in Ruffino, Giovanni (ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*. Tübingen: Niemeyer, pp. 797–806.

- Ruscelli, Girolamo 1581. *De' commentarii della lingua italiana*. [. . .] *Libri sette*. Venice: Damian Zenaro.
- Russu, Ion 1959. *Limba traco-dacilor*. Bucharest: Editura Academiei Române.
- Rusu, Valeriu (ed.) 1984. *Tratat de dialectologie românească*. Craiova: Scrisul Românesc.
- Ryan, Robert 1998. 'Des manifestations d'économie formelle et sémantique observées au sein du système adverbial du parler franco-acadien de la Baie Sainte-Marie', in Brasseur (1998a), pp. 93–104.
- Sabatini, Francesco 1964. 'Tra latino tardo e origini romanze'. *SLI* 4:140–59.
1965. 'Esigenze di realismo e dislocazione morfologica in testi preromanzì', *Rivista di cultura classica e medioevale* 7:972–98.
1968. 'Dalla scripta latina rustica alle scriptae romanze', *Studi Medievali* 9:320–58.
1978. 'Lingua parlata, scripta e coscienza linguistica nelle origini romanze', in *Atti del XIV Congresso internazionale di linguistica e filologia romanza*, I. Napoli/Amsterdam: Macchiaroli and Benjamins, pp. 445–53.
1983. 'Prospettive sul parlato nella storia linguistica italiana', in Albano Leoni, F., Gambarà, Daniele, Lo Piparo, Franco and Simone, Raffaele (eds), *Italia linguistica. Idee, Storia, Strutture*. Bologna: Il Mulino, pp. 167–201.
1996. *Italia linguistica delle origini*. Lecce: Argo.
- Șăineanu, Lazăr 1900. *Influența orientală asupra limbii și culturii române*. Bucharest: Soccec.
- Sala, Marius 1988. *El problema de las lenguas en contacto*. México: Centro de Lingüística Hispánica, Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México.
- 1998a. *Lenguas en contacto*. Madrid: Gredos.
- 1998b. 'Les langues judéo-romanes', in *LRL* (VII), pp. 372–95.
- Salazar, Michelle 2007. 'Está muy diferente a como era antes: Ser and Estar + adjective in New Mexico Spanish', in Potowski and Cameron (2007), pp. 345–55.
- Salonius, Aarne 1920. *Vitae patrum*. Lund: Gleerup.
- Salvi, Giampaolo 2004. *La formazione della struttura di frase romanza. Ordine delle parole e clitics dal latino alle lingue romanze antiche*. Tübingen: Niemeyer.
- Sánchez, Juan 2002. *Historia del español de América*. Valencia: Tirant lo Blanc.
- Sandfeld, Kristian 1930. *Linguistique balkanique. Problèmes et résultats*. Paris: Champion.
1938. 'Problèmes d'interférences linguistiques', *Actes du Quatrième Congrès International des Linguistes*. Copenhagen: Munksgaard, 58–61.
- Sanga, Glauco 1995. 'La koinè italiana', in *LRL* (II.2), pp. 81–98.
- Sanga, Glauco and Baggio, Serenella 1994. 'Il volgare nei placiti cassinesi: vecchie questioni e nuove acquisizioni', *Rivista italiana di dialettologia* 18:7–30.
1995. 'Sul volgare in età langobarda', in Banfi, Emanuele (ed.), *Italia settentrionale. Crocevia di idiomi romanzi*. Tübingen: Niemeyer, pp. 247–60.
- Sankoff, David and Poplack, Shana 1981. 'A formal grammar of code-switching', *Papers in Linguistics*, 14:3–46.
- Sanson, Helena 2005. 'Women and vernacular grammars in sixteenth-century Italy: the case of Iparca and Rinaldo Corso's *Fondamenti del parlar Toscano* (1549)', *Letteratura Italiana Antica* 6:391–431.
2007. *Donne, precettistica e lingua nell'Italia del Cinquecento. Un contributo alla storia del pensiero linguistico*. Florence: Accademia della Crusca.

2010. 'Women writers and the *Questione della Lingua* in Ottocento Italy: the cases of Caterina Percoto, La Marchesa Colombi, and Matilde Serao', *Modern Language Review*, 105(4):1028–52.
2011. *Women, Language and Grammar. Italy 1500–1900*. London: The British Academy / Oxford: Oxford University Press.
- Saramandu, Nicolae 1984. 'Aromâna', in Rusu (1984), pp. 423–76.
1988. 'Harta graiurilor aromâne și meglenoromâne din Peninsula Balcanică', *SCL* 39:225–45.
- Sas, Louis 1937. *The Noun Declension System in Merovingian Latin*. Paris: André.
- Saussure, Ferdinand de 1916. *Cours de linguistique générale* (publié par Charles Bally et Albert Séchehay; avec la collaboration d'Albert Riedlinger). Lausanne/Paris: Payot.
- [1916] 1995. *Cours de linguistique générale*. Published by Charles Bally and Albert Séchehay in collaboration with Albert Riedlinger. Critical edition and commentaries by Tullio de Mauro. Paris: Payot & Rivages.
- Scaglione, Aldo (ed.) 1984. *The Emergence of National Languages*. Ravenna: Longo.
- Scagno, Roberto, Tomasella, Paolo and Țucu, Corina (eds) 2008. *Veneti in Romania*. Ravenna: Longo.
- Schäfer-Priess, Barbara 2000. *Die portugiesische Grammatikschreibung von 1540 bis 1822. Entstehungsbedingungen und Kategorisierungsverfahren vor dem Hintergrund der lateinischen, spanischen und französischen Tradition*. Tübingen: Niemeyer.
- Scharlau, Birgit 2004. 'Situierete Konzepte: Über das Schreiben von Sprachwissenschaftsgeschichte', in Haßler, Gerda (ed.), *History of Linguistics in Texts and Concepts. Geschichte der Sprachwissenschaft in Texten und Konzepten*, vol. I. Münster: Nodus, pp. 13–32.
- Schiaffini, Alfredo 1959. 'Problemi del passaggio dal latino a l'italiano (evoluzione, disgregazione, ricostruzione)', in Gerardi Marcuzzo, Giuseppina (ed.), *Studi in onore di Angelo Monteverdi*. Modena: Società tipografica editrice modenese, pp. 691–715.
- Schlieben-Lange, Brigitte 1977. 'L'origine des langues romanes – un cas de créolisation', in Meisel, Jürgen (ed.), *Langues en contact – Pidgins – Creoles – Languages in Contact*. Tübingen: Narr, pp. 81–101.
1983. *Traditionen des Sprechens. Elemente einer pragmatischen Sprachgeschichtsschreibung*. Stuttgart: Kohlhammer.
1991. 'Okzitanisch: Grammatikographie und Lexicographie', in *LRL* (V.2), pp. 105–26.
1996. *Idéologie, révolution et uniformité de la langue*. Liège: Mardaga.
- Schmid, Beatrice 1992. 'Spanisch: Geschichte der Verschriftung', in *LRL* (VI.1), pp. 414–27.
- Schmid, Heinrich 1982. *Richtlinien für die Gestaltung einer gesamtbündnerromanischen Schriftsprache Rumantsch grischun*. Chur: Lia Rumantscha.
- Scholle, Frantz 1869. *Über den Begriff Tochtersprache. Ein Beitrag zur gerechten Beurtheilung des Romanischen, namentlich des Französischen*. Berlin: W. Weber.
- Schuchardt, Hugo 1866. *Der Vokalismus des Vulgärlateins*. Leipzig: Teubner.
1882. 'Kreolische Studien I: Ueber das Negerportugiesische von S. Thomé (Westafrika)', *Sitzungsberichte der philosophisch-historischen Classe der Kaiserlichen Akademie der Wissenschaften* 100:889–917. [Translated with notes by Tjerk Hagemeijer and John Holm, 'On the Creole Portuguese of São Tomé (West Africa)', in Holm and Michaelis (2008), pp. 131–56.]
- 1883a. 'Lucien Adam, Les idiomes négro-aryen et maléo-aryen', *Literaturblatt für germanische und romanische Philologie* 6:236–40.

- 1883b. 'Kreolische Studien II: Ueber das Indoportugiesische von Cochim', *Sitzungsberichte der philosophisch-historischen Classe der Kaiserlichen Akademie der Wissenschaften* 102:799–816.
- 1884a. 'Kreolische Studien IV: Ueber das Malaiospanische der Philippinen', *Sitzungsberichte der philosophisch-historischen Classe der Kaiserlichen Akademie der Wissenschaften* 105:111–50.
- 1884b. *Slawo-deutsches und Slawo-italienisches. Dem Herrn Franz von Miklosich zum 20. November 1883*. Graz: Leuschner & Lubensky.
1885. *Über die Lautgesetze. Gegen die Junggrammatiker*. Berlin: Oppenheim.
1909. 'Die Lingua franca', *ZRPh* 33:441–61.
- Schüle, Ernest 1978. 'Histoire et évolution des parler francoprovençaux d'Italie', in Clivio, Gianrenzo and Gasca Queirazza, Giuliano (eds), *Lingue e dialetti nell'arco alpino occidentale; Atti del Convegno Internazionale di Torino*. Turin: Centro Studi Piemontesi, pp. 127–40.
- Schwegler, Armin. 1998. 'Palenquero', in Perl and Schwegler (1998), pp. 218–91.
1999. 'Monogenesis revisited: the Spanish perspective', in Rickford, John and Romaine, Suzanne (eds), *Creole Genesis, Attitudes and Discourse. Studies Celebrating Charlene J. Sato*. Amsterdam/Philadelphia: Benjamins, pp. 235–62.
- Schwenter, Scott and Silva, Gláucia 2003. 'Anaphoric direct objects in spoken Brazilian Portuguese: semantics and pragmatics', *Revista internacional de lingüística iberoamericana* 1:99–123.
- SCL = *Studii și cercetări lingvistice*.
- Segarra, Mila 1987. *Prólogo a Josep Pau Ballot. Grammatica y apologia de la llengua catalana*. Barcelona: Alta Fulla.
- Seguin, Jean-Pierre 1972. *La Langue française au XVIII^e siècle*. Paris/Brussels/Montreal: Bordas.
- Selig, Maria 1993. 'Parodie et protocole – l'importance de la "citation" pour les premiers documents des langues romanes', in Selig et al. (1993), pp. 91–108.
1995. *Volkssprachliche Schriftlichkeit im Mittelalter – Die Genese der altokzitanischen Schriftsprache*. Freiburg: unpublished thesis.
2001. 'El problema de la tipología de los textos románicos primitivos', in Jacob and Kabatek (2001), pp. 233–48.
2008. 'Koineisierung im Altfranzösischen? Dialektmischung, Verschriftlichung und Überdachung im französischen Mittelalter', in Heinemann, Sabine (ed.), *Sprachwandel und (Dis-)Kontinuität in der Romania*. Tübingen: Niemeyer, pp. 71–85.
- Selig, Maria, Frank, Barbara and Hartmann, Jorg (eds) 1993. *Le Passage à l'écrit des langues romanes*. Tübingen: Niemeyer.
- Sénac, Philippe 2002. *Les Carolingiens et Al-Andalus (VIIIe–IXe s.)*. Paris: Maisonneuve et Larose.
- Séphiha, Haïm Vidal 1972. 'Langues juives, langues calques et langues vivantes', *La Linguistique* 8:59–68.
1986. *Le Judéo-espagnol*. Paris: Editions Entente.
- Serianni, Luca 1989. *Il primo Ottocento*. Bologna: Il Mulino.
1990. *Il secondo Ottocento*. Bologna: Il Mulino.
1999. 'Lingue e dialetti d'Italia nella percezione dei viaggiatori sette-ottocenteschi', in Tatti, Mariasilva (ed.), *Italia e Italie. Immagini tra Rivoluzione e Restaurazione. Atti del Convegno di Studi. Rome, 7–8–9 novembre 1996*. Rome: Bulzoni, pp. 25–51.
- Serianni, Luca and Trifone, Pietro (eds) 1993–94. *Storia della lingua italiana*. Turin: Einaudi.

- Shane, Sanford 1968. *French Phonology and Morphology*. Cambridge: Cambridge University Press.
- Sibille, Jean 2003. 'Francoprovençal', in Cerquiglini (2003), pp. 117–24.
- Siegel, Jeff 1985. 'Koinés and koineization', *Language in Society* 14:357–78.
1993. 'Introduction: controversies in the study of koinés and koineization', *International Journal of the Sociology of Language* 99:5–8.
1997. 'Mixing, leveling, and pidgin/creole development', in Spears and Winford (1997), pp. 111–40.
- Silva Neto, Serafim da 1952. *História da língua portuguesa*. Rio de Janeiro: Livros de Portugal.
1961. *A constituição do português como Língua nacional*. Lisbon: Universidade de Lisboa.
- 1979 (3rd edn). *História da língua portuguesa*. Rio de Janeiro: Livros de Portugal.
- Silva-Corvalán, Carmen 1994. *Language Contact and Change. Spanish in Los Angeles*. Oxford: Oxford University Press.
- Silva-Fuenzalida, Ismael 1952. *Papiamentu Morphology*. North Western University: unpublished dissertation.
- Simone, Raffaele 1998. 'The early modern period', in Lepschy (1998), pp. 149–236.
- Simoni-Aurembou, Marie-Rose 2003. 'Les langues d'oïl', in Cerquiglini (2003), pp. 137–72.
- Skok, Petar 1931. 'Byzance comme centre d'irradiation pour les mots latins des langues balkaniques', *Byzantion* 6:371–78.
1934. 'Zum Balkanlatein IV', *ZRPh* 54:175–215.
- SLI = *Studi linguistici italiani*.
- Smith, John Charles 2002. 'Middle French. When? What? Why?', in Hickey, Raymond (ed.), *Collecting Views on Language Change. A Donation to Roger Lass on his 65th Birthday*. Special issue of *Language Sciences* 24(3–4), pp. 423–45.
- Solà, Joan 1977. *Del català incorrecte al català correcte. Història dels criteris de correcció lingüística*. Barcelona: Edicions 62.
1991. 'Katalanisch: a) Grammatikographie. Gramaticografia', in *LRL* (V.2), pp. 261–81.
- Solalinde, Antonio 1915. 'Intervención de Alfonso X en la redacción de sus obras', *Revista de filología española* 11:283–88.
- Sornicola, Rosanna 1989. 'L'interferenza rivisitata', *Medioevo Romanzo* 14:435–68.
2008. 'Sul problema dei resti di ipse nella Romania', in Lazzaroni, Romano, Banfi, Emanule, Bernini, Giuliano, Chini, Marina and Marotta, Giovanna (eds), *Diachronica et Synchronica. Studi in onore di Anna Giacalone Ramat*. Pisa: ETS, pp. 537–62.
- [to appear]. 'Dislivelli stilistici e diglossia in testi notarili campani (IX–X secolo)'.
- SORS 2003 = Republic of Serbia – Statistical Office of the Republic of Serbia, *Population: 2002 Census of Population, Households and Dwellings*, 3. Religion, Mother Tongue and National or Ethnic Affiliation by Age and Sex, Belgrade.
- Spears, Arthur and Winford, Donald (eds) 1997. *The Structure and Status of Pidgins and Creoles*. Amsterdam/Philadelphia: Benjamins.
- Spevak, Olga 2010. *Constituent Order in Classical Latin Prose*. Amsterdam: Benjamins.
- Spolsky, Bernard and Shohamy, Elana 1999. *The Languages of Israel. Policy, Ideology and Practice*. Clevedon: Multilingual Matters.
- Spongano, Raffaele 1949. *La prosa di Galileo e altri scritti*. Messina/Florence: D'Anna.
- SSCU 2006 = State Statistics Committee of Ukraine, *All-Ukrainian Population Census, National Composition of Population*. Available at www.ukrcensus.gov.ua/eng/results/general/ (last accessed 15 August 2007).

- Stavans, Ilan 2003. *Spanglish. The Making of a New American Language*. New York: Rayo.
- Stefenelli, Arnulf 1995. 'Remarques sur la structure socioculturelle du latin vulgaire proto-roman', in Callebat, Louis (ed.), *Latin vulgaire-latin tardif IV*. Hildesheim: Olms-Weidmann, pp. 35–45.
1998. 'La base lexicale des langues romanes', in Herman (1998a), pp. 53–65.
- Stegagno Picchio, Luciana 1959. 'La questione della lingua in Portogallo', in Stegagno Picchio (ed.), João de Barros, *Diálogo em Louvor da nossa Linguagem. Lettura critica dell'edizione del 1540 con una introduzione su la Questione della lingua*. Modena: Società Tipografica Editrice Modenese, pp. 5–54.
- Stein, Peter 1998. 'Romanische Kreolsprachen I: Begriffsbestimmung und Bezeichnungen', in *LRL* (VII), pp. 610–18.
- Steiner, Roger 1970. *Two Centuries of Spanish and English Bilingual Lexicography (1590–1800)*. La Haye/Paris: Mouton.
- Stewart, Ian 1989. *Does God Play Dice? The Mathematics of Chaos*. Oxford: Blackwell.
- Stimm, Helmut 1977. 'Ein neuer früher Beleg des synthetischen Futur', *Zeitschrift für französische Sprache und Literatur* 5:40–52.
- Stolz, Thomas 1998. 'Romanische Kreolsprachen II: Portugiesische Kreolsprachen', in *LRL* (VII), pp. 618–37.
- Stotz, Peter 1996–2004. *Handbuch zur lateinischen Sprache des Mittelalters*. Munich: Beck.
- Straka, Georges 1990. 'Phonétique et phonématique', in *LRL* (V.1), pp. 1–33.
- Strassi, Silvia 2008. *L'archivio di Claudius Tiberianus da Karanis*. Berlin/New York: de Gruyter.
- Stussi, Alfredo 1980. 'Storia linguistica e storia politica', in Mazzuoli Porru, Giulia (ed.), *Nuovi metodi e problemi nella linguistica storica*. Pisa: Giardini, pp. 115–31.
- Swigart, Leigh 1992. 'Two codes or one? The insiders' view and the description of codeswitching in Dakar', *Journal of Multilingual and Multicultural Development* 13:83–102.
- Swiggers, Pierre 1990. 'Französisch: Grammatikographie. Grammaticographie', in *LRL* (V.1), pp. 843–69.
2001. 'Geschichte der Grammatiken und Sprachlehre romanischer Sprachen', in *LRL* (I.1), pp. 476–532.
2011. 'La (relative) standardisation de l'ancien Occitan appréhendée à travers les premières descriptions grammaticales', in Dessi Schmid et al. (2011), pp. 133–49.
- Swiggers, Pierre and Vanvolsem, Serge 1987. 'Les premières grammaires vernaculaires de l'italien, de l'espagnol et du portugais', *Histoire Epistémologie Langage* 9: 157–81.
- Sylvain, Suzanne 1936. *Le Créole haïtien. Morphologie et syntaxe*. Wetteren: de Meester.
- Szlezák, Edith 2010. 'Aspects morphosyntaxiques des variétés du français canadien parlées au Massachusetts', in Drescher and Neumann-Holzschuh (2010a), pp. 111–21.
- Tagliavini, Carlo 1972⁶. *Le origini delle lingue neolatine. Introduzione alla filologia romanza*. Bologna: Pàtron.
- Talamo, Giuseppe 1960. *La scuola dalla legge Casati alla inchiesta del 1864*. Milan: Giuffrè.
- Tarallo, Fernando 1983. *Relativization Strategies in Brazilian Portuguese*. University of Pennsylvania: unpublished thesis.
1989. 'Inside and outside relative clauses: pronominal redundancy in Portuguese', in Fasold, Ralph and Schiffrin, Deborah (eds), *Language Change and Variation*. Amsterdam: Benjamins, pp. 255–74.
- Tavoni, Mirko 1998. 'Renaissance linguistics. Western Europe', in Lepschy (1998), pp. 1–108.

