

INTERNATIONAL STANDARD

NORME INTERNATIONALE

Text-to-speech for television – General requirements

Synthèse vocale pour télévision – Exigences générales





THIS PUBLICATION IS COPYRIGHT PROTECTED

Copyright © 2013 IEC, Geneva, Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either IEC or IEC's member National Committee in the country of the requester.

If you have any questions about IEC copyright or have an enquiry about obtaining additional rights to this publication, please contact the address below or your local IEC member National Committee for further information.

Droits de reproduction réservés. Sauf indication contraire, aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie et les microfilms, sans l'accord écrit de la CEI ou du Comité national de la CEI du pays du demandeur.

Si vous avez des questions sur le copyright de la CEI ou si vous désirez obtenir des droits supplémentaires sur cette publication, utilisez les coordonnées ci-après ou contactez le Comité national de la CEI de votre pays de résidence.

IEC Central Office
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
Fax: +41 22 919 03 00
info@iec.ch
www.iec.ch

About the IEC

The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

About IEC publications

The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigenda or an amendment might have been published.

Useful links:

IEC publications search - www.iec.ch/searchpub

The advanced search enables you to find IEC publications by a variety of criteria (reference number, text, technical committee,...).

It also gives information on projects, replaced and withdrawn publications.

IEC Just Published - webstore.iec.ch/justpublished

Stay up to date on all new IEC publications. Just Published details all new publications released. Available on-line and also once a month by email.

Electropedia - www.electropedia.org

The world's leading online dictionary of electronic and electrical terms containing more than 30 000 terms and definitions in English and French, with equivalent terms in additional languages. Also known as the International Electrotechnical Vocabulary (IEV) on-line.

Customer Service Centre - webstore.iec.ch/csc

If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: csc@iec.ch.

A propos de la CEI

La Commission Electrotechnique Internationale (CEI) est la première organisation mondiale qui élabore et publie des Normes internationales pour tout ce qui a trait à l'électricité, à l'électronique et aux technologies apparentées.

A propos des publications CEI

Le contenu technique des publications de la CEI est constamment revu. Veuillez vous assurer que vous possédez l'édition la plus récente, un corrigendum ou amendement peut avoir été publié.

Liens utiles:

Recherche de publications CEI - www.iec.ch/searchpub

La recherche avancée vous permet de trouver des publications CEI en utilisant différents critères (numéro de référence, texte, comité d'études,...).

Elle donne aussi des informations sur les projets et les publications remplacées ou retirées.

Just Published CEI - webstore.iec.ch/justpublished

Restez informé sur les nouvelles publications de la CEI. Just Published détaille les nouvelles publications parues. Disponible en ligne et aussi une fois par mois par email.

Electropedia - www.electropedia.org

Le premier dictionnaire en ligne au monde de termes électroniques et électriques. Il contient plus de 30 000 termes et définitions en anglais et en français, ainsi que les termes équivalents dans les langues additionnelles. Egalement appelé Vocabulaire Electrotechnique International (VEI) en ligne.

Service Clients - webstore.iec.ch/csc

Si vous désirez nous donner des commentaires sur cette publication ou si vous avez des questions contactez-nous: csc@iec.ch.



IEC 62731

Edition 1.0 2013-01

INTERNATIONAL STANDARD

NORME INTERNATIONALE

Text-to-speech for television – General requirements

Synthèse vocale pour télévision – Exigences générales

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

COMMISSION
ELECTROTECHNIQUE
INTERNATIONALE

PRICE CODE
CODE PRIX

R

ICS 33.160.25; 33.160.99

ISBN 978-2-83220-600-3

**Warning! Make sure that you obtained this publication from an authorized distributor.
Attention! Veuillez vous assurer que vous avez obtenu cette publication via un distributeur agréé.**

CONTENTS

FOREWORD.....	3
1 Scope.....	5
2 Normative references	5
3 Terms, definitions and abbreviations	5
3.1 Terms and definitions	5
3.2 Abbreviations	7
4 Guiding principles and conventions	7
5 User requirements of visually impaired people.....	7
5.1 Users' needs	7
5.2 Navigating channels	8
5.3 Navigating TV inputs	8
5.4 Additional data services	8
5.5 Operating the TV	8
5.6 TV use	9
6 Functional requirements	9
6.1 Functionality for TV, TTS device combination	9
6.2 Functionality: TTS device/engine.....	10
6.3 Functionality: TV	10
6.4 Setting up: TV, TTS device combination	10
7 TV events and TTS data.....	11
7.1 TV context and events.....	11
7.2 TTS data per event.....	12
7.2.1 Details.....	12
7.2.2 Channel change	12
7.2.3 Additional information	13
7.2.4 Navigation and selection.....	13
7.2.5 Context switch.....	14
7.2.6 Pop-up message.....	15
8 TTS profiles.....	16
8.1 Basic profile	16
8.2 Main profile	16
8.3 Enhanced profile	17
8.4 Summary.....	17
Bibliography.....	19
Figure 1 – TV – TTS device system diagram.....	9
Figure 2 – Context event state diagram.....	12
Table 1 – Overview of profiles.....	18