- Telmon, Tullio 1985. 'Alcune considerazioni sulle parlate di Faeto e Celle alla luce di una recente pubblicazione', *Bollettino dell'Atlante linguistico italiano*, serie III, 7–10:47–51.
1992. *Le minoranze linguistiche in Italia*. Alessandria: dell'Orso.
1994. 'Aspetti sociolinguistici delle eteroglossie in Italia', in Seriani and Trifone (1993–94), pp. 923–50.
- Ternes, Elmar 1998. 'Keltisch und Romanisch / Le celtique et les langues romanes', in *LRL* (VII), pp. 266–91.
- Terracini, Benvenuto 1956. *L'Italia linguistica nell'alto Medioevo*. Turin: Gheroni.
1957. *Pagine e appunti di linguistica storica*. Florence: Le Monnier.
1964. *Saggio di un Atlante linguistico della Sardegna in base ai rilevi di Ugo Pellis*. Turin: Rattero.
- Terry, Arthur 2003. *A Companion to Catalan Literature*. Woodbridge: Tamesis.
- Tesi, Riccardo 2001. *Storia dell'italiano. La formazione della lingua comune dalle origini al Rinascimento*. Bari/Rome: Laterza.
- Teyssier, Paul 1959. *La Langue de Gil Vicente*. Paris: Klincksieck.
1980. *Histoire de la langue portugaise*. Paris: PUF.
1994. 'Portugiesisch: externe Sprachgeschichte. Histoire externe de la langue', in *LRL* (VI.2), pp. 461–72.
- Thomas, Alain 1982. "'OI" en franco-ontarien: étude socio-phonétique', *Information Communication* 3:1–27.
- Thomas, Heinz 1988. 'Der Ursprung des Wortes Theodiscus', *Historische Zeitschrift* 247:295–331.
1990. 'Zur Geschichte von theodiscus und teutonicus im Frankenreich des 9. Jahrhunderts', *Beiträge zur Geschichte der Regnum Francorum* 22:67–95.
- Thomas, Jacques 1953. 'Dialecte et patois: esquisse d'une étude sémantique', *Romanica gandensia* 1:93–117.
- Thomason, Sarah 2008. 'Pidgins/Creoles and historical linguistics', in Kouwenberg and Singler (2008), pp. 242–62.
- Thomason, Sarah and Kaufman, Terrence 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Thun, Harald and Radtke, Edgar (eds) 1996. *Neue Wege der romanischen Geolinguistik. Akten des Symposiums zur empirischen Dialektologie (Heidelberg/Mainz, 21.–24.10.1991)*. Kiel: Westensee.
- Timmermans, Linda 1993. *L'Accès des femmes à la culture (1598–1715). Un débat d'idées de Saint François de Sales à la Marquise de Lambert*. Paris: Champion.
- Țîrcomnicu, Emil 2004. *Meglenoromânii. Destin istoric și cultural*. Bucharest: Editura etnologică.
- Todoran, Romulus 1989. 'Rumänisch: Periodisierung', in *LRL* (III), pp. 460–64.
- Torres Montes, Francisco 2006. 'Política, desarrollo socioeconómico e historia de las lenguas iberorrománicas', in Ernst et al. (2003–08), pp. 1190–202.
- Toso, Fiorenzo 2006. *Lingue d'Europa. La pluralità linguistica dei Paesi europei fra passato e presente*. Milan: Baldini Castoldi Dalai.
- Toubert, Pierre 2004. *L'Europe dans sa première croissance. De Charlemagne à l'an mil*. Paris: Fayard.
- Tovar, Antonio 1964. 'Research report on vulgar Latin and its local variations', *Kratylos* 9:113–34.
- Trabalza, Ciro 1908. *Storia della grammatica italiana*. Milan: Hoepli.

- Trask, Larry 1997. *The History of Basque*. London: Routledge.
2000. *The Dictionary of Historical and Comparative Linguistics*. Edinburgh: Edinburgh University Press.
- Tremblay, Mireille 2005. 'Back en français acadien: archaïsme ou innovation?', in Brasseur and Falkert (2005), pp. 263–73.
- Treptow, Kurt 1997³. *A History of Romania*. Iași: Center for Romanian Studies. Foundation for Romanian Culture and Studies.
- Trifone, Pietro 1990. 'La svolta del romanesco tra Quattro e Cinquecento', in Bolelli, Tristano (ed.), *Studi in memoria di Ernesto Giammarco*. Pisa: Giardini, pp. 425–52.
1992. *Roma e il Lazio*. Turin: UTET.
1993. 'La lingua e la stampa nel Cinquecento', in Serianni and Trifone (1993), vol.I, pp. 425–46.
2006. 'Politica, sviluppo socio-economico e storia della lingua: Italo-romania', in Ernst et al. (2003–8), 1167–77.
- Trovato, Paolo 1984. "'Dialecto" e sinonimi ("idioma", "proprietà", "lingua") nella terminologia linguistica quattro- e cinquecentesca', *Rivista di Letteratura Italiana* 2:205–36.
1994. *Il primo Cinquecento*. Bologna: Il Mulino.
- Trudgill, Peter 1986. *Dialects in Contact*. Oxford: Blackwell.
1991. *Sociolinguistics. An Introduction to Language and Society*. London: Penguin.
2004. *New-Dialect Formation. The Inevitability of Colonial Englishes*. Oxford: Oxford University Press.
- Trumper, John 2003. 'The misunderstood double marking of direct objects and new infinitive strategies in unexpected places: a brief study of Romance variation', in Tortora, Christina (ed.), *The Syntax of Italian Dialects*. Oxford: Oxford University Press, pp. 229–49.
- Tsiapera, Maria and Wheeler, Garon 1993. *The Port-Royal Grammar. Sources and Influences*. Münster: Nodus.
- Tuailon, Gaston 1983. *Le Francoprovençal. Progrès d'une définition*. Saint-Nicolas: Centre d'études francoprovençales René Willien.
2001. *La Littérature en francoprovençal avant 1700*. Grenoble: ELLUG.
- Turculeț, Adrian 1989. 'Rumänisch: Grammatikographie. Grammaticographie', in LRL (III), pp. 481–91.
- Tuten, Donald. 2003. *Koineization in Medieval Spanish*. Berlin/New York: Mouton de Gruyter.
- Uddholm, Alf 1954. *Formulae Marculfi. Etudes sur la langue et le style*. University of Uppsala: unpublished thesis.
- Urciuoli, Bonnie 2008. 'Whose Spanish? The tension between linguistic correctness and cultural identity', in Niño-Murcia, Mercedes and Rothman, Jason (eds), *Bilingualism and Identity. Spanish at the Crossroads with Other Languages*. Amsterdam: Benjamins, pp. 257–77.
- Urișescu, Dorin 1984. 'Graiul din Țara Oașului', in Rusu (1984), pp. 390–98.
- Ursini, Flavia 2002. 'La Dalmazia e l'Istria', in Cortelazzo et al. (2002), pp. 357–74.
- Ursu, Nicolae Alexandru (ed.) 1969. *Dimitrie Eustatievici Brașoveanu. Gramatică românească, 1757, prima gramatică a limbii române*. Bucharest: Editura Științifică.
- Väänänen, Veikko 1959. *Le Latin vulgaire des inscriptions pompéiennes*. Berlin: Akademie-Verlag.
1967. *Introduction au latin vulgaire*. Paris: Klincksieck.

1982. *Introduzione al latino volgare*. Bologna: Pàtron.
- Valdman, Albert (ed.) 1977. *Pidgin and Creole Linguistics*. Bloomington/London: Indiana University Press.
1978. *Le Créole: structure, statut et origine*. Paris: Klincksieck.
- 1998 'Revitalisation du cadien et enseignement du français langue étrangère aux Etats-Unis', in Brasseur (1998a), pp. 279–292.
- Valkhoff, Marius 1931. *Les Mots français d'origine néerlandaise*. Amersfoort: Valkhoff et Cie.
- Vallette, Paul (ed.) 1924. *Apulée, Apologie*. Paris: Les Belles Lettres.
- van Acker, Marieke 2007. *Vt quique rustici et inlitterati hec audierint intellegant. Hagiographie et communication verticale au temps des mérovingiens (VIIe–VIIIe s.)*. Turnhout: Brepols.
- van Coetsem, Frans 1988. *Loan Phonology and the Two Transfer Types in Language Contact*. Dordrecht: Foris.
2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Heidelberg: Winter.
- van Deyck, Rika, Sornicola, Rosanna and Kabatek, Johannes (eds) 2004. *La Variabilité en langue. I. Langue parlée et langue écrite dans le présent et dans le passé*. Ghent: Communication & Cognition.
- (eds) 2005. *La Variabilité en langue. II. Les Quatre variations*. Ghent: Communication & Cognition.
- van Name, Addison 1869–70. 'Contributions to creole grammar', *Transactions of the American Philological Association* 1:123–67.
- van Uytanghe, Marc 1976. 'Le latin des hagiographes mérovingiens et la protohistoire du français', *Romanica Gandensia* 16:5–89.
1977. 'Latin mérovingien, latin carolingien et scripta romana rustica, rupture ou continuité?', in Despy, Georges and Pohl, Jacques (eds), *D'une déposition à un couronnement*, 476–800. Brussels: Editions de l'Université de Bruxelles, pp. 65–88.
1984. 'Histoire du latin, protohistoire des langues romanes et reconstruction', *Francia* 11:579–613.
1987. *Stylisation biblique et condition humaine dans l'hagiographie mérovingienne, 600–750*. Brussels: Académie royale des sciences, des lettres et des beaux-arts de Belgique.
1989. 'Les expressions du type "quod vulgo vocant" dans les textes latins antérieurs au concile de Tours et aux serments de Strasbourg: témoignages lexicologiques et sociolinguistiques de la "langue rustique romaine"?', *ZRPh* 105:28–49.
1991. 'The consciousness of a linguistic dichotomy Latin-Romance in Carolingian Gaul: the contradictions of the sources and of their interpretation', in Wright (1991), pp. 114–29.
1994. 'La Bible et l'instruction des laïcs à l'époque mérovingienne: des témoignages textuels à une approche langagière de la question', *Sacris erudiri* 34:67–123.
1995. 'La langue de la "Vision de Baronte" (678/679). Un spécimen de latin protoroman dans une phase cruciale de la diachronie?', in Callebat, Louis (ed.), *Latin vulgaire–latin tardif IV*. Hildesheim: Olms-Weidmann, pp. 577–609.
2000. 'Rome, Romania, Germania. Recente inzichten in de genese van het Europa der talen', in *Academia Analalecta*. Brussels: Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, pp. 3–24.

2001. 'Biographie, II, C, III,2, Die Hagiographie, Lateinische, col. 1236–1322', in *Reallexicon für Antike und Christentum. Sachwörterbuch zu Auseinandersetzung des Christentums mit der Antiken Welt*. Stuttgart: Hiersemann.
2005. 'Les voies communicationnelles du message hagiographique au haut Moyen-Age', *Settimana di Studio del Centro Italiano di Studi sull'Alto Medioevo LII. Comunicare e significare nell'alto medioevo*. Spoleto: CISAM, pp. 685–731.
2008. 'Quelques observations sur la communication linguistique dans la Romania du IXe siècle', in von Moos (2008), pp. 317–37.
- Vandeloise, Claude 1993. *La Couleur des prépositions (Langages 110)*. Paris: Larousse.
- Vanelli, Laura 2005. 'Le varietà friulane occidentali: tra conservazione e innovazione', in Benincà, Paola and Vanelli, Laura (eds), *Linguistica friulana*. Padua: Unipress, pp. 381–402.
- Varvaro, Alberto 1968. *Storia problemi e metodi della linguistica romanza*. Naples: Liguori.
1979. 'Capitoli per la storia linguistica dell'Italia meridionale e della Sicilia. I. Gli esiti di -ND-, -MB-', *Medioevo romanzo* 6:189–206.
1983. 'Sulla nozione di area isolata: il caso della Lucania', in *La parola nel tempo*. Bologna: Il Mulino, pp. 127–44.
2004. *Identità linguistiche e letterarie nell'Europa romanza*. Rome: Salerno Ed.
2005. 'La latinizzazione delle province come processo di lunga durata', in Kiss, Sandor, Mondin, Luca and Salvi, Giampaolo (eds), *Latin et langues romanes. Études de linguistique offertes à József Herman*. Tübingen: Niemeyer, pp. 115–33.
2007. 'Goscelin de Saint Bertin e la più antica descrizione dell'area linguistica galloromanza', *Medioevo romanzo* 31:164–67.
- Velázquez, Isabel 2003. *Latine dicitur. Vulgo vocant. Aspectos de la lengua escrita y hablada en las obras gramaticales de Isidoro de Sevilla*. Logroño: Fundación San Miguel de la Cogolla.
- Vendryès, Joseph 1925. *Language. A Linguistic Introduction to History*. London: Kegan Paul, Trench and Trubner.
1968. *Le Langage. Introduction linguistique à l'histoire*. Paris: Michel.
- Veny, Joan 1985. *Introducció a la dialectologia catalana*. Barcelona: Enciclopedia catalana.
- Verdo, Rémy 2010. 'Reconfiguration du latin mérovingien sous les carolingiens: étude sociolinguistique des diplômes royaux et des réécritures hagiographiques (VIIe – IXe siècle)', *Positions des thèses soutenues par les élèves pour obtenir le diplôme d'archiviste paleographe*. Paris: École des Chartes, pp. 257–65.
- Versteegh, Kees 2008. 'Non-Indo-European pidgins and creoles', in Kouwenberg and Singler (2008), pp. 158–86.
- Videsott, Paul 2009. *Padania scrittologica. Analisi scrittologiche e scrittometriche di testi in italiano settentrionale antico dalle origini al 1525*. Tübingen: Niemeyer.
- Vidos, Benedek 1965. *Prestito, espansione e migrazione dei termini tecnici nelle lingue romanze e non romanze. Problemi, metodo e risultati*. Florence: Olschki.
- Vieliard, Jeanne 1927. *Le Latin des diplômes royaux et chartes privées de l'époque mérovingienne*. Paris: Champion.
- Vigo, Pasquale 1971. *Istruzione e sviluppo economico in Italia nel secolo XIX*. Turin: Ilte.
- Vincent, Nigel 1996. 'Appunti sulla sintassi dell'infinito coniugato in un testo napoletano del Trecento', in Benincà, Paola, Cinque, Guglielmo, de Mauro, Tullio and Vincent, Nigel (eds), *Italiano e dialetti nel tempo. Saggi di grammatica per Giulio C. Lepsky*. Rome: Bulzoni, pp. 389–409.

2001. 'The decline and fall of the Latin language or does linguistics matter to historians?', Lecture given at the Centre for Late Antiquity, University of Manchester, seminar on Language and Identity in Early Medieval Europe.
- Vinet, Marie-Thérèse 1984. 'La syntaxe du québécois et les emprunts à l'anglais', *Revue de l'Association québécoise de linguistique* 2:221–42.
- Vitale, Maurizio 1984 (3rd edn). *La questione della lingua*. Palermo: Palumbo.
1986. *L'oro nella lingua. Contributi per una storia del tradizionalismo e del purismo italiano*. Milan: Ricciardi.
- Völker, Harald 2001. 'Die Skriptaforchung als eine Philologie der Varietäten. Zur Negation mit *ne* ... *nient* in den altfranzösischen Urkunden der Grafen von Luxemburg 1237–1281', in Gärtner, Holtus, Rapp and Völker (2001), pp. 75–104.
2003. *Skripta und Variation. Untersuchungen zur Negation und zur Substantivflexion in altfranzösischen Urkunden der Grafschaft Luxemburg 1237–1281*. Tübingen: Niemeyer.
2006. 'Politique, développement socio-économique et histoire des langues: Galloromania', in Ernst et al. (2003–08), pp. 1178–90.
- Voßler, Karl 1929. *Frankreichs Kultur und Sprache*. Heidelberg: Winter.
- VR = *Vox Romanica*.
- Vrabie, Emil 1992. 'Slavic influence on Romanian: a case of exaggeration', *General Linguistics* 32:105–10.
- Vurpas, Anne-Marie 1995. 'Les scriptae francoprovençales', in *LRL* (II.2), pp. 389–405.
- Wallace, Rex 2011. 'The Latin alphabet and orthography', in Clackson, James (ed.), *A Companion to the Latin Language*. Oxford: Blackwell, pp. 9–28.
- Wallensköld, Axel (ed.) 1968. *Les Chansons de Conon de Béthune*. Paris: Champion.
- Wallington, Katherine 2008. '"... tradotte da dotta penna": New evidence for a reappraisal of the 1788 Zatta edition of Le memorie di Carlo Goldoni', *The Italianist* 28:203–16.
- Waquet, Françoise 2002. *Latin or the Empire of a Sign. From the Sixteenth to the Twentieth Centuries*. London/New York: Verso.
- Ward-Perkins, Bryan 2005. *The Fall of Rome and the End of Civilization*. Oxford: Oxford University Press.
- Wartburg, Walther von [1934] 1971. *Evolution et structure de la langue française*. Bern: Francke.
1950. *Die Ausgliederung der romanischen Sprachräume*. Bern: Francke.
- 1922–2003. *Französisches Etymologisches Wörterbuch*. Basel: Zbinden.
- Weimar, Peter 1972. 'Zur Entstehung des sogenannten Tübinger Rechtsbuches und der *Exceptiones legum Romanorum* des Petrus', in Wilhelm, Walter (ed.), *Studien zur europäischen Rechtsgeschichte. Festschrift Coing*. Frankfurt: Klostermann, pp. 1–24.
- Weinreich, Max 1956. 'The Jewish languages of Romance stock and their relation to earliest Yiddish', *RPh* 9:403–28.
- Weinreich, Uriel 1953. *Languages in Contact. Findings and Problems*. New York: Linguistic Circle of New York.
1954. 'Is a structural Dialectology possible?', *Word* 10:388–400.
1968. 'Uniligisme et bilinguisme', in *Encyclopédie de la Pléiade. Le Langage*. Paris: Gallimard, pp. 647–88.
- Wells, Colin 1984. *The Roman Empire*. London: Fontana.
- Wendt, Heinz Friedrich 1960. *Die türkischen Elemente im Rumänischen*. Berlin: Akademie-Verlag.
- Wenzel, Siegfried 1990. 'Reflections on New Philology', *Speculum* 65:11–18.

- Wexler, Paul 1981. 'Jewish interlinguistics: facts and conceptual framework', *Language* 57:99–149.
1988. *Three Heirs to a Judeo-Latin Legacy. Judeo-Ibero-Romance, Yiddish and Rotwelsch*. Wiesbaden: Harassowitz.
- Whinnom, Keith 1956. *Spanish Contact Vernaculars in the Philippine Islands*. London: Hong Kong University Press.
1965. 'The origin of the European-based creoles and pidgins', *Orbis* 15:509–27.
- Whitney, William 1881. 'On mixture in language', *Transactions of the American Philological Association* 22:5–26.
- Wilhelm, Raymund 2011. 'Lombardische Schreibtraditionen im 14. und 15. Jahrhundert', in Dessi Schmid *et al.* (2011), pp. 151–69.
- Windisch, Rudolf 1989. 'Rumänisch: Varietätenlinguistik des Rumänischen', in *LRL* (III), pp. 464–80.
1993. 'Le passage à l'écrit et la constitution d'une identité nationale: l'exemple du roumain', in Selig *et al.* (1993), pp. 149–56.
- Winford, Donald 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.
2005. 'Contact-induced changes: classification and processes', *Diachronica* 22:373–427.
2010. 'Contact and borrowing', in Hickey, Raymond (ed.), *The Handbook of Language Contact*. Chichester: Wiley–Blackwell, pp. 170–77.
- Winkelman, Otto 1989. 'Rumänisch: Lexikographie. Lexicographie', in *LRL* (III), pp. 492–507.
- (ed.) 1993. *Stand und Perspektiven der romanischen Sprachgeographie*. Wilhelmsfeld: Egert.
- Winniffrith, Tom 2002. 'Vlachs', in Clogg, Richard (ed.), *Minorities in Greece. Aspects of a Plural Society*. London: Hurst, pp. 112–21.
- Wolf, Joseph and Crook, John 1989. *Rechtsurkunden in Vulgärlatein aus dem Jahren 37–39 n. Chr.* Heidelberg: Winter.
- Wolfram, Herwig 1998. *Das Reich und die Germanen. Zwischen Antike und Mittelalter*. Berlin: Siedler.
- Woll, Dieter 1994. 'Portugiesisch: Grammatikographie. Gramaticografia', in *LRL* (VI.2), pp. 649–72.
- Wood, Ian 1990. 'Administration, law and culture in Merovingian Gaul', in McKitterick (1990), pp. 63–81.
- Woolf, Greg 1998. *Becoming Roman. The Origins of Provincial Civilization in Gaul*. Cambridge: Cambridge University Press.
- Wright, Roger 1982. *Late Latin and Early Romance in Spain and Carolingian France*. Liverpool: Francis Cairns.
- (ed.) 1991. *Latin and the Romance Languages in the Early Middle Ages*. London: Routledge.
1993. 'Complex monolingualism in early Romance', in Ashby, William and Mithun, Marianne (eds), *Linguistic Perspectives on Romance Languages*. Amsterdam/ Philadelphia: Benjamins, pp. 378–87.
1994. 'La muerte del ladino escrito en Al-Andalus', *Euphrosyne* 22:255–68.
- 1995a. *Early Ibero-Romance. Twenty-one Studies on Language and Texts from the Iberian Peninsula between the Roman Empire and the Thirteenth Century*. Newark: Juan de la Cuesta.
- 1995b. 'L'ensemble latino-roman du septième siècle', in Callebat, Louis (ed.), *Latin vulgaire–latin tardif IV*. Hildesheim: Olms-Weidmann, pp. 103–12.
1996. 'Latin in Spain: early Ibero-Romance', in Nielsen and Schøsler (1996), pp. 277–98.