INTERNATIONAL ELECTROTECHNICAL COMMISSION

TEXT-TO-SPEECH FOR TELEVISION – GENERAL REQUIREMENTS

FOREWORD

- 1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
- 2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
- 3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
- 4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
- 5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
- 6) All users should ensure that they have the latest edition of this publication.
- 7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
- 8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
- 9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

International Standard IEC 62731 has been prepared by IEC technical committee 100: Audio, video and multimedia systems and equipment.

The text of this standard is based on the following documents:

FDIS	Report on voting
100/2070/FDIS	100/2109/RVD

Full information on the voting for the approval of this standard can be found in the report on voting indicated in the above table.

This publication has been drafted in accordance with the ISO/IEC Directives, Part 2.

The committee has decided that the contents of this publication will remain unchanged until the stability date indicated on the IEC web site under "<http://webstore.iec.ch>" in the data related to the specific publication. At this date, the publication will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

TEXT-TO-SPEECH FOR TELEVISION – GENERAL REQUIREMENTS

1 Scope

This International Standard specifies the text-to-speech functionality for a (broadcast) receiver with a text-to-speech system. Such a system may be one device, i.e. a receiver with an integrated text-to-speech generator, or may be two devices, i.e. a receiver interfacing with an external text-to-speech device. This International Standard applies only to completely functional stationary (or semi-stationary) digital TV receivers such as set top boxes, integrated digital TVs, recorders and other products whose primary function is to receive TV content. Where this standard refers to TV, this will be shorthand for all such receivers.

This International Standard does not apply to products that are capable of receiving TV as a secondary function (e.g. PCs or game consoles with digital television receivers). It also does not apply to sub-assemblies (e.g. PC tuner cards).

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

(void)

3 Terms, definitions and abbreviations

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1.1

context

one specific function of a TV

EXAMPLE Watching TV, EPG, etc.

3.1.2

DTV broadcast event

set of related broadcast streams with a defined start and end time, commonly referred to as a TV programme

Note 1 to entry: DTV events have typically the following properties associated with them: start time, end time or duration, content synopsis, additional content information, parental rating and availability of subtitles or audio description.

3.1.3

DTV service information

metadata describing broadcasting content and its scheduling and timing details

Note 1 to entry: The purpose of DTV service information is to aid the end user to select and schedule viewing and recording, and also to select the equipment configuration.

3.1.4

DTV broadcast event classification

general category of programme/event content, or its classification

EXAMPLES Movie (drama), news/current affairs, talk show, sports (football), etc.

3.1.5

EPG filter

filter that organises or reduces the list of displayed EPG items according to certain criteria

EXAMPLES Of criteria are to show only:

- programmes with a certain content type;
- favourites;
- programmes that are audio described;
- programmes for a given time period (for instance "today", "tomorrow", "next 7 days").

3.1.6

event

trigger to start an action

3.1.7

list

collection of items

3.1.8

menu

subsequent order of items

3.1.9

receiver

device capable of receiving or handling digital television signals

3.1.10

service

sequence of programs under the control of a broadcaster which can be broadcast as part of a schedule

3.1.11

subtitle

textual representation of the dialogue (and frequently additional auditory information), typically shown at the bottom of the screen

Note 1 to entry: Subtitles can be a textual rendering in the same language as the spoken dialogue, or can provide a written translation in a different language.

Note 2 to entry: In some parts of the world subtitles are called "(closed) captions", and subtitling is referred to as "(closed) captioning".

Note 3 to entry: This standard uses the term subtitles throughout.

3.1.12

TTS audio

audio output by the TTS engine in correspondence with TTS data

Note 1 to entry: If the TV uses an external TTS converter, TTS audio is interpreted as TTS data.

3.1.13

TTS data

(text) data converted into TTS audio information by the text-to-speech engine

3.2 Abbreviations

For the purposes of this document, the following abbreviations apply.

DTV	digital television
EPG	electronic programme guide
STB	set top box
TTS	text-to-speech
TV	television
UI	user interface

4 Guiding principles and conventions

This standard describes the required basic behaviour for a TV text-to-speech combination in a basic profile, but also provides for enhanced profiles. It also gives a short introduction into the basic problems of visually impaired people: i.e. what are the problems visually impaired people experience when using and watching TV?