1997. 'Translation between Latin and Romance in the Early Middle Ages', in Beer, J. (ed.), *Translation. Theory and Practice in the Middle Ages*. Kalamazoo: Western Michigan University, pp. 7–31.
 1998. 'Il latino: da madrelingua nativa a lingua straniera', in Herman (1998a), pp. 77–85.
 1999. 'Periodization and how to avoid it', in Blake, Robert, Ranson, Diana and Wright, Roger (eds), *Essays in Hispanic Linguistics dedicated to Paul M. Lloyd*. Newark: Juan de la Cuesta, pp. 25–41.
 2000. 'Latino e romanzo: Bonifazio e il Papa Gregorio II', in Herman, József, Marinetti, Anna and Mondin, Luca (eds), *La preistoria dell'italiano*. Tübingen: Niemeyer, pp. 219–29.
 2001. 'La chronologie relative des nouvelles *scripta* et des nouvelles langues', in Banniard, Michel (ed.), *Langages et peuples d'Europe. Cristallisation des identités romanes et germaniques*. Toulouse: Presse Universitaire, pp. 261–70.
 2003. *A Sociophilological Study of Late Latin*. Turnhout: Brepols.
 2004. 'La representación escrita del romance en el Reino de León entre 1157 y 1230', in Fernández Catón, José María (ed.), *Orígenes de las lenguas romances en el reino de León, siglos IX–XII*. León: Archivo Histórico Diocesano, pp. 273–91.
 2006. 'Language and Religion in early medieval Spain', in Bremer, Ernst *et al.* (eds), *Language of Religion – Language of the People. Medieval Judaism, Christianity and Islam*. Munich: Fink, pp. 115–26.
 2008. 'The monolingual Latin glossaries of the Iberian Peninsula: can they help the Romanist?', in van Acker, Marieke *et al.* (eds), *Latin écrit – roman oral? De la dichotomisation à la continuité*. Turnhout: Brepols, pp. 137–58.
 - Wüest, Jakob 1979. *La Dialectalisation de la Gallo-Romania. Problèmes phonologiques*. Berne: Francke.
 - Wuilleumier, Pierre 1963. *Inscriptions latines des Trois Gaules*. Paris: Centre national de la recherche scientifique.
 - Wunderli, Peter 1965. 'Die ältesten romanischen Texte unter dem Gesichtswinkel von Protokoll und Vorlesen', *VR* 24:44–63.
 - Zamboni Alberto 1998. 'Cambiamento di lingua o cambiamento di sistema? Per un bilancio cronologico della transizione', in Herman (1998a), pp. 99–127.
 1999. 'Dal latino tardo agli albori romanzi: dinamiche linguistiche della transizione', in *Settimana di Studio del Centro Italiano di Studi sull'Alto Medioevo XLV*. Spoleto: CISAM, pp. 619–98.
 - Zamora, Juan Clemente 1977. 'Interferencia recíproca: receptividad y productividad', *Word* 28:132–38.
 - Zamora Vicente, Alonso 1967. *Dialectología española*. Madrid: Gredos.
 - Zawadowski, Leon 1961. 'Les dialectes d'origine différente en contact', *Orbis* 10:293–307.
 - Zimmermann, Klaus 1992. *Diglosia y poliglosia*, in *LRL* (VI.1), pp. 341–54.
- ZRPh = *Zeitschrift für romanische Philologie*.

Index

- ablative case 100, 105, 112, 139
- Academia de Ciencias* 266
- Academia de la Llingua Asturiana* 341
- Academia de Matemáticas* 250
- Academia Orthográfica Portuguesa* 266
- Académie française* 173, 256, 259, 264
- Academies 145
- Acadia 363, 364
- Accademia della Crusca* 257, 259, 264, 267
- Accarisio 244
- acceptance 148
- Accessibility Hierarchy 434
- accommodation 97
- accusative
 - case 93, 139
 - and infinitive 113, 139, 140
- Acker, van 104, 137
- acrolect 344
- acronyms 213
- Adam, N. 264, 419
- Adams, J. 12, 15, 22, 25, 26, 27, 33, 43, 62, 66, 134
- Adams, M. 333
- address forms 217
- adjectives 202, 208, 224, 227, 236
- administrative terminology 212
- adoption 144
- Adrianople, Battle of 38
- Adriatic 6
- adstrate 98, 378–387
- adverbs 223–225, 230, 232, 233
- Æneid* 61
- affixes 114, 205, 211, 233; *see also* prefixes;
 - suffixes
- affrication 198
- Afonso Henriques 184
- Afonso I (of Portugal) 184
- Africa 6, 9, 13, 14, 16, 18, 19, 29, 33, 34, 38, 39, 130,
 - 190, 287, 316, 356, 357, 360, 363, 365–367, 368, 369, 395, 430
 - northern 193, 290, 353
 - West 65, 68, 74, 130, 376, 422, 430
- African languages 190, 192, 376–378, 430, 431
- Afrikaans 436
- Agostiniani 199
- agreement 398
- Agrigento 21
- Ahlqvist 239
- AIS 331
- Aki Yerushalayim* 399
- Al Andalus* 84, 288
- Alamans 202, 208
- Albania 286, 287, 311, 315
- Albanian 12, 13, 22, 193, 201, 284, 307, 315, 359
- Alberti, Leon Battista 243
- Albertini Tablets* 24
- Albi 173, 174
- Albigensian crusade 174, 176, 297
- Alcáçovas/Alcazovas-Toledo, Treaty of 362
- Alcobaça 251
- Alcuin 46, 48, 70, 105, 130, 131, 157
- Alemannic 284, 287, 295, 299, 350, 355
- Alessandri 249
- Alexander VI 248
- Alfieri 268
- Alfonso VI 183, 184
- Alfonso X (the Wise) 51, 109, 122, 182, 185, 248,
 - 289
- Algarve 292
- Algeria 367
- Alghero 292, 293, 302, 304

- Alicante 316
 Alinei 237
 Alleyne 424, 428, 429
Alliance israélite universelle 368, 399
 Allobrogians 43
 allomorphy 235
 Alonso 249
 Alpes Maritimes 303
 alphabet 132, 134, 135
 Cyrillic 218, 254, 256, 280, 310, 341
 Roman 218, 310, 341
 Alphonsine Renaissance 185
 Alpine (language) 193, 195
 Alps 6, 353
 Alsace 295, 315, 350
 Alsatian 315, 350
 Alsatians 358
 Altieri Biagi 258
 Alto Adige 287, 300, 307, 315, 358
 Alunno 244
 Alvar 177, 191
 Alvar Ezquerro 250
 Álvares 253
 Alvernii 44
 Ama y Borbón 265
 Amaral 439
 Amazonians 357
 ambiguity 74
 America 7, 107, 356, 363, 367, 370, 372,
 373–376
 American languages 192, 358
 Americanisms 343
 Americas 289, 290, 316
 Amerindian languages 194, 431
amestáu 314
 Amiens 37
 Ammianus Marcellinus 38, 39, 40, 337
 Amoretti 278
 Ampurias 22
 analogy 229
 analytic structures, analyticity 192, 403
 Anatolia 290
 Andalusia 177, 338; *see also* Spanish, Andalusian
 Andes 375
 Andler 92
 Andorra 293
 Andreescu 240
 Andreose 6, 282, 332, 361
 Andrieux-Reix 106
 Angenendt 79
 Angers 170
 Angevin 169
 Angevins 43
 Anglicisms 114
 Anglo-Norman 122, 168
 Anglo-Saxon 49, 135
 Angola 124, 316, 362, 365, 366, 376, 377
 Angolar, *see* Portuguese, of Angola
 animacy 217
 animals 194
 Anipa 131
 Anselmo 240
 anterior marker *ya* (Chabacano) 418
 Antioch 38, 40
 Antonine Wall 18
 Antoninus Pius 20
 Apamea 40
apertus sermo 78
 Apostles 73
 Appel 50, 189
Appendix Probi 29
 Apuleius 20
 Aquincum 29
 Aquitaine 131, 393
 Aquitania 9, 22
 Aquitanian 193
 Arab conquest 176, 248, 288, 353
 Arab occupation 248
 Arabia 337
 Arabic 3, 80, 84, 134, 135, 161, 183, 192, 226–228,
 229, 232, 285, 288, 355, 368, 385
 Moroccan 384
 Arabisms 135
 Arabs 63, 178, 183, 206, 211, 248, 287, 288
 Aragon 31, 44, 119, 121, 176, 180, 250, 289, 291,
 302, 313, 362
 Upper 19
 Aragonese 44, 109, 122, 167, 176, 177, 181, 248,
 288, 289, 313, 347
 Aramaic 193
 Aran 298
 aranés 298
 Arawak languages 374
 Arborea 166, 167
 archaism 12, 96
 architecture 146
Archivio glottologico italiano 278, 280
 Ardennes 39
 areal linguistics 320, 329–333
 Arends 427, 430
 Argentina 357, 361, 362, 369, 370, 386; *see also*
 Spanish, Argentinian
 Aribau 275
 Ariosto 242
 Ariza 180
 Arles 175

- Armenia 40
 Arnaldi 31
 Arnauld 261
 Arno 6
 Aromanian 229, 233, 235, 254, 287, 311
 Aromanians 213, 254, 309, 311, 312, 315
 moscopoleani 311
 muzăcheari 311
 pîndeni 311
arpitan (arpeitan) 299
 Arquint 254
 Arras 37
 Arrighi 364
Artes Lectoriae 131
 article (definite) 136, 194, 201, 230, 234,
 339, 371
 emergence 30–32, 35–36
 ILLE 32
 IPSE 30
 Aruba 407
 Arvinte 254
 Ascoli, G.I. 23, 154, 192, 245, 278–279, 280, 298,
 317
 Asia 6, 18, 289, 402, 422
Asociación de Academias de la Lengua Española
 391
 aspect 221–222, 223, 225
 continuous 392
 aspirates, aspiration 198
 Aspromonte 230
 assimilation 29
Association de co-opération culturelle et technique
 393
Association des écrivains de langue française 393
Association des universités partiellement ou
 entièrement de langue française 393
Association Vidas Largas 399
 Asturian 248, 290, 313, 324, 326, 327, 341, 347
 Asturias 176, 226, 290, 291, 292, 313, 314, 327
 Asturo-Leonese 288, 289, 290, 326
 Atanasov 225, 311, 312, 313
 Athens 20
Atlant linguistich dl ladin dolomitich y di dialec
 vejins 322
 Atlantic 250, 363
 Atlantic Provinces 372, 380
Atlas linguistique de la France 154, 170, 317, 318,
 346
 atlases (linguistic) 318
 Atsma 76, 106
 Audin 13
 Auer 145, 150, 386, 411
 Auerbach 65, 73
 Augustine 25, 39, 66, 67, 68, 73, 74, 77, 78,
 130; *see also* *City of God*
 Auroux 63, 263, 404, 417
Ausbau 148, 160
 Australia 362
 Austrasia 88
 autonomías 109
 autonomization 433, 442
Autoridad Nacional del Ladino i su Kultura 399
 Autun (Saône-et-Loire) 37
 Auvergnans 43, 44
 Auvergnat, *auvergnat* 173, 298
 Auvergne 39, 50
 Auxerre (Yonne) 37
 auxiliaries 112, 136
 Avas 31, 47, 64, 104, 106
 Avar 286
 Avignon 122
 Avitus of Vienne 137
 Avolio 300, 328
 Aymarâ 358, 373, 374
 Ayres-Bennett 123, 124, 134, 259, 260, 264
 Azores 363

bable 291
 Babylon 40
 Bachmann 3, 364, 404, 405, 407, 410, 412, 413,
 418, 439, 440
 Badia i Margarit 178
 Bagaldi 230
 Baggio 89
 Baggioni 395, 404
 Baglioni 308, 343
 Bagno 437
 Bailey 415, 416
 Baker 427, 429, 436
 Bakker 388, 389
 Bal 190
 Baldinger 148, 177
 Balearic Islands 178, 226, 292, 341
 Balkan languages 232
 Balkan Romance 161, 225
 Balkans 285, 290, 311, 368
 Ballot i Torres, Josep Pau 266
 Baltic states 359
 Banat 214, 309, 315
 Banchieri 258
 Banitt 367
 Banniard 5, 46, 58, 60, 61, 64, 67, 69, 70, 72, 73,
 74, 75, 76, 77, 78, 80, 81, 82, 83, 84, 85, 88,
 89, 91, 92, 93, 96, 99, 101, 106, 107, 115, 136,
 156, 157, 167
 Bantu languages 376, 377, 389, 427

- Barbieri 241
 Barcelona 115, 158, 178, 179, 251, 266, 275, 316, 346
 Baretti 269
 Barra Jover 160
 Barrère 351
 Barros 252
 Barthélemy 264
 Bartoli, D. 257
 Bartoli, M. 164, 282, 308, 320, 330, 331, 332, 333
 Basile, Giambattista 258
 basilect 344
 Basilicata 304
 Basque 12, 19, 20, 22, 109, 193, 195–197, 234, 284, 287, 289, 295, 296, 315
 Basque regions 287
 Basque-Navarrese Country 19
 Basques 358
 Bassi 268
 Battisti 230
 Bavarian 287, 355
 Bavoux 390, 398
 Baxter 437
 Bayer 77
 Béarn 174, 297
 Béarnais 173
 Beauzée 263
 Bec 174, 274
 Beccaria 250
 Beckmann 91, 93
 Bédarida 268
 Beinke 275
 Belardi 254
 Belgium 6, 285, 294, 296, 315
 Bello 391, 393, 403
 Bello Rivas 147, 185
 Bellunese 316
 Belorussian 359
 Beltrán 174, 180, 185
 Bembo 166, 241, 242
 Bénabu 18
 Benedictine, Benedictines 158, 164, 166
 Benelux countries 316
 Benevento 80, 205, 210
 Beni 257
 Beniak 378, 380, 382–383, 395
 Bénin 362
 Benincà 254, 269, 279, 280, 333
 Benkner 254
 Bentahila 384, 385, 386
 Berber 13, 22, 285
 Berbers 354
 Berg, van den 364
 Bergamasque 346
 Bergé 96
 Bernard 404
 Berneri, G. 258
 Berni 242
 Bernstein 203, 209
 Berrichon, *berrichon* 169, 293
 Berrichons 43
 Berruto 313, 369
 Berry 69
 Berschin 237, 342
 Bessarabia 309, 310
 Besslingen 39
 Betsiamites 389, 390
 Beumann 70
 Bianchi 164
 Biber 92
 Bible 112, 189, 240, 241, 261
 Bickerton 426, 427, 428, 429, 441
 Bifrun 254
 bilingualism, bilinguals 21, 22, 23, 41, 55, 85, 119, 187, 188, 189, 192, 193, 204, 206, 210, 212, 222, 235, 236, 251, 253, 268, 280, 286, 297, 313–316, 327, 335, 349, 350, 356, 370, 379, 386, 410
 Bill 101 395
 Biloa 376, 377, 390
 bioprogram 426–428, 429
 Bischoff 48
bisiacco 301
 Bitola 311
 Blake 115
 Blanc 379, 395
 Blanche-Benveniste 68
 Blas Arroyo 190, 192
 Blasco Ferrer 12, 166, 167, 178, 179, 270, 341
 Bloomfield 422
 Boccaccio 126, 165, 242, 306
 Bodvarsson 364
 Boerio 279
 Boiardo 242
 Boisvert 269
 Bolivia 361, 370, 373, 374
 Boll 404
 Bollée 397
 Bologna 158, 171, 258
 Congress of 249
 University of 159, 165
 ‘Bologna Discourse’ 159
 Bolognese 346
 Bolognese Renaissance 165
 Bonaduz 356
 Bonaire 407

- Bonfante 66, 156, 288
 Boniface 80
 Bonifacio 285, 303, 304
 Bonvini 431
 Boretzky 433
 Borghi 278
 Börner 185
 borrowing 3, 114, 187–225, 430
 ‘clusters’ in 195, 197, 231
 of convenience vs. borrowing
 of necessity 190
 lexical 418; *see also* loanwords
 morphological 234
 phonological 234
 semantic 228
 syntactic 234, 442
 Borst 73
 Bosnia 306, 313, 316
 Bosnia-Herzegovina 304
 Bossong 183, 191
 Boulogne-sur-Mer 37
bourbonnais 293
 Bourbons 289, 293, 304
 Bourciez 317
 Bourges 170
 Bourghiba 393
 Bourguignon, *bourguignon* 169, 294
bozal 364
 Brâncuș 201
 Brașov 254
 Brazil 124, 316, 357, 361, 362, 363, 372, 373, 374,
 387, 392, 431
 ‘Brazilian’ 123
 Brea 275, 276
 Bresse 299
 Breton 22, 284, 287, 294, 295, 296, 315
 Bretons 358
 Britain, Britannia 9, 13, 18, 19, 22, 38, 40, 287,
 304, 337, 353, 354, 364
 British language 22
 British Isles 193
 Brittany 22, 193, 315, 393
 Broglio 278
 Brown 66, 73, 91, 102
 Bruch 202, 208
 Brucioli 241
 Brun-Trigaud 61
 Brun, A. 168
 Brun, R. 390
 Brunel 173
 Brunetto Latini 345
 Brunhilda 39
 Brunhölzl 48
 Bruni 48, 123, 241
 Brunot 145, 171, 238, 247, 259, 260, 263, 265, 267,
 271, 272
 Brussels 47, 287, 315
 Bucharest 254, 281
 Bucovina 309
 Budai-Deleanu 272
 Budapest 255, 281
 Buenos Aires 63, 387
 Buffet 260
 Bugeac 309, 310
 Buisset 51
 Bulgaria 309, 310, 311, 315
 Bulgarian 201, 213, 214, 226, 285, 359
 Bulgarians 80, 81, 213, 309, 316, 359
 Bulgars 286
 Bulu 376
 Burgos 158, 181, 182
 Burgundians 202, 204, 209
 Burgundy 52, 168, 171, 202, 208, 294
 Buridant 167
 Burkina Faso 362
 Burnett 120
 Burundi 362
 Busse 191
 Butler 91, 92
 Byzacena 287
 Byzantine period 230
 Cabral, A. 365
 Cabral, P. 363
 Cáceres 290
 cacuminalization 23
 Cadore 26
 Caesar, Julius 13
 Cæsarius of Arles 76, 78
 ‘cafone’ 346
 Cagliari 166
 Cairo 40
 Cajun 397
 Čakavian 221
 Calabria 7, 22, 196, 230, 297, 304
 Calasetta 304
 Calboli 80, 189
 Calderón de la Barca 262
 California 364, 369
 Callæcia 183
 calques, calquing 189, 206, 211, 213, 214, 219,
 225, 228, 232, 236, 340, 369, 375, 379
 Calveras 275
 Calvet 59
 Calvin 241
 Calvinists 256

- Cambrai, Peace of 249
 Cameron 73
 Cameroon 362, 376, 390
 Camfranglais 390
 Camões 126, 253
 Campania 6, 304
 Campanian 164
 Campbell 397
 Campidanese 167, 270, 301, 302, 329
 Camproux 274
 Canada 362, 366, 372, 378–383, 385, 388, 393, 394
 Canary Islands 194, 338, 362, 397
Cancioneiro Geral 253
 Canfield 372
 Cano Aguilar 177, 249
 Cano González 324, 326
 canon law 175
 Cantabria 181, 226, 290, 291, 326
 Cantacuzino 261
Cantares Gallegos 276
 Cantemir 262
Cantigas de Santa Maria 185
 Canut 377
 Cape Verde 362, 363, 364, 366
 Caracausi 23
 Caragiale 281
 Carbonell 275
 Carbonia 304
 Cardona 147, 238
 Caribbean 357, 372, 373; *see also* Spanish, Caribbean
 Carinthian 284
 Carles 36, 50
 Carloforte 304
 Carlos III 266
 Carolingian(s) 79, 81, 82, 83, 84, 87, 88, 117, 118, 119, 130, 170
 dynasty 204, 210–211
 Empire 48, 53
 minuscule 46, 48
 reforms 46, 48, 52, 53, 100, 101, 102, 131, 137, 138, 157, 336
 Renaissance 69
 Carolus-Barré 155
 Carrié 65, 72
Carta consolare pisana 166
Carta dei dialetti d'Italia 322, 328
Carta di Arborea 166
 Cartago 269
 Carthage 100
 Carthaginian settlements 21
cartilhas 252–253
cartillas 250
cartinhas 252; *see also* *cartilhas*
 Carvalhão Buescu 251, 252
 Casanova 251, 275
 Casati Law 279
 case system 202, 207, 208, 217, 221; *see also* *under names of individual cases*
 loss/weakening of 35, 93, 163, 366
 Caspe, Compromise of 250
 Cassano 191
 Castellani 165, 277, 306
 Castellum Tidditanorum 17
 Castelveccchi 243
 Castiglione 241
 Castile 51, 109, 112, 115, 120, 122, 176, 184, 185, 196, 288, 290, 291, 292, 302, 326, 327, 342, 362
 New 126
 union with Aragon 122, 250, 291, 292
 Castilian, *see* Spanish
 Castilianization 374
 Castilians 359
 Castillo Lluch 183
 castrapo 314
 Castro, R. de 276
 Catach 247
 Catalan 7, 32, 33, 44, 109, 115, 119, 120, 121, 122, 141, 167, 174, 175, 176, 177, 179, 180, 181, 203, 205, 206, 209, 210, 211, 227, 233, 250–251, 265–266, 276, 285, 288, 289, 291, 292, 293, 295, 296, 297, 302, 307, 313, 316, 322, 331, 341
 Menorcan 266
 revival 275
 Catalans 43, 51
 Catalonia 44, 56, 118, 178, 179, 205, 210, 250, 265, 266, 275, 289, 293, 298, 301, 314, 317, 322, 359
 Catania 21
 Cateau-Cambrésis, Treaty of 249
 Catherine des Médicis 246
 Catholic Monarchs 248, 275, 289, 362, 367
 Catholics, Catholicism 240, 255, 281, 299, 350
 CAUANNUS 26
cauchois 294
 Caudmont 238
 Cava dei Tirreni 35
 Cavigelli 356
 Cazacu 255, 261, 262, 271
Cazania 262
 Celle San Vito 298, 299
 Celtiberian, Celtiberians 22, 199
 Celtic 3, 6, 9, 13, 14, 22, 26, 129, 193, 194, 199, 200, 234, 284
 Celts 354
Cena Trimalchionis 113
 Central African Republic 362

- Central America 372, 373
 Cernăuți 310
 Cerquiglini 145, 167
 Certeau 273
 Cervantes 126, 262, 409
 Cervera 266
 Cesari 268
 Cesarotti 269
 Chad 362, 366
 Chalons-sur-Saône 37
 Champagne 168, 171, 172
 Champenois, *champenois* 169, 294
 Changeux 62
 Channel Islands 294
Chanson de Roland 49, 61
 Charlemagne 27, 46, 52, 53, 130, 169, 205, 210
 Charles III 265, 289
 Charles V 248, 289
 Charles-Picard 15, 16, 17
 Charleston 369
 Charpentier 398
 Charpin 103
Charte de la langue française 395
Chartes 151
 'chartes' 151
 Chartier 240
 Chartres 245
 Chaudenson 401, 402, 407, 423, 425, 430, 432, 433, 436
 Cherubini 279
 Chervel 247
 Chia 390
 Chiac 390
 Chiappelli 237
 Childebert 39
 Chile 361, 370, 372
 China Sea 363
 Choi 374
 Chomarat 64
 Christianity, Christians 10, 19, 40, 66, 72, 73, 75, 76, 79, 80, 84, 87, 91, 101, 115, 119, 136, 157, 176, 203, 204, 209, 226, 228, 248, 288, 355, 358
Chroniques of Jean Froissart 51
 chronostratigraphic atlas 322
 Chur 299, 350, 356
 Church 54, 55, 72, 75, 80, 119, 138, 241, 250, 261, 281, 354, 358
 Church Slavonic 215, 233, 255
 Cicero 242
 Cinonio, *see* Mambelli
 Cintra 150
 Ciobanu 359
 Ciolac 343
 Cioni 278
 Cirta 17
cismontincu 305
City of God 66, 67, 68, 78
 classification 333–334
 Claudin 240
 Claudius Terentianus 24
 clause linking 182
 cleft constructions 427
 Clement VIII 241
 clergy 157
 clitics 136
 loss of 439
 Close 281
 Clovis 204, 209
 Cluniac reform 157, 177
 Cluny 183
 clusters, *see* borrowing
 Clyne 396
 co-official languages 284
cocoliche 369, 386–387
 code-switching 146, 327, 370, 384–387, 430
Codex Iuris Civilis 175
 codification 148, 238–256, 275, 282
 Coimbra 184, 251
 Colho, Adolpho 404
collectivités d'Outre-Mer 362
 Collins 79, 84
 Cologne 18, 37
 Colombat 239
 Colombia 361, 373
 Colon 178
 Colón 322
 colour terms 203, 205, 208, 210
 Columbus 363
 Comas 250, 266
commedia dell'arte 346
 Commonwealth 393
 communicative networks 150
 Communism, Communists 309
 Como 25
 Comoros 362
Compañía de Jesús 248
 comparative method 28
 complementation 426
 complementizers 139, 433, 441
 comprehension, *see* intelligibility
 computers 321
 Condado de Portugal 184
Condaghi 166
 conditional
 emergence of new paradigms 35, 52

- Conférence des chefs d'état et de gouvernement ayant en commun l'usage du français* 394
- Confessions* 77
- Congo, Democratic Republic 362, 376
- Congo, Republic 362, 377, 385
- conjugation 97
- conjunctions 229, 233
- Connecticut 368, 369
- Conon de Béthune 171, 345
- Conquistadors 358
- Consentius 130
- consonantal weakening, *see* lenition
- Constantine 18
- Constantine, Peace of 79
- Constantinople 39
- contact 3, 19, 187–225, 373–390, 401, 423, 424, 430, 431, 432, 433, 443, 444
indirect 189
- contaminations 200
- Conti, A. 268
- Contini 324, 328, 329
- Canto navale pisano* 164
- convergence 143, 144, 147, 149, 150, 151, 156, 167
- convergence interference 147
- Coppino Law 279
- copula 414, 415
- Corbie 170
- Córdoba 84
- Corella 250
- Coresi, Deacon 256
- Corne 427, 429
- Cornwall (Ontario) 383
- Corominas 19, 20
- corpora, corpus 59, 127, 143
- Corpus Iuris Civilis* 159
- Corrà 316
- Correas 249
- Correctio* 157
- Corsica 9, 56, 285, 296, 301, 303, 304, 305, 314, 342, 348
- Corsican 198, 295, 305, 306, 313, 334, 339, 342
- Corsicans 316
- Corso, R. 244
- Cortelazzo 258, 268, 270
- Cortese 258
- Coseriu 92, 144, 145, 146, 147, 175, 232, 235
- Costa Rica 361
- Costin 261
- Costums de Tortosa* 179
- Côte d'Ivoire 362, 399
- Côtis-Capel 293
- Coulmas 237
- Coulon 37
- Counter-Reformation 241, 250
- courtly lyric 297
- coutumiers* 175
- Covarrubias 31, 264
- Cravens 111, 134
- Creangă 281
- creativity 319
- Cree 388–390
- Cremona, J. 245, 308, 332, 343
- creoles 3, 109, 194, 350, 351, 363, 365, 366, 400–444; *see also* bioprogram;
decreolization; *patois*; relexification
hypothesis; substrate; superstrate
- African grammar 419
- American French 420
- Angolar 438
- articles 419–420
- Cape Verde 366, 414, 415, 438
- Caribbean 415, 424, 429, 430, 434, 436, 442
- Chabacano 418
- corrupted forms of Romance 402–403
- definition 412
- Dominican French 423
- English-lexifier 407, 415, 424
- French-lexifier 363, 405, 406, 419–422, 424, 429
as fully-fledged natural languages 412–417, 443
- genesis 401, 403, 407, 418–431
- grammatical functions 413–414
- Guadeloupe 435
- Guyana 434, 435, 443
- Haitian 421, 434, 441
- history of Creole Studies 401–417
- India 424
- Indian Ocean 420, 434
- Jamaican 441
- Lesser Antilles French 421
- lexifiers 401, 402, 405, 408, 411, 413, 414, 420, 421, 422, 436, 441
- Louisiana 434, 435
- Martinique 434, 443
- Mauritian 419, 420, 427
- monogenesis theory 422–426, 428
- nativized 422
- noun phrase in 419–420
- optimal systems 419–422
- origin of term 402, 406
- pan-Afro-American dimension 415
- Papiamentu 405, 407, 409, 413–414, 419, 421, 423, 425, 438, 441, 442
- Philippine Spanish 418, 423; *see also* creoles, Chabacano
- pidgin–creole life cycle 429–431

- creoles (cont.)
 Portuguese-lexifier 364, 406, 415, 422, 424,
 429, 438
 relexification 424, 426, 428, 429
 restructuring 433–443
 Réunion 435, 436
 São Tomense 405, 438
 São Tomense nasalization 405
 São Tomense sound changes affecting /r/ 405
 Saramaccan 425
 Seychelles 435, 436
 Spanish-lexifier 424, 429, 438, 442
 Sranan Tongo 425
 tense, mood, aspect (TMA) markers 414,
 421, 422, 426, 427, 435, 443
 Upper Guinea Portuguese 427
 verb system 427
 as vernacular languages 404–409
- creolization 402, 432
- Crevatin 302
- Crimea 304
- Croatia 303, 313, 315
- Croatian 214, 218, 219, 220, 221, 222, 223, 284,
 285, 307, 308, 313, 315, 355
- Croce 250, 258, 319
- Croissant 297
- Crook 29
- Crusades 226
- Cuba 342, 361, 364
- Cuervo 403
- Cugno 322
- Cumans 308
- Curaçao 407
- Cushitic languages 389
- Cyrenaica 286
- Cyrino 439
- D. Belica 311
- D'Agostino 250
- D'Anglejan 395
- Dachsprache* 156, 161
- Dacia 9, 262, 272, 287
- Daco-Romanian 213, 254–256
- Dahmen 311, 312, 342
- Dakar 385
- Dalmatia 22, 29, 282, 285, 287, 288, 304, 308,
 353, 354
- Dalmatian 162, 202, 282, 285, 287, 303,
 308, 355
- Daman 362
- Danet 264
- Danish 401
- Dante 43, 65, 165, 242, 306
- Danube 13, 18, 38, 40, 213, 214, 218, 229, 286,
 287, 335, 353, 355
- Upper 22
- Dardel 28, 133
- Dardi 267, 268, 278
- Dark Ages 65
- Darwin 11
- database 322
- dative 97, 100, 230
- dativo greco* 230
- Dauphinois 169
- Dauzat 317, 320
- Davies, E. 384, 385, 386
- Davies, M. 127
- Dazzi 254
- de Bartholomaeis 32
- De Blasi 164
- De Fort 279
- De Mauro 277, 279, 306
- de Morais Madureira Feijó 266
- De Ordine* 25
- De Orthographia* 48, 131
- De Rooij 441
- de Sale, Flaminio 276
- De Vulgari Eloquentia* 43, 165
- De Wailly 266
- Débax 32, 33
- DeCamp 351, 410, 415, 416, 440
- declension 97
- decreolization 436–443
- Decretum* 175
- Dees 152, 154, 155
- degemination 29
- DeGraff 402, 412, 428
- deictics 128
- del Rosso, P. 244
- Della Valle 256
- Delton 302
- demonstratives 339
- Densuşianu 256, 311
- dental fricatives 228
- départements d'Outre-Mer* 362, 393
- deponent verb 97, 139
- derogatory forms 205, 211
- Derrer 175
- des Pepliers, J.-R. 261
- Désirat 263
- Desportes, Philippe 259
- determiner system 420, 426
- Detges 427
- Dettori 270
- Devoto 163
- Deyck, van 59

- diachronic phonology 320
 diachrony 152, 154, 319
 dialect 237, 283, 294, 304, 305, 307, 316, 317, 318,
 406, 410–411, 416
 continuum 161
 levelling 343, 401, 432
 simplification 432
 dialectalization 169, 361
dialecto fronterizo 191
 dialectology 59, 321, 322, 329
 Diallo 396
Dialogues 78
 diamesic variation 145, 149
 diaphasic variation 145, 184
 diaspora 367
 diastratic variation 77, 145, 184, 344, 348
 diasystem 94, 95, 96, 100, 105, 146, 149, 150, 351
 diatopic variation 145, 152, 154, 184, 336, 344, 345
 Díaz 100, 177
Dictionnaire correctif du français du Canada 395
Dictionnaire de nos fautes contre la langue
 française 395
Dictionnaire du français de Côte d'Ivoire 396
Dictionnaire du français québécois 395
Dictionnaire du judéo-espagnol 399
Dictionnaire historique du français québécois 395
Dictionnaire québécois d'aujourd'hui 395
 Diez 59, 317, 403
 diffusion 144
 diglossia 24, 86, 87, 127, 158, 245, 297, 299, 302,
 305, 313–316, 327, 328, 336, 343, 344, 348–353,
 395, 401, 410, 441
 conflictual 314
 contaminated 313
 Dignano 285, 302
 digraph 129, 135
 dilalia 313
 Dima 282
 diminutive 320
 suffixes 128
 diocesan divisions 178
 Diocletian 9
 Dionisotti 239
 Dioscorides 200
 diphthongization, diphthongs 35, 47, 181, 207,
 326, 347
 discourse markers 233
 discourse tradition 144, 146, 160
 divergence 147, 149, 151
 divergence interference 147
 Djibouti 362
 Dniester 310
 Dobrogea 309, 311, 312
 Dolbeau 67
 Dolce, L. 244
 Dolet 247
 Dolomites 254, 300
 Dom Dinis 185
 Dominican Republic 361
Don Quijote 126
Donat 174
 Donatus 130
Donatz 174
Donatz proensals 174
 Donner 14, 15
 Donzé 261
 Doria 282, 355
 dos Reis Gonçalves Viana 267
 Dossens 134
 Douro / Duero 291
 Drăganu 201
 Drescher 366
 Drusi 241
 Drusus 13
 du Bellay, Joachim 246
 Du Marsais 263, 264
 Du Tillot, Guillaume 270
 du Wes, Gilles 247
 Duala 376
 Duarte 439, 440
 Duarte i Montserrat 179
 Duarte Nunes de Leão 342
 Dubois, Jacques 247
 Dubrovnik 162, 285; *see also* Ragusa
 Dumestre 377
 Dumézil 65
 Dumoulin 107
 Duponchel 396
 Durante 165
 Dutch 367, 401, 407, 409, 425, 442
 Duval 102
 Dyer 341, 359
 East Timor 124, 362
 Eberenz 123, 178
 ecclesiastical administrative areas 168
 Eckert 124
 Ecuador 361, 372, 373, 388; *see also* Spanish,
 Ecuadorean Sierra
 Edict of Nantes 369
 Edictum Rhotari 29
 education system 75, 366, 398; *see also* schools
 Edwards 88
 Efoua-Zengue 390
 égalité 351
 Egerland 206, 211

- Egypt 24, 38, 394
 Egyptian 22
Einsiedel Linear Version 162
 Eisenstein 240
 El Carche 285
 El Salvador 361
 El-Amrouni 14, 16
 El-Kef 16
 elaboration 148, 160
 Elcock 1, 361
elegantia 73
 Elia 238
 Elipand of Toledo 80
 Elizaincín 190, 191, 387
 Eltis 430
 Elwert 192
 Elymian 21
 Embedded Language islands 384
 emigration 304, 311, 312
 Emilian 328
 eastern 26
 Emilian Glosses, *see* Glosses of San Millán
 Emiliano 47, 115, 124, 137
 Eminescu 281, 341
 Engadine 299
 Engadinish 33, 254
 England 172, 250
 Englebert 321
 English 82, 85, 114, 117, 129, 132, 139, 141, 190,
 192, 220, 236, 267, 295, 308, 360, 364, 367,
 369, 378–384, 385, 393, 394, 395, 397, 399,
 401, 414, 418, 432, 434
 Afro-American 415, 416, 417, 436, 441
 Australian 117
 Jamaican 441
 Middle 123
 Negro 416
 White American 415
 Enlightenment 107, 263, 264, 267, 270,
 271, 347
 Entrains-sur-Nohain 37
 epithetic consonant 304, 324, 326, 347
 EPOTOSOROVIVS 13
 Equator 363
 Equatorial Guinea 361, 362, 367, 377
 Erice 21
 Ernst 1
 Estienne 246, 247
 Ethnologue database 362, 373
 Etruscan 6, 22, 193, 195, 197–199, 234
Etymologies (Isidore of Seville) 100
 Eucharius 26
 Eugippius 42
Eulalia Sequence 31, 157, 169, 170
 Europe 361, 362
 Eustatievici Braşoveanul 272
 evidentials 226
 Ewald 169
 Ewe 434
 Ewondo 376
 ‘exploitation colonies’ 363, 376–378
 Extremadura 177, 291
extremeiro 326

 /f/ > /h/ 195, 196, 234, 326
 Fabra 122
 Fabre d’Olivet 274
 face-to-face discourse 216
 factorization principle 426
 Faeto 298, 299
 Falkowski 100
 Far East 363
 Farete 260
fârşeroţi 311
 Fasana 302
 Fascism, Fascists 304, 307
fassano 300
 Fava, G. 165
 Febrer i Cardona 266
 Félix Carneiro Souto-Major 266
 Félibrige 274, 297
 Feral 390
 Ferdinand of Aragon 248, 250, 362
 Ferguson 245, 350, 351, 374, 410, 441
 Fermo 328
 Fernández Rei 314
 Fernández-Ordóñez 177
 Fernández-Vest 62
 Fernando III 182
 Ferrara 304
 Ferreira 254
 Ferrer i Gironès 266, 275
 Ferro 311
 feudalism 172
 Feugère 63
 Filastrius 29
 Filipi 313
 final consonants 338; *see also* /m/, /s/, /t/
 loss of 93
 finite clauses 230
 Fiorelli 246
 Firenzuola 243
 First World War 62, 214, 296, 307, 309, 312
 Fischer 217
 Fishman 59, 191, 313, 336
 Flanders 170, 295

- Fleischman 5, 34, 137, 441
 Flemings 358
 Flemish 287, 295, 296, 315
 Fleury-sur-Loire 117
 Flobert 90
 Florence 166, 240, 278
 Florentine 165, 166, 243, 258, 278, 306, 313
 fourteenth-century 242
 Florescu 271, 280
 Florida 364, 369
 Flydal 146
foederati 202, 211
 Fohlen 369
 Folena 267
 folk etymology 134, 320
 folk-tales 101
 Fontaine 74, 78
 Fontanella de Weinberg 432
 Formentin 32
Formulae Andecavenses 29
Foros de Castelo Rodrigo 150
 Forro 366
 Fortunio 244
Forum iudicum 178
 Fosso 390
 Fouché 346, 347
 Fourquet 64, 88
 Fragnito 241
 Franc-Comtois, *franc-comtois* 169,
 287, 294
français commun 351
français familier 351
français populaire 351
français régional 351
 France 6, 22, 31, 34, 51, 56, 107, 115, 116,
 117, 118, 120, 122, 139, 151, 157, 158, 159,
 167, 171, 174, 175, 176, 177, 183, 226,
 240, 245, 250, 270, 272–273, 283, 285,
 287, 290, 293, 296, 297, 298, 303, 305, 313,
 316, 317, 318, 332, 347, 351, 352, 358, 359,
 364, 369
 decline of dialects 273
 linguistic unification 272–273
 southern 274
 Franciscato 324, 327, 348, 349
 Franche-Comté 168, 294
Francia 84
 Francian, *francien* 44, 169, 172, 245
 Francians 43
francitan 314
 Franco-Provençal 169, 175, 203, 209, 245, 283,
 285, 287, 296, 298, 299, 307, 313, 314, 317,
 321, 350, 358
 Franco, F. 275, 289, 290, 292, 293, 341
 François I 176, 246, 295
 Franconian 286, 295
 Francophonie 393
 Frangache 390
 Frank 55, 106, 153, 155, 162, 166, 171, 173
 Frankish 199, 203–210, 211, 355
 Franks 40, 182, 203, 206, 211, 340
Franxa exterior 292
Französisches Etymologisches Wörterbuch 231
 Fredegar, Chronicle of 30, 34
 Frederick II 165
 free morpheme constraint 387
 Freire 365
 French 1, 7, 11, 33, 82, 85, 93, 99, 102, 103, 104,
 107, 109, 113, 114, 116, 119, 120, 122, 123, 124,
 129, 132, 137, 138, 139, 140, 141, 147, 152, 153,
 167, 173, 175, 176, 180, 194, 199, 202, 203,
 204, 205, 206, 208, 209, 210, 211, 212, 229,
 233, 238, 241, 246, 258–261, 263–264, 267,
 274, 275, 281, 282, 283, 284, 285, 287, 293,
 294, 295, 296, 297, 299, 306, 307, 314, 315,
 316, 328, 332, 333, 339, 342, 345, 350, 351, 358,
 360, 361, 362, 363, 364, 366, 367, 368–369,
 371, 372, 376, 378–383, 384, 388–390,
 393–396, 396–397, 398, 399, 401, 404, 418,
 419, 420, 421, 423, 441; *see also* *français*,
 Norman, *Oil*
 Acadian 372, 380, 381, 390, 395
 Belgian 395
 Canadian 384, 433
 Canadian Montreal 434
 Canadian Quebec (*québécois*) 342, 372, 378,
 380, 395, 434
 Classical 167
 Congolese 360
 corrupt 402, 433
 dialects 43, 203, 208
 as diplomatic language 271
 early French 32
 Judæo-French 367
 grammars 247, 258–261
 as international language 270
 in Italy 268–269
 lexis 259
 middle 108, 110, 114, 122, 123
 old 108, 109, 114, 121, 167
 orthography 247–248
 Parisian norm 272
 popular 421, 433
 proto- 83, 86, 99, 100, 106
 Renaissance 107
 sixteenth-century pronunciation 247