Providing text-to-speech functionality for a broadcast receiver, e.g. TV or STB can be of great help to (visually) disabled people. Such speech functionality may be integrated in the receiver or may be external to the receiver in a separate device.

In general as the guiding principle, when building a TTS interface in the context of this standard, implementers should aspire to achieve functional equivalence of the user experience. This means that a person operating the device using the speech interface should have access to similar information and be able to accomplish similar tasks as with a graphical UI.

The main features of this International Standard are:

- basic functional description for a TV-TTS device combination or TV with integrated TTS;
- profiles for different levels of TV-TTS functionality;
- targeted towards the digital TV application.

In this standard mandatory requirements are specified; optional and informative features are also included.

A claim of conformity with this standard requires conformity with all mandatory requirements. A TV-TTS device combination or a TV with a TTS that is integrated may provide options for a user to enable or disable product features.

5 User requirements of visually impaired people

5.1 Users' needs

This subclause 5.1 explains the needs of visually impaired people as the primary target users for a TV with TTS. Unless these needs are met, the system is not accessible to this user group. Visually impaired people experience access barriers in the course of the following activities when watching TV:

- a) following TV programming, e.g. the TV series;
- b) using a remote control;
- c) not being able to see subtitles;
- d) navigating channels;

- e) navigating TV inputs;
- f) using additional data (text) services provided by the broadcaster, e.g. an EPG;
- g) daily operation of the TV and initial setup of the TV for use.

Items a), b) and c) are outside the scope of this standard. Item c) further relates to the fact that in some countries foreign language programmes are being translated via subtitles. For users who cannot see the subtitles, supplementary audio services are sometimes used to deliver an audio version of the subtitles. This standard elaborates on the remaining four items, i.e. d), e), f) and g), in 5.2 to 5.6.

NOTE 1 For DVB systems, item a) is already solved by audio description. Also, the use case of providing supplementary audio services to deliver an audio version of the subtitles is covered in the DVB-SI specification ETSI EN 300 468.

NOTE 2 For ATSC systems, the audio system includes a visually impaired (VI) associate service which allows a complete programme mix containing music, effects, dialogue, and additionally a narrative description of the picture content, see ATSC A/53 part 5 and part 6.

5.2 Navigating channels

The problem is a user does not know which channel the TV displays, i.e. the user gets “lost during navigation”. The TV is displaying navigation data on the screen but the user is unable to see it. Such data are for example:

- channel number,
- service name,
- (DTV broadcast) event name.

5.3 Navigating TV inputs

The problem is that a user is unable to select the required input to the TV, e.g. the user wishes to select DTV or a specific external input linked to a recording or other device. The choice is shown on the screen but the user is unable to see it.

5.4 Additional data services

With digital TV a broadcaster may transmit additional data (text) services to augment TV programming, provide additional information on programming, or provide news. Such additional data are:

- information about whether audio description, subtitling is available,
- (next) (DTV broadcast) event name,
- (DVB-) event information (enhanced description of the (DTV broadcast) event),
- EPG data.

The items above are listed in order of importance with the most important item appearing first. It is noted that this data provides additional convenience in using the TV, but that is non-essential for the primary function of watching TV, and selecting channels.

5.5 Operating the TV

User settings are another needed function besides navigation. This can be done through buttons on the remote control (out of scope for this specification), but also via on-screen menus. For visually impaired people on-screen menus are typically of little use.

A distinction exists between initial setup and daily operation of the TV. Initial setup is typically a onetime operation during the lifetime of a TV. Daily operation is more frequent and more important. Consequently a distinction among menu items for daily operation exists, those addressing specific accessibility functions, and TV setup menu items. However, the most frequently used keys are “volume”, “channel up/down”, and number keys.

5.6 TV use

Use characterization of a TV helps in determining implementation profiles. Navigating channels, for example, is done most often when watching TV, as well as commands like volume up and down. This may be supported by additional data services, but does not affect the primary functions of the TV. Changing the TV's system settings is not done very often, except perhaps for changing sound or video settings or switching audio description on and off. Such settings may have an easy access mode through a special menu. TV installation is typically performed only once during the lifetime of the TV. Often, visually impaired people can benefit from specialized support for installing the TV, i.e. it is part of the service when buying a new TV. Understanding this life- and usage cycle of a TV helps with defining the most effective and efficient solutions and is reflected in the profiles. In the following paragraphs, we refer to “basic”, “main” and “enhanced” profiles as further defined and detailed in Clause 8.