- French (cont.)
 spoken 352, 353
 standard 245, 407, 421, 432
 Swiss 395
 women's language and usage 260, 264
 written 353
- French Academy, *see Académie française*
- French Polynesia 362
- French revolution 123, 145, 295, 297, 347, 351
- French Southern and Antarctic Lands 362
- frequency 192
- Fribourg 350
- fricativization 197
- Friuli 46, 300, 301, 304, 307, 316, 317
- Friulian 162, 190, 204, 205, 209, 210, 213, 254,
 276, 284, 285, 287, 301, 307, 313, 314, 323,
 324, 325, 327, 347, 348, 349
- Friulians 304, 316
- Froissart 122
- fronteiriço* 387
- Fuchs 91, 403
- fueros* 126
- Fulda 47
- function words 204, 205, 210, 232, 233
- Furetière 264
- Furs de Valencia* 179
- 'fused lects' 386, 387
- Fusina 306
- fusion language 191
- future tense 99, 136, 352, 375; *see also*
 subjunctive
 future in the past 99
 Latin 'be' + active future participle 32
 new HABEO future 32–34, 35, 48, 52
 perfect 99
 relics of Latin synthetic future in early
 Romance 32
VOLO + infinitive 33
- G. Belica 311
- Gabon 362, 366
- Gabriele, Giacomo 244
- Gadet 92
- Gaeng 100
- Gaeta 32, 35
- Gagauz 284, 359
- Gagliardi 270
- Gáldi 271
- Galiani 269, 270, 271
- Galicia 49, 135, 184, 185, 291, 292, 314, 341, 342, 349
- Galician 109, 110, 121, 135, 141, 177, 183, 186, 191,
 248, 275–276, 289, 291, 292, 313, 314, 326,
 327, 342, 346, 349
- grammars and dictionaries 276
- revival 276
- standardization 276
- Galician-Portuguese 176, 181, 183, 275, 291, 327
- Galileo 257, 258
- Gallesano 302
- Gallia; *see also* Gaul
 Gallia Lugdunensis 9
 Gallia Belgica 9, 286
- Gallina 250
- gallo* 294, 338
- Gallo-Romance 26, 43, 45, 118, 119, 121, 134, 139,
 151, 167, 176, 177, 180, 194, 202, 203,
 205–212, 234, 287, 318, 330, 331, 332
- Gallura 166, 301
- Gallurese 329
- Gamillscheg 36, 201, 202, 203, 208, 209
- García Mouton 321, 322
- García Turza 120
- García, C. 191
- Garcilaso de la Vega 262
- Gartner 317
- Gascon 120, 173, 195, 196, 298
- Gascons 43
- Gaul 18, 26, 29, 32, 38, 39, 70, 77, 79, 84, 134, 135,
 136, 141
 Cisalpine 9
 Narbonese 9
 southern 56
 The Three Gauls 9
- Gaulish 22, 199
 inscriptions 29
- Gaunt 274
- Gazeta de Lisboa* 267
- Geary 65, 82
- Gemeinsprache* 411
- gender 163, 366, 398
- genealogical conception of languages 319
- Generalitat de Catalunya* 298
- generative grammar 90, 321, 332, 333
- generative phonology 320
- genitive case 93, 97, 100, 136, 139, 140
- Genoa 301, 304, 305
- Genoese 285, 304
- Gerba 14
- Gérin 390
- German 14, 85, 162, 192, 216, 226, 240, 267, 276,
 281, 282, 284, 287, 300, 306, 307, 314, 315,
 325, 350, 356, 358, 359, 410
- Germania 13
 Lower (*Germania Inferior*) 18, 286
 Upper (*Germania Superior*) 286
 Upper and Lower 9

- Germanic 3, 13, 22, 36, 41, 45, 56, 63, 75, 79, 80,
 81, 82, 88, 134, 167, 168, 169, 176, 192, 202,
 233, 234, 284, 286, 323, 353, 354, 355, 404
 incursions 36, 40, 163
 kingdoms 337
 Germanicus 13
 Germanization 168
 Germans 287, 309
 Germany 174, 279, 316, 359
 Gerov 286
Gesta Berengarii 118
 Getic 22
 Gevgelija 311, 312
 Gheție 256, 261
 Giannelli 197, 198
 Gibbon 64
 Gibraltar 285, 304
 Gifford 31
 Gil Vicente 253
 Gili Gaya 264
 Gilliéron 167, 317, 318, 320, 328, 330, 346
 Gimeno 144
 Gingrich 369
 Giorgini 278
 Giovanardi 241
 Girard 263, 264–265
gîria 403
 Gleick 96
 Gleßgen I, 95, 144, 174, 176
 glides 170, 229
 glossaries 131
 glosses 157
 Glosses of San Millán, *Glosas emilianenses* 33,
 120, 157
 Goa 362
 Godding 102
 Goebel 144, 149, 152, 153, 154, 172, 322, 342
 Gold Rush 369
 Golden Age 107, 131, 249, 262
 Goldiş-Poalelungi 271
 Goldoni 258, 268
 Golembeski 397
 Gómez Redondo 182
 Gonçalves, M. 266
 Gonçalves, P. 377
 Góngora 262
 Gonzaga 244
 González González 349
 González Ollé 182
 González, Count Fernán 181
 Gonzalo de Berceo 159
 Gonzalo Rubio 193
 Gonzzone di Novara 119
 Goodman 422, 425
 Goody 62
 Google 127
 Gopeš 311
gorgia 198, 328
 Gorizia 307
 Goscelin of Saint-Bertain 43
 Gospels 101, 102
 Gossen 151, 152, 153, 172
 Gothic 205, 211–212, 286
 Goths 340
 Gouron 175
 Grado 301
 Grafström 50, 174
 grammar 141
 grammarians 131, 239
 grammaticalization 33, 34, 222, 224, 427, 434,
 435
 grammaticography 174
grămoşteni 311
 Granada 226, 248
 Kingdom of 362
 Granda 190, 236, 425
 Grassi 279, 318, 319
 Gratian 175
 Graubünden, *see* Grisons
 Graur 201, 217
 Graus 101
 Greci 316
 Greco-Roman traders 176
 Greece 38, 290, 311, 312, 315
 Greek 7, 10, 20, 21, 22, 30, 38, 129, 134, 135, 147,
 156, 159, 161, 163, 189, 192, 193, 215, 229–231,
 233, 234, 235, 237, 239, 240, 243, 246, 252,
 255, 257, 262, 271, 280, 281, 284, 285, 307,
 315, 335, 336, 346, 368
 Greeks 308
 Green, J. I, 100, 332
 Greenberg 333
 Grégoire, abbé 273, 295, 351
 Gregory I 78
 Gregory of Tours 39, 40, 47, 70, 71,
 77, 344
 Gregory V 119
 Griffith 241, 257, 267, 268, 277
 Grignani 165
 Grimm, Jacob 403
 Grisons 299, 300, 303, 306, 307, 315, 350, 356
 Gröber II, 405, 406, 409
 Gross 254
 Grosseto 304
 Gruber 174
 Grübl 147

- Guadeloupe 362, 364
 Guanche 194
 Guaraní 190, 236, 357, 358, 373, 374
 Guardia Piemontese 297, 298
guardiolo 298
 Guatemala 361, 374
 Guernsey
 Guey 13
 Guido of Pisa 46
 Guinea 362
 Guinea-Bissau 124, 362, 364, 365, 366
 Guṭu Romalo 280
 Guyane 362
 Guyon 98
 gypsies 359

 /h/ in loanwords 206, 215
 Habsburgs 256, 272, 289, 309, 343
 Hadrian 18, 37
 Hadrian's Wall 18, 335
 Hafner 145, 167
 Hagemeyer, Tjerk 405
 Haiman 254
 Haiti 351, 362, 363
 Haitian, *see* creoles, Haitian
haketi 368
 Hall I, 118, 133, 134, 197, 238, 241, 257, 259, 260,
 263, 269, 273, 275, 402, 412, 414, 422
 Hannick 80
 hard pronunciation of affricates 213
harpeitan, *see* *arpitan*
 Harris-Northall 134
 Harris, M. I, 2, 254
 Harris, T. 367, 368, 372
 Hartmann 55, 106, 153, 162, 166, 171, 173
 Hasdeu 201
 Haskins 159
 Haubrichs 88
 Haudricourt 320
 Haug 80, 88
 Haugen 148, 160, 188, 191
 Hausa 376
 Hausmann 163
Haut Conseil de la Francophonie 394
 Hauteville 321
 Hawaii 426
 Hawkesbury 383
 Hazard 268
 Hebrew 22, 189, 191, 193, 240, 252, 288, 399
 Heidegger 66
 Heilmann 254
 Heinzelmänn 76
 Heitman 341

 Heliade-Rădulescu 271, 280, 281
 Heller 384, 386
 Hennemann 395
 Henri II 246
 Henri IV 258
 Henriquez 409
 Henry of Burgundy 183
 Henry, A. 31
 Herman 5, 34, 58, 60, 63, 77, 80, 86, 89, 91, 92,
 93, 98, 100, 101, 103, 110, 112, 116, 118, 130,
 134, 139, 156, 163, 335, 344
 hermitism 76
 Hernández 265
 Hernández González 250, 265
 Herța 309
 Hesseling 409, 415
 Hibernia 13
 Hilary of Poitiers 69
 Hilty 158, 170, 202, 203, 204, 208, 209
 Hinskens 145, 150
 Hirsch 240
 Hispanic Latin 176
 Hispano-Americans 357
Historiae Francorum 39
 historical dialectometry 144
 'historical language' 146
 historical linguistics 431
 historical pragmatics 144
 historical sociolinguistics 144
 historical-comparative paradigm 402, 408, 412,
 415, 418, 419, 428, 443
 historiography 159, 181, 182
 Hjelmslev 419, 420, 421, 422
 Hochberg 372
 Hodcroft 31
 Hofmann 103
 Hohl 21
 Holm 366, 405, 422, 427, 428, 429, 436, 437, 438,
 439, 440
 Holocaust 290, 399
 Holtus I, 2, 149, 254, 348
Homilies d'Organyà 32, 119, 178
Homilies on Ezekiel 78
Homilies on the Gospels 78
 homogenization 316, 343
 homonymic clash 318, 330
 homonymy 74
 Honduras 361
 Honnorat 274
 Hope 188, 231, 267
 Hordé 263
 horizontal communication 79
 Hotin 309

- Hubschmid 235
 Hucbald de Saint-Amand 170
 Huesca 291
 Hugo, Victor 61, 126
 Huguenots 369
 Hull 381
 Humanism 153, 239, 249, 252, 257
 Humboldt 403, 418
 Hundred Years War 172
 Hungarian 202, 216, 229, 281, 284, 285, 286, 315, 323, 414
 Hungarians 308, 309, 359
 Hungary 281, 309, 310, 315
 Kingdom of 255, 256
 Hunnish 286
 Hunnius 204, 206, 209, 211, 411, 421
 Huns 204, 209
 Hupka 237
 Hurren 222
 hybrid texts 146
 hybridization 387–390
 hymnography 47
 hypercorrection 147, 148, 154, 172
 hypothetical clauses 230

 Iași 261
 Ibeas de Juarre 31
 Iberian 22, 193
 Iberian Peninsula 6, 12, 22, 38, 44, 47, 49, 110, 111, 119, 134, 139, 157, 159, 175, 176, 177, 180, 181, 182, 183, 204, 206, 209, 211, 212, 226, 248, 285, 288, 326, 331, 332, 355, 362
 Ibero-America 431
 Ibero-Romance 26, 109, 113, 118, 120, 121, 135, 137, 141, 162, 167, 173, 176, 186, 200, 203, 208, 226, 228, 236, 285, 287, 332
 western 49
 Icelandic 192
 idealism 319, 320
Idéologues 263
 idiolect 146
 idiom chunk 102
 Île de France 121, 168, 170, 171, 172, 245, 294, 295, 313, 407
 Iliescu 190
 Illinois County 397
 illiteracy 72
 Illyrian 22, 193, 195
 immigration 304, 346, 348, 350, 359
 imperative 216, 375
 imperfect indicative tense 99, 230, 352, 395
 imperfect learning 204, 210
 Imperia 303

 inanimate subject 113
 India 39, 363
 Indian Ocean 363, 397
indios 357
 indirect object 377
 Indo-European 6, 11, 28, 193, 284
 (proto-) 110, 118, 133
Indovinello veronese 163
 Industrial Revolution 368
 industrialization 296, 321
 inequality 72
 infectum 97, 98, 99
 infinitive 97, 233
 inflection 136, 138, 398
 Ingoberga 39
 Ingunde 39
inhumatio ad sanctos 102
innere Sprachform 418
 innovation 144
 inscriptions 200
Institut d'Etudis Occitans 297
Institution de la religion chrestienne 241
Instituto Arias Montano 399
 intelligibility 335, 345
 mutual 337, 342, 344, 432
 interference 147, 148, 189, 382–383
 negative 147
 International Phonetic Alphabet 46
 intonation 86, 106, 128
 'invisible hand' 144
 Ionescu 310
 Iordan 189
 Iraola 291
 Isabel I of Castile 248, 249, 362
 Isidore of Seville 78, 100, 119
 Islam 79, 288, 312, 355, 359
 Isleño 397
 isoglosses 95, 97, 322, 323, 325, 326, 327, 328, 329, 338
 isolated areas 331
 Israel 115, 290, 399
 Istria 282, 288, 304, 306, 308, 313
 Istrian 285, 287, 302, 303
 Istro-Romanian 190, 214, 218, 219, 220, 221, 223, 224, 225, 236, 285, 311, 313
 Istro-Romanians 311
 Italian 1, 7, 15, 33, 49, 51, 63, 113, 121, 124, 129, 134, 137, 139, 140, 141, 162, 163, 164, 166, 167, 178, 203, 205, 209, 210, 211, 218, 233, 237, 238, 243, 244, 246, 256–258, 267–269, 281, 282, 283, 284, 285, 287, 299, 301, 302, 303, 305, 306, 307, 308, 313, 314, 315, 328, 342, 343, 345, 349, 358, 362, 369, 371, 386

- Italian (cont.)
 central 11
 as diplomatic language 245
 dictionaries 256–257
 grammars 243–244, 257
 Gallicisms 267
 lexicography 256
 literary language 242, 245
 middle 123
 old 121
 orthography 243
 proto-Italian 84, 87
 regional varieties 307, 313, 314
 spoken 244
 spread of standard language 276–280
 written language 277
- Italian dialects 43, 49, 162, 163, 166, 204, 206, 209, 211, 244, 245, 269–270, 277, 279, 303, 313, 358; *see also* Italo-Romance; Campanian; Emilian; Lombard; Marchigiano; Sicilian; Piedmontese; Pugliese; Romagnol; Venetian
- central 26
 central and southern 302
 dialectology 279, 280
 dictionaries 270, 279
 grammars 243
 literary traditions 258
 northern 26, 203, 208, 302, 334
 southern 33, 339
- italiano popolare* 369
- Italians 49, 306, 316
- Italic 17, 22, 26, 193
- italo-americano* 369
- Italo-Romance 26, 33, 47, 56, 162, 167, 202, 208, 227, 303, 315, 331, 332, 342, 345
 northern 11
 southern 11, 118, 119, 120, 121, 122
- Italy / Italian Peninsula 6, 7, 8, 9, 11, 13, 29, 56, 78, 84, 87, 106, 107, 110, 111, 112, 115, 118, 119, 121, 122, 123, 131, 159, 162, 163, 164, 165, 166, 174, 175, 176, 202, 208, 230, 240, 241, 243, 245, 250, 259, 270, 271, 283, 284, 297, 298, 299, 302, 305, 306, 307, 313, 314, 315, 316, 317, 323, 328, 332, 340, 351, 352, 358, 359, 370
 education 279
 Italianization of Peninsula 276–280
 Kingdom of 301
 linguistic minorities 277
 north 26, 33
 Republic of 306
 south 20, 33, 225, 315
 Unification of 277, 302, 303, 306
- use of French 268–269
- IULIUS GEDOMO (C.) 13
- IULIUS OTAVANEVVS (C.) 14
- IULIUS RUFUS (C.) 13
- Ivănescu 201, 217, 312
- Ive 282
- Izzo 197
- Jaberg 318, 319
- Jackendoff 91
- Jacobinism 145
- Jakobson 187, 191
- James, E. 354
- Janson, T. 85, 115, 134
- Japan 362
- jargon* 403
- Jaume I 179
- jerga* 403
- Jerome 69, 70
- jers 215
- Jesuits 250, 271, 374
- Jews 120, 182, 213, 248, 290; *see also* Sephardic communities
- Jireček 286
 Line 287
- Joan 250, 251, 266, 275
- João III 252
- João VI 373
- Jocs Florals* 275
- Jodl 163
- Jolivet 88
- Joly 263
- Jones, M. 7, 63, 295
- Jones, V. 258
- Jordanova 107
- Joseph, J. 237
- joual* 366
- Joubert 409, 442
- Jourdain 434
- Jouveau 274
- Judæa 18
- judezmo*, *see* Ladino
- Juilland 320
- Julia 273
- Juneau 395
- Jungmann 101
- Jura (Canton) 294
- jurisprudence 159
- Justinian 34, 175
- Kabatek 45, 49, 115, 118, 144, 146, 147, 148, 149, 151, 159, 161, 175, 180, 181, 182, 184

- Kabuverdianu, *see* creoles, Cape Verde
 Kahane, H. 229
 Kahane, R. 229
 Kahl 311, 312, 313
 Kairun 16
 Kant 66
 Kany 375, 391
 Karlsson 146
 Kato 437, 438, 439
 Kattenbusch 341
 Kaufman 188, 191, 379, 402, 427, 430, 431
 Kay 274
 Keller 144
 Kerswill 145, 150
 Kiesler 54
 Kihm 421, 427, 439
 Kimbundu 376, 378
 King 282, 379–382, 390, 397
 Kitts 265
 Kiwitt 367
 Klausenburger 93
 Klee 374
 Kloss 148, 156, 160, 161
 Knights of St John 173
 Knights Templar 158, 173
 Koch 60, 90, 101, 106, 146, 148, 149, 156, 157, 163, 165, 241, 417, 421
koinè padana 165
 koiné, koineization 143, 147, 151, 158, 161, 167, 173, 297, 301, 372, 400, 432–433
 Kontzi 156, 168
 Kopitar 201
 Körner 149
 Kovačec 218, 219, 220, 222, 224
 Kramer 86, 91, 342
 Krefeld 124, 156, 163, 359
 Kremnitz 274
 Kriyol (Guinea-Bissau) 414, 421, 427, 438
 Krk 282, 285, 308, 313
 Krusch 39, 40, 47, 69
 Kukenheim 239
 Kurds 359

 L'Aquila 302
 L'Isle Madame 395
 /l/ > /r/ 201
 Li 191, 359, 360
 L2 191, 360
La Decadència 250
 La Maddalena 304
 La Ramée 247
 La Rioja 291
 La Spezia–Rimini Line 29, 328, 334

 Labov 59, 68, 92, 144, 154, 187, 320
 Ladin 26, 120, 162, 204, 209, 254, 276, 284, 285, 287, 300, 307, 314, 317, 323, 325, 358
ladin dolomitan 300, 341, 348
ladin standard 300
 Ladino 120, 189, 290, 367; *see also* Spanish, Judeo-Spanish
 Ladins 314, 315
 Lafage 376, 377
 Lafont 274
 Lagueux 372
 Laliena 79
 Lancelot 261
 landscape 194
 Langobard, Langobardic 163, 164, 204, 207, 210–211
 Langobards 211
 Langres 37
 language change 401, 402, 408, 418, 427, 429, 430, 431, 443, 444
 language death 222, 370, 396–399
 language loyalty 191
 language mixture 418, 419, 433
 language names 151
 language planning 148
 language shift 430
 langue 319
langue d'oc, *see* Occitan
langue d'oïl 45, 47, 55, 245, 246, 313, 340; *see also* French
 Languedoc 50–50
 Languedocian, *languedocien* 102, 173, 298
 Laon 170
 Lapesa 181, 228, 238, 249, 262, 265
 laryngeals 227
Las razós de trobar 174
 Lass 133, 134, 137
 Late Antiquity 65, 67, 72, 75, 79, 89, 135
 lateral vs. central areas 330
 Lathuillère 260
 Latin 3, 4–5, 7, 8, 10, 11, 13, 14, 15, 16, 19, 20, 21, 22, 23, 24, 25, 29, 30, 33, 38, 41, 47, 49, 52, 53, 54, 55, 57, 59, 60, 61, 62, 63, 64, 65, 66, 68, 71, 72, 73, 74, 75, 77, 78, 81, 82, 83, 84, 85, 86, 88, 89, 90, 91, 92, 93, 94, 95, 98, 104, 108, 110, 112, 113, 114, 115, 116, 117, 118, 119, 120, 126, 129, 130, 131, 133, 134, 135, 137, 138, 140, 147, 156, 157, 158, 159, 160, 161, 163, 164, 165, 166, 169, 173, 175, 176, 178, 179, 182, 188, 189, 190, 193, 214, 232, 233, 236, 238, 239, 240, 244, 246, 247, 250, 251, 252, 253, 255, 257, 258, 267, 285, 286, 295, 303, 306, 320, 335, 336, 339, 340, 344, 350, 354, 361, 368, 390, 391, 403, 420, 442

- Latin (cont.)
 African accent 21
 (of) Aquitaine 26
 Christian 105
 Classical 5, 32, 53, 55, 252
 diachronic variation 25
 diastatic variation 25, 42
 diatopic variation 7, 25, 26, 27, 42, 54
 expansion 8, 21
 (of) Gallic countryside 26
 inscriptions 25
 late 5, 26, 27, 46, 48
 late Latin regional variation 26, 27
 learnèd language 238–240, 257, 270
 literary norm 24, 54
 manuscripts 25
 medieval 33, 46, 47, 48, 119, 120, 239
 normative 43, 52
 orthography 4, 10, 45–51
 prestige 239
 roof language 7
 spelling of [w] 4
 spoken 25, 28, 35, 46, 47, 52, 53, 54
 substandard 24–27, 35, 36, 54, 55
 sub-substandard 27–36
 teaching of Latin grammar 271
 uneducated 10
 vulgar 5, 24, 27, 403, 404, 408; *see also* Vulgar Latin; ‘Leonese Vulgar Latin’
 written 35, 46
 Latin America 6, 63, 236, 317, 342, 357, 360, 378, 391
 Latina 304
 Latinity 64, 82, 83
 Latinization 7, 8, 10, 12–21
 Latinos 346
 latinum circa romançum 47, 345
 Latium 361
 Lausberg 1, 29, 32, 156, 317, 320
 Lausberg Zone 331
 Law 83; *see also* legal
 Lazard 106
 Lázaro Carreter 264, 265
 Lazcano 239
 Lazio 305
 Le Goff 101
 Le Jan 65
 Le Jars de Gournay, Marie 260
Le Monde 47
 Le Page 407, 415, 417, 428, 433, 442
 learned influence 188
 Lebanon 367, 394
 Lebsanft 144
 Lecce 230
 Ledgeyway 6, 139, 230, 371, 378, 439
 Ledoix-Serrigny 34
 Lefebvre 379, 428
 Lega Grigione 300
 legal
 documents 158, 159, 163, 164, 166, 167, 179, 182
 terminology 150, 183
 length 339
lengua castellana, *see* Spanish
lengua española, *see* Spanish
lenguas generales 350, 358
 lenition 28–30, 35, 47, 194, 328, 333, 334
 Lentner 73
 Lenz 405, 413
 León 41, 44, 115, 122, 137, 141, 288, 290, 291, 292, 326
 Kingdom of 176, 181, 184
 Leonese 44, 121, 132, 176, 177, 180, 181, 182, 248, 291, 345, 347
 ‘Leonese Vulgar Latin’ 177
 Lepschy, A.-L. 245
 Lepschy, G. 237, 245
 Leptis Magna 21
 Lesser Antilles 407; *see also* creoles; Lesser Antilles French
 Leto, Pomponio 239
 Lévi-Strauss 62
 Levison 39, 40
 Lewitt-Gibbon 65
 lexeme 97
 lexical diffusion 111
 lexicography 254, 256
 lexicon, *see* lexis
Lexikon der romanistischen Linguistik 348
 lexis 24, 27, 86, 185, 259, 319, 332, 369; *see also* loanwords
 lexical variation (in Latin) 26, 27
 ‘ley Morano’ 289
Leys d’amors 174
 liaison 338
Liber Historiae Francorum 344
 Librandi 164, 241
 Libri 278
Libri di banchieri fiorentini 164
Libri Historiarum 70, 77
Libro de buen amor 51
 Liburnio, Nicolò 244
 Libya, Libyan 13, 14, 16, 17, 22
 Libyan-Berber 13, 22
 Libyans 354
 Lidzbarski 14, 15

- Liébana 31
 Liège 45
 Lightfoot 108
 Liguria 305, 317
 Ligurian
 ancient 6, 22, 193, 195
 modern 201, 285, 295, 303, 304
 Ligurians 304
 Lima 392
 Lima, Luís Caetano de 266
Limba Sarda Comuna 302
 Limousin, *limousin* 102, 173, 298
 Lingala 385
lingua cortegiana, *lingua cortigiana* 165, 241
lingua franca 308, 424, 425
lingua geral 374
lingua lombarda 164, 165
lingua mixta 78, 100
linguagem dos musseques 365
 linguistic atlases 154
 linguistic geography 317–329
 linguistic islands 236, 323, 326
 Lipski 364, 372, 377, 378, 398, 436
 Lisbon 158, 184, 185, 251, 292; *see also*
 Portuguese
 Lissus 286
 literacy 92, 240, 244, 252
 liturgy 101
 Liver 162, 254
 Lleal 177
 Lleida 44, 178
 Llera 291
 Llull 179
Lo codi 175
 loanwords 114, 232, 286, 320, 378
 Gallicisms 23, 36, 265, 267, 269
 Greek neologisms 271
 Lodge 121, 148, 167, 170, 259, 273
 Löfstedt, B. 100
 Löfstedt, E. 68, 93, 103
 logographic reading and writing 45–51, 52, 132
 Logudorese 167, 270, 301, 302, 329
 Logudoro 32
loi Deixonne 296
 Loire 47, 56, 372
 Lollius Urbicus (Q.) 17, 18, 37
 Lombard 284, 303, 308, 323, 328, 342
 Lombardi, L. 244
 Lombards 44, 78, 87
 Lombardy 134, 205, 210, 304, 305, 317, 325, 342
 Lope Blanch 236
 López García 112
 López Morales 265
 Loporcaro 29
 Lorenzino 429
 Lorenzo 183
 Lorrain, *lorrain* 169, 294, 315
 Lorraine 171, 206, 294, 295, 315
 Los Angeles 346, 383
Los Muestrros 399
 Lot 53, 65
 Lotharingian 315
 Louis IX 295
 Louis XIII 258
 Louis XIV 263
 Louisiana 363, 364, 368, 372, 378, 381, 389, 396, 397
 Loyola 248
 Luanda 365, 377
 Lubin 219
 Lucan 66
 Lucania 6, 12, 331
 Lucca 31
 Lucchi 271
 Lüdi 350, 394, 395
 Lüdtke 83, 85, 90, 101, 115, 156, 157, 158, 168, 250,
 403
 Lugdunum 168
 Lumetti 276
 Lurçat 96
 Lusignan 43
 Lusitanian 22
 Luther 240
 Luti 278
 Lutz 254
 Luxembourg 39, 295
 Grand Duchy, 296
 Lynch 374
 Lyon (Lugdunum) 13, 39, 299
 Lyonnais 169

 /m/ (word-final) in Latin 338
 Ma'a 389
 Macau 362
 Macedo 365
 Macedo-Romanian(s), *see* Aromanian,
 Aromanians
 Macedonia 9, 214, 286, 311, 312, 315
 Macedonian 214, 225, 233
 Machado 269
 Mackey 188, 191
 Mactar 15, 16, 17
 Madagascar 362, 390, 398
 Madao 270
 Madaura 20
 Madeira 253, 363
 Madrid 183, 317

- Maffei 268
 Magalhães de Gândavo 252
 Magellan 363
 Maggi 258
 Maghreb 316
 Magie 21
 Magna Curia 165
Magna Graecia 230
 Magra–Rubicon line 163
 Maia 183, 184, 185, 275
 Maiden 121, 124, 202, 208, 222, 333
 Maine (USA) 368
 Mainz 286, 299
 Maior 271
 Mair 348
 Majorcan 278
 Malato 258
 Malayan 419
 Malerbi 241
 Malherbe 66, 258, 259, 260
 Mali 362, 377
 Malkiel 6, 8, 233, 236
 Malmberg 236
 Malovište 311
 Malta 308, 343
 Maltese 308, 415
 Mambelli 257
 Mandouze 74
 Manila 423
 Manitoba 388, 389
 Mannessey 92
 Mansceaux 43
 Manuel I 253
 Manuzio 239
 Manzoni 278
 maps, linguistic, *see* atlases (linguistic)
 Maramureş 213, 348
 Marano Lagunare 301
 Maraschio 242
 Marazzini 241, 258, 268
 Marcato, C. 254
 Marcato, G. 301
 Marcellesi 342
 Marche 305
 Marchello-Nizia 103
 Marchigiano 328
 Marco Polo 345
 Marcos Marín 124
 Marghine 329
 Maria of Aragon 253
 Marie de France 159
 Mariño 183
 Marques 251, 252, 253
 Márquez-Sterling 181
 Marrou 68, 274
 Marsala 21
 Marseille 22, 47, 316
 Martelli 243
 Martí i Castell 275
 Martin 181
 Martín Abad 240
 Martin the Human 179
 Martín Zorraquino 177
 Martin, J.-B. 245
 Martinell 269
 Martinet 187, 320, 321
 Martinique 362
 Martino 12
 Martins 183
 Martinus 409, 438, 442
 mass communication 76
 Massa 238, 302
 Massachusetts 368, 396, 397, 398
 Massignon 381
 Massobrio 322
 Matarrese 267
 Matthews 123
 Mattoso Câmara 275
 Maurer 413, 414, 421, 427, 439, 441, 442
 Mauretania 9
Mauretania
 Caesariensis 287
 Sitifensis 287
 Tingitana 287
 Mauritius 364
 Maximus Thrax 17
 Mayan 374
 Mayotte 362
 Mazzarella Farao 270
 -mb- > -mm-, *see* weakening, of postnasal
 consonants
 McKitterick 80, 81, 83, 169
 McLaughlin 208
 McWhorter 429, 436
 measure phrases 220, 221
 media 296, 343
 Media Lengua 388
 Mediterranean 8, 40, 79, 177, 250, 367
 Meglen 312
 Megleno-Romanian 214, 225–226, 233, 254, 287,
 311, 312
 Megleno-Romanians 309, 311, 312, 315
 Meier 176
 Meigret 247
 Meillet 188, 191
 Meisterfeld 175

- Meli 258
 Melillo 298
 Melo 253
 Mendes Drumond Braga 253, 267
 Menéndez Pidal 31, 33, 47, 50, 84, 120, 145, 152,
 177, 180, 181, 182, 183, 326, 345
 Menorcan, *see* Catalan
 MENTE 225
 Meo Zilio 387
 Mérida 29
 Merlo 168, 197
 Merovingian(s) 69, 70, 71, 80, 82, 83, 84, 88, 105,
 134, 169, 170, 204, 210, 337
 charters 30
 mesoclisys 33
 Mesopotamia 38, 193
 Messapic 6, 22, 193, 195
 Messina 22
 Messner 124
mestizos 424
 metaphor 320
 Metastasio 267
 Métis 388, 389
 metonymy 320
 Metzeltin 1, 2, 277
 Meuse 56, 286
 Mexico 236, 361, 364, 374, 387, 391
 Mexico City 384
 Meyer-Lübke 1, 214, 217, 317
 Meyer, B. 276
 Michaelis 405, 414, 435
 Michif 388–390
 Micu 271, 272
 Middle Ages 7, 44, 45, 46, 49, 51, 53, 55, 64, 75,
 84, 89, 107, 114, 116, 127, 130, 131, 154, 160,
 161, 162, 177, 178, 180, 183, 185, 237, 258
 Miestamo 146
 Migliorini 164, 241, 257, 267, 268, 277
 migration 63, 296, 316–317, 343; *see also*
 emigration; immigration
 Mihăescu 229, 286
 Milan 41, 258, 346
 military lexicon 205, 210, 212
 Millar 17
 Miller 62
 Millet-Gérard 80
 Milroy, J. 111, 417
 Milroy, L. 111
 Minerva 264, 267
 Mingas 365, 378
 Minho 183, 184, 291, 342
 Mioni 149
 Miracle 275
 Miranda, G. 249
 Mirás 276
Mirèio 274
 Miret i Sans 32
 Miron 262
 missionaries 357
 Missouri 397
 Mistral 274, 297
 Mitterrand 394
 Mixteco 374
 Mochica 358
 Modena 258
modernismo 393
 Moesia 286, 353
Moesia Inferior 286
Moesia Superior 286
 Mohács, Battle of 255
 Moldavia, *see* Moldova
 Moldova 213, 228, 254, 255, 261, 271, 281
 Principality of 308, 309, 310
 Republic of 218, 284, 306, 309, 310, 315, 316,
 341, 359
 Soviet Socialist Republic of 310, 315, 341
 Moldovan dialects 310
 Moldovan language 282, 310, 341, 359
 Moldovans 317
 Molière 126, 260
 Moluccas 363, 423
 Monaci 32, 33, 51
 Monaco 296, 303
 monarchy 172
 monasteries 150, 164, 171, 177, 180, 184, 185
 monasticism 75
 Monco 366
 Moncton 390
 Monfalcone 301
 Monfrin 155
 Mongin 318
 Mongols 308
 Monjour 49, 183, 184, 185
 monogenesis theory 422–426, 428
 monolingualism 86
 monosyllabism 320
 Montagnais 389
 Montaigne 260
 Montana 388
 Montanyés 290
 Montecassino 46, 47, 48, 164, 166
 Montemor 253
 Monti, P. 279
 Montoliu 275
 Montpellier 158
 Montréal 366, 382, 385

- Moorish 226
 Moors 182
 Moos 53
 Morala 132, 154
Moralia in Job 78
 Moran 178, 275
 Morf 44, 168
 Morgana 267
 Morlachs 313
 Mørland 33
 Morocco 367, 385
 Morosi 298
 morpheme 94, 97, 114, 420
 morphology 52, 86, 94, 97, 98, 102, 110, 113, 115,
 136, 138, 139, 141, 163, 185, 191, 230, 236, 239,
 240, 332, 418; *see also* borrowing
 change in 92
 derivational 114, 205, 211–212, 215, 233, 369,
 388, 395
 inflectional 194, 207, 216, 235, 236, 352
 nominal 112
 verbal 112
 morphosyntax 94, 112, 115
 Morpurgo Davies 406
 Mortara Garavelli 243
 Mossé 88
 Mougeon 378, 379, 382–383, 395, 398
 movement rules 426
 Mozambique 316, 362, 365, 377
 Mozarabic 80, 84, 285, 288
 Mozarabs 180, 182, 184, 185, 226
 Mozia 21
 Mufwene 402, 406, 407, 413, 422, 425, 427, 430,
 432, 433, 436, 437
 Muggia 301
 Mühlhäusler 402, 427
 Muljačić 189, 237, 282
 Müller 62, 336, 352
Mulomedicina Chironis 26
 multilingualism 24, 159, 187, 248, 313–316, 430,
 432
Mundart 411
 Munteanu 438
 Muntenia 254, 308, 348
 Muntenian 214, 254
 Muntzel 397
 Murcia 177, 285
 Murciano 110
 Murecine Tablets 29
 Muslims 119, 126
 mutation 144
muwaššahas 288
 Muysken 189, 388, 422, 427, 428
 Myers-Scotton 379, 384, 385, 389
 /n/ > /r/ 201
 Nadal 250
 Nahuatl 358, 374
 Nandriş 217
 Naples 22, 33, 258, 269
 Napoleon 273
 Napoleonic Wars 373
 Naro 425, 426, 436
 nasalization 405
 Navarre 176, 196, 291
 Navarrese 176, 180
 Navarro-Aragonese 177, 291
 Navia 326
 -ND- > -nn-, *see* weakening, of postnasal
 consonants
 Neacşu de Câmpulung 254
 Neagoe 214
 Neapolitan 121, 269, 270, 304, 346, 347
 Nebrija 131, 249
 negation 215, 414, 434, 435
 Nehama 399
 Neiescu 221
 neo-Romanized areas 287
 Neogrammarians 193, 318, 404, 410, 411, 412,
 418
 Nerva 38
 Neto 184, 238, 251, 275
 Neumann-Holzschuh 249, 366, 434, 437, 444
 Neustria 204, 210
 neuter 35, 100, 105, 112, 217
 neutralization 218, 219
 New Brunswick 362, 363, 372, 380, 390
 New Caledonia 362
 New England 368
 New France 368
 New Hampshire 368
 New Mexico 364, 384
 New Orleans 364
 'New Philology' 144
 New Testament 29, 240, 241, 254, 261, 408
 New York 346, 383
 Newfoundland 372, 380
 Nicaragua 361
 Niccolini, G. 278
 Nice 22
 Niculescu 215, 256, 262, 287, 333
 Niederehe 183, 248, 249, 250, 265
 Nielsen 124
 Nièvre 37
 Niger 362
 Nîmes 41

- Niño-Murcia 375
 Nisibis 34
 Nithard 118, 169
 Nocera 32
 nominative case 89, 139
 non-linear dynamic system 96
 Norberg 77, 85, 118
 Nord-Pas de Calais 315
 Norden 73
 Noricum 9, 42, 286, 354
 'norma culta' 374
 Norman England 56
 Norman French 169, 172, 205, 211, 220, 330, 347
 Anglo-Norman 44, 55
 Normandy 52, 152, 168, 171, 172, 294
 Normans 43, 205, 210, 226
 Norsemen 203, 208
 North America 364, 396
 North Bay 383
 North Carolina 369
 North Dakota 388
 Norton 240
 Notia 312
 noun phrase 419–420
 Nova Scotia 363, 372, 380
 Novati 119
 null subject parameter 437, 439
 number 163, 218
 numerals 218, 220, 221, 232
 Numidia 9, 15, 18, 287
 Nunes do Leão 252
 Nuorese 301

 oblique case 139
 Occam's razor 63
 Occitan 7, 11, 33, 56, 99, 117, 119, 120, 121, 147,
 157, 167, 168, 169, 173, 174, 176, 177, 180,
 185, 189, 199, 203–210, 211, 226, 227, 245,
 246, 274–275, 283, 285, 293, 294, 295, 296,
 297, 298, 313, 314, 316, 330, 331, 332, 341,
 347; *see also* Gascon
 old 33, 50, 250
 Provençal 165, 173, 183, 185, 298
 restandardization 274
 Occitania 55, 56
 Occitanist 176
 occupations 194
 Odessa 310
 Odoacer 211
 Oea 20
 Oesterreicher 60, 90, 91, 92, 106, 125, 145, 146,
 148, 149, 158, 160, 167, 403
 Office québécois de la langue française 379, 395

 Ohrid, Lake 311
 Oil 287, 293, 294, 295, 296, 297, 299, 313
 Oise 31
 Okeju 379
 Old High German 49
 Old Testament 241
 Oliva, F. 269
 Olivar 72
 Oliveira, F. 252
 Olivença 251
 on 206, 211, 232, 352
 Oña 31
 Ong 62
 onomasiology 215, 320
 onomastics 13
 Ontario 378, 380, 381, 382–383, 395
 Onu 255, 261, 262, 271
 Orbanic 313
 Organisation internationale de la Francophonie
 394
 Oribasius 26
 Oristano 304
 Orlando 369
 Orlando furioso 242
 Orlando innamorato 242
 orléanais 293
 Orléans 170, 245
 Orthodox Church, *see* Romanians, Orthodox
 orthography 45–51, 115, 117, 120, 141, 160, 172,
 180, 185, 206, 240, 243, 247–248, 252,
 265, 270, 274, 342, 352, 407, 408; *see also*
 spelling
 reform of 252
 Oscan 6, 22
 Osco-Umbrian 195
 Ostrogoths 78, 206, 211
 Otheguy 383
 Otranto 230
 Ottoman Empire, Ottomans 228, 255, 256, 267,
 268, 271, 290, 309, 367
 Ottonian 87
 Oudin, César 249
 Oudin, Charles 261
 Ourique 184
 Overbeke 192

 Paccagnella 258, 268
 Paciaudi 271
 Pacific Ocean 363
 Padanian koinè 242
 Paden 174
 Padley 239, 261
 Padua 31

- palaeo-Sardinian 193, 195
 palatalization 35, 194, 198, 218, 324, 326, 339, 372
 Romance graphies 49
 palatals 324
 Palenquero 421, 438
 paleographic reform 48
 paleographical analysis 155
 Palermo 21, 258
 Palestine 193
 Pallars 19
 Panama 361
 Pannonia 29, 40, 206, 211, 229, 286, 353, 354
 papacy 304, 337
 Papal court 118
 Papanti 280
 Papen 389
 Papiamentu, *see* creoles
 parachuting 325
 Paraguay 190, 357, 358, 361, 363, 373, 374
 parataxis 398
 Pariente 261
 Paris 41, 47, 66, 100, 141, 158, 160, 168, 170, 173,
 245, 246, 295, 346, 347, 351
 Parisian 122
 Parma 270
 Parodi 257
 Pascal 66
 Pasqualino 270
passato remoto 230
passé simple 352
 passive 47, 97, 98, 100, 113, 344, 352, 377
 new Romance formation 35, 52
 past anterior 352
patois 283, 285, 293, 296, 299, 317, 328, 406, 407
 Patota 244
 Patriarchi 270
 Pau 174
 Paul IV 241
 Paul the Deacon 46, 80, 81
 Paul, H. 410, 411, 412, 416, 417, 443
 Pavia 46, 158
 Pays de Caux 294
 Pazos 250, 251, 266, 275
 Pechenegs 308
 Pedro 137
 Pei 69, 93, 100
 Peitgen 94, 106
 Pellandra 267
 Pellegrini 231, 300, 322, 328
 Pembroke (Ontario) 383
 Penny 107, 108, 111, 176, 177, 180, 196, 238, 290,
 367
 pentaglossia 350
 Pepin 69
 Percoto 276
 Percyvall 249
 Perea 322
Peregrinatio (Aetheriae) 26, 65, 68, 113
 Pérez González 84
 perfectum 98, 99
 Périn 76
 periphrastic structures 320, 352
 past 35
 Perl 366, 378
 Perotti 239
 Perrot 390
 Persia 38, 40
 person 98, 352
 personal names 134, 212
 Pertusi 239
 Peru 357, 361, 373, 374
 Peter III 179
 Petersen 17
 Petrarch 165, 242, 306
 Petrovici 220
 Petrucci, A. 157
 Petrucci, L. 55
 Petrucci, P. 216, 217, 255
 Pfänder 421, 427, 443
 Pfister 36, 45, 170, 174, 175, 176
 Phanariots 228, 271, 280
 Philip II 250
 Philip II Augustus 171, 345
 Philip V 266
 Philip the Fair 295
 Philippide 201, 286
 Philippines 289, 361, 363, 423
 phonetic change 132
 phonetic erosion 318, 330
 phonetic laws, *see* sound laws
 phonetic mutilation 320
 phonetics 52, 110, 111, 185
 phonographic writing 132
 phonological adaptation 227
 phonology 206, 235, 369, 389; *see also*
 borrowing
 Pi de Cabanyes 275
 Piacenza 270
Pianta della Madonna 164
 Picard 79, 169, 170, 172, 204, 209, 294, 295, 296,
 330, 345, 347
 Picardy 170, 171, 172, 206, 294
 Picenian 6
 pidgins, pidginization 400, 401, 402, 412, 421,
 426
 English 422

- nativized pidgin 426
- non-native pidgins 422
- origin of term 402
- pidgin-creole life cycle 429–431
- Portuguese 422, 423, 424, 425, 426
- Spanish 423
- Piedmont 268, 284, 298, 299, 304, 305, 317, 325
- Piedmontese 268, 269, 270
- Pierrard 321
- Piètri 73, 76, 79
- Pilotus 247
- Pindus 311
- Pinheiro Freire da Cunha 266
- Pinker 62, 92
- Pinkster 103, 139
- Pipino 269, 270
- Pirona, Jacopo 276
- Pisa 301, 304, 305, 348
- Pisans 301
- pitch 128
- Pitz 88, 355
- placenames 134, 195, 197, 200, 204, 205, 209, 211, 212, 286, 302, 355
- Placiti cassinesi* 30, 45, 118, 119, 146, 157, 163
- Placitum capuanum* 163
- Plangg 254
- plants 194, 195
- Plautus 24, 113
- Pléiade* 246, 260
- Ploog 399
- pluperfect 99
- plural 219, 235, 324, 333, 334
- Po Plain 338, 347
- Poema de mio Cid* 157
- poetry 94, 126
- Poggi Salani 244
- Poghirc 201
- Pohl 65, 80
- Poitevin 168, 169
- poitevin-saintongeais* 294
- Poitevins 43
- Poitiers 104
- Poles 339
- Polish 213
- political fragmentation 84
- polymorphism 95, 102, 105, 179
- polysemy 318
- Pombal 374
- Pommeau 96
- Pompei 24, 29
 - inscriptions 24
- Pons Rodríguez 177
- Pontes 439
- Poplack 384, 387
- popular etymology, *see* ‘folk etymology’
- Pordenone 325
- Porru 270
- Port Royal Grammar 257, 261, 263
- Porta, C. 258
- Portugal 47, 56, 115, 137, 177, 184, 240, 245, 251, 288, 291, 292, 316, 317
- Portuguese 1, 7, 11, 33, 47, 49, 51, 63, 99, 107, 110, 114, 121, 124, 129, 137, 141, 150, 177, 183, 186, 190, 191, 194, 211, 212, 227, 238, 251–254, 266–267, 275, 276, 283, 288, 291, 342, 348, 358, 360, 361, 362, 363, 365–366, 367, 370, 372, 373–376, 377, 387, 392–393, 401, 404, 405, 408, 413, 414, 423, 424, 425, 442
 - African 378
 - Asian 424
 - of Angola 366, 405, 431
 - Brazilian 343, 371, 372, 378, 437, 439, 441
 - Brazilian Vernacular Portuguese 401, 436, 437, 438, 440
 - Céara 440
 - corrupt 402, 425
 - European 343, 371–372, 373, 438, 439, 440
 - grammars 252–253
 - Gulf of Guinea 427
 - Helevética 440
 - Lisbon 251
 - literacy 252
 - middle Portuguese 123
 - Mozambique 360
 - null arguments in Brazilian Portuguese 439
 - orthography 252, 267
 - São Paulo 405, 439
 - standardization 267
 - West African 424, 438
 - written language 251
- positivism 319
- Posner 1, 276, 407, 420, 433, 434, 437
- possessives 392
- Pottier 358
- Pountain 7, 30, 63, 130, 160, 384, 386, 387, 442
- power relationships 189
- Pozzi 243
- pragmatics 104
- Prats 250
- pre-Roman languages 10, 13, 16, 19, 21–24, 42
- prefixes 234
- preposition stranding 380–381, 397
- prepositions 105, 215, 227, 230, 233, 366, 380–381, 395

- Presbyterianism 369
 present tense 74, 98
 prestige 189, 203, 209, 305, 317, 329, 344, 345, 358, 359
 preterite 74, 96, 98, 138, 141
 'pretoguès' 365
 priests 76
 Primo de Rivera 289
 Prince 379
 Prince Edward Island 363, 372, 379–382
 Principato 247
 Principense 366
 printing 153, 160, 240, 256, 392
Privilegio logudorese 166
 pro-drop 203, 209
 proclisis 371, 378
Proconsularis Zeugitana 287
 progressive/habitual aspect 421
Promessi sposi 278
 pronouns 136, 139, 366, 371, 383, 392
 conjunctive 397
 disjunctive 397
 subject 203, 204, 209
 pronunciation 92, 118, 131, 180, 395
Prodocimi 7
Prose della volgar lingua 242
 prose text 175
 Protestantism, Protestants 299, 369
 proto- (for proto-varieties of individual Romance languages, see the entries for the respective languages)
 proto-Romance 132, 133, 134, 156, 177
 Proust 62
 Provençal, *see* Occitan
 Provence 119, 121, 316
 Prunay 264
 Prüßmann-Zemper 352
 Prut 341
 Pseudo-Apuleius 200
 Puerto Rico 361, 364; *see also* Spanish, Puerto Rican
 Puglia 6, 230, 304
 Pugliese 298, 299, 304
 pull-chain 123
pumontincu 305
 punctuation 243
 Punic 14, 15, 16, 20, 21, 22, 193
 neo-Punic 15
 Punic War
 First 8
 Third 9
 Puoti 268
 Puppo 267
 Puquina 358
 purism 191, 268
 Puşcariu 201
 push-chain 123
 Putte 407
 Pyrenees 19, 177, 284, 291, 297
Pyénées Orientales 293
 Pyrrhus, King 8
 Quadrilateral 309, 311
 Quarmaro 304
 Quebec 6, 362, 363, 364, 368, 389, 394, 395, 434; *see also* French, Canadian, Quebec (*québécois*)
 Quechua 190, 236, 357, 373, 374–376, 388
 Quercy 174
questione della lingua 131, 162, 165, 166, 241–245, 251, 267, 277
 Quevedo 262
 Quilis 124, 157, 177, 182
 Quondam 244
 /r/ in French 346
 Rab 313
 Racan 259
 Radtke 321, 322
 Raetia 9, 202, 208, 286
 Rætic 6, 193
 Ræto-Romance 139, 161, 226, 254, 284, 317, 350, 356
 Ragusa 162, 282, 308, 355; *see also* Dubrovnik
 Ragusan 162, 308
 Raible 157, 158, 160, 163
 Raimbaut de Vaqueyras 120, 142
 Raimon Vidal 174
 Raimundo 184
 Ramajo Caño 249
 Ramus, *see* La Ramée
 Raposo 439
 Rastier 91
 Ravenna 134, 163, 211
 Ravier 106
 Raynaud 151
 Raynouard 317
 Razós 174
 re-Romanization 287, 288
Real Academia de Buenas Letras 266
Real Academia de la Lengua 122
Real Academia Española 127, 264, 265, 342, 390
 Récatas 191
 Reconquest, *Reconquista* 180, 183, 226, 248, 288, 289, 356, 357
 in Sicily 356

- reconstruction 132, 133
- reducciones* 374
- reflexives 99, 113, 136, 352
- reformatio in melius* 69, 83
- Reformation 240, 254, 256
- Reformed church 350
- regional languages 284
- register 76, 105
- Régnier-Desmarais 263
- Reichenkron 62, 201
- Reinecke 436, 437
- relativization 420, 426, 433, 434, 437–440, 441, 442
- relexification hypothesis 426, 428, 429
- Remacle 45, 151, 155, 169, 172
- Remada 14
- Remarques [...] sur la langue françoise*, see Vaugelas
- Renaissance 107, 165, 239, 241, 245, 257, 368, 403
- Renaixença* 122, 275, 293
- Renzi 6, 122, 157, 272, 282, 332, 361
- Res gestae* 38
- Restaut 263, 266
- Reuillon-Blanquet 264
- Réunion 362, 364
- Revel 273
- Rexurdimento* 276
- Reyes Católicos*, see Catholic Monarchs
- Rézeau 395
- Rhine 13, 18, 37, 38, 40, 254, 335, 353, 355
 - Valley 202, 208, 286
- Rhineland 22
- Rhône Valley 274
- rhotacism 201, 234
- rhythm 128
- Richardson 240, 241, 242, 243, 244, 245, 246
- Riché 75, 76
- Richelet 264
- Richelieu 259
- Richter 29, 58, 94, 103, 106
- Rickard 239, 245, 273, 275
- Rickford 440
- Rico 266
- Riiho 371
- Rimini 29
- Rio de Janeiro 141, 392
- Rio Grande do Sul 370
- Rioja 177, 180
- Riojan glosses 128, 142
- Ripert 274
- Ristori 268
- Ritmo Laurenziano* 32, 33
- Ritmo su S. Alessio* 164
- ritual 101
- Rivarol 263
- Rivarola 238
- River Plate 370, 386
- Roberts 204, 209, 437
- Roche gude 274
- Rodríguez, F. J. 276
- Rogge 275
- Rohlf's 23, 197, 203, 208, 209, 230, 254, 288, 319, 328, 331, 332, 334
- Röllig 14, 15
- Rom, Roma, Romany 192, 359
- Romagnol 306, 328
- Romaine 144
- Roman
 - Eastern Empire 75
 - Empire 9–10, 12, 13, 17, 18, 19, 21, 27, 34, 36, 37, 38, 40, 41, 42, 52, 56, 63, 64, 72, 73, 75, 76, 77, 78, 91, 94, 98, 117, 134, 139, 238, 391
 - law 159, 165, 175, 179
 - political rule 7
 - provinces 9, 21, 168
 - religions 10
 - Republic 73
- Romance languages
 - classification 28
 - dialectal continuum 44
 - diatopic variation 52
 - emergence of vernaculars 7, 8, 53, 55
 - family tree 11
 - genesis 400, 402, 403
 - geographic diffusion 6
 - morphological differences 48
 - mutual intelligibility 44
 - new graphies 45–51
 - number of 317
 - official spoken standard/norm 46
 - phonetic differences 48
 - proto- 28, 35, 61, 85, 95, 96, 98, 118
 - shared features 27, 36
 - spoken language 55
 - syntactic differences 48
 - textual production 55–56
 - written language 55
- Romance philology 402, 403, 404, 406, 407, 409, 443
- romance, romanz* 336
- Romand 169
- România (the) 28, 143, 156, 180, 283, 363
- Romania 6, 240, 284, 287, 305, 306, 309, 310, 311, 312, 316, 323, 332, 358, 359, 394
- 'Greater Romania' 361

- Romania (cont.)
 Hungarian (Magyar) population 256
 Saxon population 256
 Szeckler population 256
 Turkish rule 255
- Romania Novissima* 360
- Romania Submersa* 286
- Romanian Academy 281
- Romanian 7, 11, 30, 32, 33, 34, 116, 120, 123, 124, 141, 159, 161, 194, 201, 202, 212, 213, 214, 215, 216, 217, 218, 223, 226, 228, 229, 231, 232, 233, 234, 235, 236, 237, 238, 254–256, 261–262, 271, 280–282, 283, 287, 310, 315, 323, 331, 332, 334, 339, 341, 343, 348, 356, 359; *see also* Aromanian; Daco-Romanian; Istro-Romanian; Megleno-Romanian; Muntenian; Moldovan dialects; Moldovan language; Wallachian dialects 254, 310
 grammars 254, 272
 Greek neologisms 271
 lexicography 254
 old 123
- Romanians 308, 315, 317; *see also* Aromanians; Istro-Romanians; Megleno-Romanians
 Orthodox 256
- Romanization 10, 11, 12, 13, 15, 16, 19, 20, 39, 40, 63, 168
- Romans 200, 204, 209, 356
- Romansh 11, 162, 203, 205, 209, 210, 254, 276, 282, 283, 284, 287, 299, 300, 306, 315, 350; *see also* *Rumantsch grischun*
- Romantic movement, Romanticism 64, 176, 273, 289
- Rome 18, 29, 31, 32, 37, 38, 39, 40, 41, 66, 100, 118, 119, 184, 206, 211, 277, 304, 305, 325, 332, 337, 340, 347
- Romeo 157
- Römische Volkssprache* 408
- Rona 191
- Ronsard 260
- Roques 318
- Rosetti 123, 124, 201, 217, 254, 255, 256, 261, 262, 271, 280
- Rosier 321
- Rossellonese 178
- Rossi, G. 270
- Rossi, M. 86, 106
- Rottet 381, 397
- Rouen 168
- Rouergue 173, 174
- Rousselle 65, 72
- Roussillon 178, 295, 341
- Rovigno 285, 302
- Rozzo 241
- Rubin 191, 374
- Ruffino 321, 322
- Ruiz, J. 51
- Rumantsch Grischun* 300, 341
- Ruscelli 244
- Russia 310
- Russian 212, 213, 218, 284, 310, 316, 359
- Russians 310, 316
- Russu 201
- rustica romana lingua* 81, 88
- Rusu 311
- Rwanda 362
- Ryan 379
- /s/ (word-final)
 in Latin 30, 338
 in Romance 338
- Sá de Miranda 253
- Sabater 250, 251, 266, 275
- Sabatini 31, 115, 137, 156, 157, 158, 163
- Sabir 424
- saccio* 45
- Saco y Arce 276
- Sahrawi Arab Democratic Republic
 362, 367
- Șăineanu 255
- Saint Bernard Parish 377, 397
- Saint Lawrence 389
- Saint-Amand 170
- Saint-Armand-les-Eaux 170
- Saint-Barthélémy 362
- Saint-Cyran-du-Jambot 69, 71
- Saint-Denis 170, 171
- Saint-Gilles 175
- Saint-Martin 362
- Saint-Pierre and Miquelon 362
- Saint-Riquier 170
- Saintes (Charente-Maritime) 13
- saints, cult of 76
- saints, lives of 76, 82, 101
- Sala 23, 187, 191, 373
- Salamanca 44, 290, 326
- Salazar 384
- Salcedo 388
- Salentino 230
- Salonika, *see* Thessaloniki
- Salonius 33
- Salvi 90, 371
- Salviati 257, 280
- Salzburg 355
- San Francisco 369

- San Marcos, University of 392
 San Marino 306
 San Martín 291
 San Millán de la Cogolla 48, 177, 181
 Sánchez 238
 Sandfeld 191, 201, 217, 235
 sandhi 128
 Sanga 89, 163, 164
 Sankoff 387
 Sanson 153, 162, 176, 241, 244, 264, 269, 278, 361
 Santa Cruz de Coimbra 251
 Santiago de Compostela 181
 Santo Domingo de Sillos 181
sao 45
 São Paulo 393
 São Tomé and Príncipe 362, 364, 366
 São Tomense 366
 Sardinia 8, 9, 12, 22, 55, 56, 178, 285, 292, 304, 307, 316, 328
 Sardinian 22, 32, 33, 166, 178, 196, 202, 218, 233, 270, 285, 301, 304, 307, 313, 318, 324, 329, 331, 332, 334, 339, 341, 348; *see also* palaeo-Sardinian
 Sardinians 314
 Sardo-Corsican area 301
 Sarmatian 22
 Sartre 66
 Sas 93, 100
 Saši 309
 Saskatchewan 388
 Sassanid empire 337
 Sassarese 329
 Sassari 301
 Saussure 237, 319, 321, 410, 420
 Sauvageot 188
 Savaria 40
 Savoia 198
 Savoy 245, 299
 House of 299, 302
 Savoyard 169
 Sawa 376
 Saxons 354; *see also* saši
 Saxony 241
sayagués 345
 Scaglione 237
 Scagno 305
 Scandinavian 203, 208
 Schäfer-Priess 252
 Scharlau 403
 Scherre 436
 Schiaffini 64, 119
 Schleicher 11
 Schlieben-Lange 144, 145, 151, 274, 417, 441, 444
 Schmid, H. 265, 300, 341
 Schmitt, C. 1, 2, 359
 Schneider 434, 437, 444
scholasticus sermo 78
 Scholle 404
 schools 75, 307; *see also* education system
 Schøsler 124
Schreiblandschaften 151
Schriftsprache 411
 Schuchardt 10, 11, 24, 319, 402, 404, 405, 408, 415, 418, 419, 422, 424
 Schüle 298
 Schwegler 421, 424
 Schweickard 1
 Schwenter 439
 Schwytzerdütsch, *see* Swiss German
 science 159
 scientific terminology 183, 190
 scientific texts 182
Școala ardeleană 271
 Scotland 13
scripta 45, 81, 84, 86, 88, 115, 143, 151, 299, 303
scripta latina rustica 156
 scriptology 151
scriptoria 83
 scriptorial features 152, 153, 155
 Scriptures 240, 241
 Scupi (Skopje) 286
 Second World War 290, 307, 312, 321, 359, 386
Sefarad 399
 Segarra 266
 Segesta 21
 selection 144
 Selig 49, 106, 148, 157, 158
 Selinunte 21
 Sella 326
 semantic change 140
 semantics 103, 104, 111, 136, 141
 Semitic 226
 Sénac 79
 Sénégal 362, 376, 385, 393
 Senghor 393
 Senigallia 6
 Sephardic communities 290, 316, 367, 399
 Séphiha 189, 367
 Septimus Severus 17, 21
 Serbia 309, 310, 313, 315
 Serbian 214, 284, 285
 Serbs 309, 359
 serial verbs 427
 Serianni 269, 277, 279
sermo altus 65
sermo humilis 65, 73

- sermo piscatorius* 73
sermo rusticus 82
 sermons 67, 68, 82
Servisc de Planificazion y Elaborazion dl Lingaz Ladin 300
 'settlement colonies' 363, 370, 373–376
 Severinus 42
 Seville 182
 Seychelles 362, 396
 Shane 333
 Sharp 269
 Shohamy 399
 sibilants 112, 367
 Sicanian 195
 Sichel 21
 Sicilian 110, 165, 226
 Sicilian School of poetry 165
 Sicily 8, 9, 226, 230, 285, 288, 304
 Sicinius Pudens 20
 Sidonius Apollinaris 77
 Siegel 147, 148, 433
 Sigiramus 69, 70, 81
Siglo de Oro, *see* Golden Age
 Silva 439
 Silva-Corvalán 383
 Silva-Fuenzalida 413
 Simone 261
 Şincai 271, 272
 Sinnemäki 146
 Sirmium 354
 Sissano 302
 Sixtus V 241
 Skok 286
 slang 365
 Slav, Slavonic 3, 63, 159, 161, 189, 192, 202, 212, 221, 222, 225, 229, 232, 234, 255, 261, 262, 271, 285, 288, 307, 346, 353
 slaves 357, 358, 363
 Slavs 286, 287, 302, 308, 316, 354
 Slovenia 303, 315
 Slovenian 213, 284, 307
 Slovenians 287, 358
 Smith, J. C. 114, 122
 Smith, N. 422, 428
Société des écrivains coloniaux 393
 sociolinguistics 57, 58, 59, 65, 66, 91, 93, 94, 106, 108, 110, 111, 125, 134, 148, 167, 187, 188, 231, 320, 321, 322, 327, 335
 sociophilology 141
soi tiers 232
 Solà 266, 275
 Solalinde 182
 Sornicola 30, 106, 203, 209
 sound atlas 322
 sound change 404, 405, 406, 408, 411, 433
 sound laws 318, 319, 320, 329, 331
 Southern Cone 372
 SOV order 113
 Spain 29, 34, 78, 79, 84, 87, 100, 106, 108, 115, 120, 190, 240, 245, 248, 251, 270, 275, 285, 288, 289, 292, 313, 316, 317, 326, 342, 351, 352, 358, 359, 362, 363, 364, 367, 391, 399
 Hither and Further Spain 9
 Spallanzani, 268
 Spanglish 369, 386
 Spaniards 317
 Spanish 1, 7, 11, 33, 51, 56, 63, 85, 99, 102, 107, 108, 109, 112, 113, 115, 117, 120, 121, 122, 123, 124, 126, 129, 132, 134, 138, 139, 140, 141, 150, 152, 167, 175, 176, 177, 178, 179, 183, 184, 185, 190, 191, 194, 195, 196, 197, 205–212, 227, 228, 232, 234, 238, 248, 249, 250, 251, 252, 253, 254, 262–263, 266, 267, 283, 288, 289, 290, 291, 292, 293, 302, 313, 314, 315, 316, 317, 326, 327, 328, 339, 342, 343, 347, 349, 350, 357, 358, 360, 361, 363, 364, 367, 369, 370, 371, 372, 373–376, 377, 378, 383–384, 386–387, 388, 390–392, 396, 397–398, 401, 404, 405, 409, 413, 419, 423, 432, 442
 Andalusian 109, 176, 346, 372
 Argentinian 391
 Caribbean 401, 436, 437
 corrupt 402, 423–424
 dictionaries 264–265
 Ecuadorian Sierra 190
 grammars 249–250, 265
 Gallicisms 265, 269
 Judæo-Spanish 115, 189, 190, 191, 213, 214, 228, 282, 285, 290, 367–368, 399; *see also* Ladino
 Latin American varieties 326, 372, 403
 lengua castellana 249
 lengua española 249
 middle 123
 morphology 249
 old Castilian 33, 50–50, 121, 265
 old Spanish 33, 109, 121, 405
 orthography 249, 265
 Peninsular 367
 proto-Castilian 87
 Puerto Rican 372
 Venezuelan 117
 Spanish America 338, 350, 357
 Spanish Civil War 317

- Spanish empire 364
 Spanish Marches 176
 Spanish Succession, War of 265
 Spano 270
 spelling 115, 128, 131, 132, 136, 155, 158, 391, 392
 Speroni 246
 Spevak 139
 spirants, spirantization 197, 198, 234
 Split 287
 spoken language 45, 46, 52, 53, 54, 244, 338,
 352, 409
 Spoleto 205, 210
 Spolsky 399
 Spongano 258
 Sprachbund 201
 standardization 7, 148, 237, 238, 240, 243, 267,
 276, 282, 314, 341, 343, 401, 406, 407, 409,
 416, 417
 vernacular/standard divide 410–411
 status 189, 190, 340–343
 Stavans 386
 Ștefan 261
 Ștefanini 263
 Stefanelli 5, 91, 103
 Stegagno Picchio 251
 Stein 402, 403, 406, 407
 Steiner 250
 Stewart 96, 106
 Stimm 34
 Štivor 316
 Stobi 286
 Stotz 30, 58, 104
 Straka 347
 Strasbourg 37, 355
Strasbourg Oaths 30, 33, 45, 52, 117, 118, 121, 134,
 136, 146, 157, 169, 170, 204, 210
 Strassi 24
 stratigraphic interpretation 318
 stress 35, 203, 207, 208, 228, 229, 324
 structuralism, structuralists 187, 193, 234, 320, 321
 American 400, 412, 415, 421, 443
 structured accommodation 219, 220
 Stussi 124
 style 94
 subject 352
 subjunctive 98, 99
 future 99, 138
 imperfect 99
 pluperfect 352
 subordinate clauses 128
 substrate, substratum 3, 23–24, 98, 168, 178, 187,
 189, 190, 192–201, 234, 235, 373–378, 404, 418,
 422, 427, 428, 429, 430, 433, 434, 438
 Südtirol, *see* Alto Adige
 suffixes 98, 194, 215, 227, 229, 230, 233, 320
 Suleiman the Magnificent 255
 Sulpicio Verulano 239
Summa Trecensis 175
 superstrate 36, 98, 168, 187, 192, 194, 234, 418,
 422, 427, 430
 suppletion 221, 222, 224
 supra-regional 170, 172
 supra-regionality 148, 154, 160, 162, 169, 177
 Surmiran 276
Surselva 299
 Surselvan 32, 34, 254, 276
 Šušnjevića 218, 219
Sutselva 299
 SVO order 113
 Swigart 385
 Swiggers 174, 244, 247, 249, 266, 272, 276
 Swiss Confederation 283, 300
 Swiss German 315, 350, 410
 Swiss Romandy 296, 298, 299
 Switzerland 6, 245, 282, 285, 287, 294, 296, 299,
 306, 307, 313, 315, 316, 342, 350
 Sylvain 434
 synchrony 319, 321
 syncope 200
 syntactic structure 94
 syntagmeme 102
 syntax 86, 103, 110, 112, 113, 115, 136, 185, 207,
 228, 230, 234, 239, 240, 332, 377, 388; *see also*
 borrowing
 syntheticity 403
 Syria 193
 Székely 309
 Szlezák 397
 Szombathely 40

 /t/ (word-final) 324
tabarchino 285
 taboo 338
 Tabouret-Keller 433, 442
 Tagalog 418
 Tagliavini 2, 162, 254, 256, 282
 Taino 374
 Talamo 279
 Talian 370
 Tampa 369
 Tanzania 389
 Țara Oașului 213
 Tarallo 437, 438, 439
 Taranto 7, 8
 Tarraco 41
 Tarragona 178

- Tassoni, Alessandro 258
 Tatar 284
 Tatars 308
 Tavoni 239, 243, 248, 249
 technical terminology 234
 Telmon 277, 298
 Tempea 272
 tense, mood, aspect (TMA) markers 414, 421,
 422, 426, 435, 443
 Teramo 31
 Teresa of Castile 183
 Ternateño 423
 Ternes 200
 Terni 29
 Terracini 163, 318, 319, 329
 terrain 323
 Terrebonne-Lafourche 397
 Terry 251, 265
 Tertullian 34
 Tesi 241
 text messages 142
 text tradition 144
 texts 336
 Teyssier 251, 253, 269, 275, 276, 372
 theatre 345, 346
 Theodoric 211
 Theodulf of Orléans 130
 therapy 318, 320
 Thessaloniki 290, 312, 399
 Thiers 342
thiotisca lingua 82, 88
 third declension 93
 Third Republic 351
 Thomas, A. 398
 Thomas, H. 82
 Thomas, J. 406, 407
 Thomason, 188, 191, 379, 402, 427, 428, 430,
 431, 435
 Thrace, Thracia 286
 Thracian 22, 201
 Thraco-Dacian 193, 194, 199, 200, 201
 Thun 321, 322
 Tiber 6
 Tiberius 13
 Ticino (Canton) 284, 303, 306, 307, 342
 Tiddis 17
 Timmermans 260
 Timoc 309, 310
 Țircomnicu 313
 Tirso de Molina 262
 Tizzano, Gaetano 244
 Todoran 123, 124
 Togo 362, 377
 Tok Pisin 414, 415, 427
 Toledo 120, 158, 160, 176, 182, 183, 289
 Tolomei, Claudio 243
 Tolosa 32
 Tomasella 305
 Tongres 37
 toponyms, *see* placenames
 Tordesillas, Treaty of 363
 Torres 166
 Torres Montes 177
 Tortosa 178
 Tory 247
 Toso 305
 Toubert 65
 Toul 37
 Toulouse 158, 174
 Touraine 69
tourangeau 293
 tourism 348
 Tourmon, Antoine 264
 Tours 40, 47, 71, 170, 245
 Council of 53, 81
 Tovar 41
 Trabalza 243, 244
 Trajan 17, 24, 37
 Transcarpathia 310
 translation 84, 85, 126, 159
 Transnistrian Republic 310
 transparency principle 433
 transport routes 168
 transposition interference 147
 Transylvania 229, 254, 255, 256, 261,
 271–272, 281, 282, 284, 309, 315, 323,
 325, 359
 Trask 22, 107, 196
 Trastámaras 250, 292
Tre corone 165, 166
 Tremblay 381
 Tremiti islands 304
 Trentino dialect 287, 300, 304, 325, 328
 Trentino 284, 300, 304, 307
 Trento 305, 316
 Treptow 280, 281
 Trévoux 264
 Trier 41, 355
 Trieste 285, 301, 307, 325, 347, 358
 Trifone 162, 243, 304
 triglossia 177, 303, 327
 Tripoli 20
 Tripolitania 287
 Trissino 243, 244
trobadors 157, 165, 174, 274, 291
trovadores 275

- Trovato 237, 241
 Trudgill 59, 92, 432, 433, 442
 Trumper 230
 Tsanga 377
 Tsiapera 261
 Tuaillon 245
 Tucker 395
 Tuki 376
 Tunisia 9, 15, 16, 367, 393
 Tupi 374
 Turculeț 254, 272
 Turin 268, 346
 Turkey 290, 312, 399
 Turkish 192, 215, 228–229, 232, 233, 236, 284, 359, 399
 influence on Romanian 255
 Turks 308, 316, 359
 Tuscan 122, 165, 166, 203, 204, 208, 209, 234, 241, 242, 245, 269, 279, 302, 303, 304, 305, 306, 325, 328, 332, 342, 343, 347
 fourteenth-century 242, 257, 267
 spoken 245
 Tuscany 56, 121, 165, 166, 245, 277, 305, 325, 348
 Tuten 177, 181, 182, 432
 Twelfth-Century Renaissance 119, 138, 159
 typology 333
 Tyrol 300
 Tyrolean 284, 299, 323

uariatio 68
 Uc Faidit 174
 Udaina 282, 308
 Uddholm 93
 Udina, *see* Udaina
 Udine 307, 328, 349
 Ukraine 309, 310, 315
 Ukrainian 213, 284, 285, 310, 323, 359
 Ukrainians 309, 310, 316
 Ullastra 266
 ULUCCUS 26
 ULULA 26
 Umbria 32, 305
 Umbrian (ancient) 6, 22
 Umbrian Confession Formula 33
 Unamuno 126
 Union of Soviet Socialist Republics (USSR) 282, 309, 310, 315, 341
 United States of America (US) 316, 361, 362, 364, 368–369, 383–384, 386, 387, 388, 399, 444
 Universal Grammar 428
 universals 422
 universities 159, 171

 unmarkedness 433
 urban centres 346
 urban usage 349
 urbanization 304
 Urciuoli 383
 Ureche 261
 Urițescu 213
 Urraca (daughter of Alfonso VI of Castile) 184
 Ursicinus 38
 Ursini 308, 316
 Uruguay 361, 362, 387
 Uytfanghe, van 58, 72, 77, 78, 80, 81, 83, 85, 89, 91, 104, 116

 Väänänen 29, 64, 68
 Val d'Aosta 245, 284, 296, 298, 299, 307, 358
 Val d'Aran 297, 298
 Val di Non 300
 Val di Sole 300
 Val Gardena 300, 315
 Val Müstair 299
 Val Venosta 299
 Valdés 248
 Valdman 397, 416, 421
 Valence 175
 School 175
 Valencia 177, 178, 179, 251, 289, 292, 293, 316, 341
 Kingdom of 178
 Valencian 115, 178
 Valkhoff 190
 Valla 239
 Valladolid 183
 Valle 302
 Vallette 20
 Vallisnieri 268
 Valpuesta 31
 van Coetsem 190
 van Name 415
 Van Raemdonck 321
 Vandeloise 92
 Vanuatu 362, 398
 Vanvolsem 244
 variation 133, 143, 145, 161, 167, 348–353
 Varlaam 262
 Varvaro 5, 11, 12, 13, 20, 22, 43, 44, 195, 282, 331, 361, 390
 Vatican City 306
 Vaugelas 259–260
 vectoriality 150
 Veglia, *see* Krk
 Vegliote 282
 velarization 213

- Velius Longus 130, 142
 Venantius Fortunatus 105
 Vendryès 177, 189
 Venetan 285, 301, 304, 308, 324, 325, 327
 Venetian 213, 219, 268, 282, 285, 287,
 288, 300, 301, 302, 303, 308, 346, 347,
 349, 355, 370
 Venetians 316
 Venetic 6, 22, 195
 Veneto 204, 210, 300, 304, 305, 307, 316, 317,
 349, 370
 Veneto-Giulian 303
 Venezia Giulia 301
 Venezuela 112, 361; *see also* Spanish,
 Venezuelan
 Venice 258, 268, 300, 346
 Republic of 300, 304
 Veny 178
 verb paradigm 236
 verb-second 203, 209; *see also* word order
 Verdo 137
 Vergil 47, 242
 Vermont 369
 vernacular 47, 48, 49, 53, 54, 55, 56, 238, 239, 240,
 241, 242, 245, 246, 250, 252, 255, 256, 270,
 406, 407, 409, 423
 vernacular vs. standard divide 410–411
 vernacularization 163
 Verona 31
 Veronese Communes 284
 Veronese, G. 239
 Versailles 394
Verschriftlichung 160
Verschriftung 158
 Versteegh 401
Verstümmelungen 403
 vertical communication 65, 70, 71, 72, 75,
 76, 77, 78, 79, 80, 81, 84, 85, 88, 89, 90,
 95, 104
 Vezin 106
 Vian, Josef Anton 276
 Vicente, G. 253
 Videsott 154, 163, 164
 Vidos 188
 Vidossi 330
 Vieliard 30, 32
 vigesimal counting 194, 200
 Vigo 279
 Vikings 170
 Villalobos 249
 Villalón 249
 Villers-Cotterêts, Ordinance of 153, 246,
 274, 295
 Villon 346
 Vincent, N. 1, 2, 65, 90, 92
 Vinet 380
 Virginia 369
 Visigothic 174
 Visigoths 78, 79, 87, 140, 176, 177, 182, 203, 204,
 206, 208, 209, 211
Vita Leudegarii 77
Vita Radegundis 104
Vita Richarii 105
 Vitale 241, 267, 268
 Vlachs, *see* Aromanians
vlasi, *see* Aromanians
 Vléry 107
 vocabulary 114, 115, 136, 137, 138, 141, 182
 vocative 15, 139, 216, 217, 236
 Voghera 245
 voicing 200
 Voivodina 309, 310, 312
volgare beneventano 164
volgare illustre 165
volgare italico 163, 165
volgare toscano 164
 Völker 149, 154, 167, 172
volkstaal 407, 408
 Voßler 172, 319
 vowels
 centralization 194
 fronting 200
 height 74
 lengthening 207
 quantity 74
 reduction 74
 rounding of front vowels 194
 shortening 74
 systems 35
 unstressed 206, 334
 Vulgar Latin 57
 Vulgate 134; *see also* Bible
 Vurpas 169

 /w/ in loanwords 206
 Wackernagel's Law 378
 Waldensians 297, 298
 Wallace 4
 Wallachia 228, 254, 255, 261, 271, 281, 308
 Wallachian 254
 Wallensköld 345
 Wallington 268
 Wallis and Futuna 362
 Wallonia 170, 171, 206, 294
 Walloon 169, 170, 204, 209, 295, 296
 Walloons 315

- Walsers 284
 Waquet 240, 241
 Ward-Perkins 42
 Wartburg 8, 36, 168, 202, 203, 207–226, 333
 Way of Saint James 180, 181
 weakening 198, 324, 328
 of postnasal consonants 23, 195, 234
 Webster 395
 Weimar 175
 Weinreich 59, 146, 187, 188, 191, 192, 225,
 237, 321
 Weinrich 320
 Welland 378
 Wells 18
 Wendt 255
 Wenker 317
 Wenzel 144
 Werminghoff 81
 Wexler 191
 Wheeler 261
 Whinnom 422, 423, 424
 Whitney 418
 Wilhelm 165
 William of Aquitaine 174
 Windisch 159
 Windisch 348
 Winford 190, 390
 Winkelmann 254, 322
 Winniffrith 312
 Wolf 29
 Wolfram 36
 Woll 252
 Wolof 376, 385
 Wood 75
 Woolf 18, 19
 word order 48, 136, 139, 194, 339, 426
 SOV 340
 SVO 37
 typology 333
 word-formation 183, 194, 233
 Worms 37
 Wright 5, 43, 44, 46, 48, 49, 58, 63, 72, 78,
 81, 82, 83, 84, 86, 88, 89, 91, 93, 103, 106,
 108, 115, 116, 117, 118, 119, 131, 137, 138, 141,
 151, 155, 157, 167, 177, 182, 237
 writing 303
 ‘writing landscapes’ 151, 152
 written (vs. spoken) language 45, 52, 189, 317,
 343, 409, 411
 Wüest 169
 Wuilleumier 13
 Wunderli 157, 158
 Würzburger Federprobe 162

Xogos Florais 276
 -xss- for -x- 15

 Yiddish 359
 yod 339
 Yucatán 236
 Yugoslavia 214, 304, 309, 312
 Yunga 358

 Zadar 162
 Zamboni 65, 89
 Zamora 290, 326
 Zamora Vicente 326
 Zamora, J. 190, 191
 Zara 308
 Zaragoza 44, 130
 Treaty of 363
 Zawadowski 191
 Žejane 219, 221, 313
 Zeno 211
 Zentella 383
 zeta 135
 Zimmerman 350
 Zuccagni Orlandini 279