Key operations for a minimum TTS implementation on a receiver for TV use are as set out in the basic profile defined in this standard. This basic profile shall include:

- a) channel number, name and event information – key for a user to identify which service has been selected;
- b) availability of audio description – key for a user to know about the availability of this service feature;
- c) availability of subtitles – key for a user to know about the availability of this service feature;
- d) basic EPG – allow the user to navigate through the EPG, if such data is present in the broadcast, to identify which future events are available to them;
- e) context changes – key for a user to understand if the TV went to another state or when a pop-up message appears;
- f) the main profile shall in addition to all the items from the basic profile include receiver menu functions (allows the user to navigate receiver operations and functions).

Additional operations that shall be included in the enhanced profile, in addition to all those from the basic and main profiles, are:

- g) event Information – provide the event synopsis;
- h) additional EPG data – allows the user to get more info on the service or event;
- i) operations of a recording device – allows the user to record future events, possibly selected via the EPG. Play/pause a recorded event.

6 Functional requirements

6.1 Functionality for TV, TTS device combination

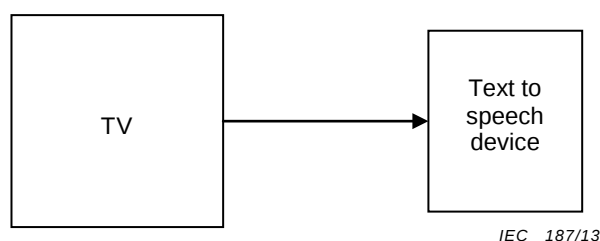


Figure 1 – TV – TTS device system diagram

The TV-TTS device system diagram is illustrated in Figure 1. As shown in the figure the TTS device is a separate function from the TV, which can be implemented on a device connected with a TV-TTS device interface, or may also be integrated in the TV.

The functionality requirements for a TV with TTS combination are:

- the delay between an event and the resulting TTS audio related to that event shall be such that they are perceived as belonging tied together;
- priority TTS audio shall overrule currently playing TTS audio information;
- the user should be able to stop currently playing TTS audio;
- the user shall be able to repeat the current or previous TTS audio;
- the user shall be able to mute the TTS audio;
- the user shall be able to switch on/off the TTS function;
- the language of the TTS audio shall be the same as set for the TV's UI, except when signalled differently. The TTS device/engine may choose to pronounce the text or to indicate failure in case it does not support the signalled language;
- TTS audio may not need to literally represent the related visual information on the screen as long as the meaning of the visual information stays intact.

6.2 Functionality: TTS device/engine

The TV, in principle, only outputs text strings towards the TTS device.

The TTS device shall follow these outputs:

- an external TTS device should be designed to be fully accessible to visually impaired users without being dependent on the TV;
- the volume level of the TTS device/engine shall be changeable by the user. The TTS device shall announce the new volume level. It should be possible to do this independent from the TV volume;
- the user should be able to adjust speech characteristics like speed, pitch, voice type, when applicable;
- the TTS engine should announce abbreviations as such, letter by letter, rather than as a normal word. Example: "TTS" should be announced "T T S" ("tee tee es") instead of "tts" ("tetes") where relevant. The TTS engine may also pronounce common acronyms in full, e.g. "sub" could be spoken as "subtitles" where appropriate. This standard does not identify what is understood to be an abbreviation, rather leaves this at the discretion of implementers;
- the TTS engine should announce numbers in a manner suited to the context, e.g. as natural number, digit-by-digit, etc.;
- the TTS engine is considered to determine the context.

6.3 Functionality: TV

The TV determines the user interface, i.e. what is displayed on the screen, and how the TV interacts with the user. The TV therefore also determines which text is sent to the TTS engine.

The user should be able to control, via the TV TTS, settings like volume, speed, pitch, voice type.

NOTE In Europe, to fulfil basic accessibility needs, the TV is expected to comply with SELFC.

6.4 Setting up: TV, TTS device combination

TTS audio guidance may be given during the installation of a receiver. TTS audio guidance for receiver installation should adhere to the following requirements when implemented:

- the requirements from 6.1 should apply;
- the user should be able to easily switch on the TTS audio at the beginning of the installation;

- all essential information to complete the installation procedure should be available as TTS audio. Special attention should be given to inform the user the progress of slow processes such as channel searches.

Additional suggestions for TTS audio guidance during receiver installation may be found in DTG-UG024.

7 TV events and TTS data

7.1 TV context and events

The TV can output data in response to events and events can take place in different contexts. The following TV contexts can exist (this list is not necessarily exhaustive):

- watching TV: the user watches TV;
- list: the user browses a (channel) list;
- (acc-)menu: the user browses through the (accessibility) menu structure of the TV. If the menu contains accessibility services, it is called an accessibility menu, or acc-menu. An example of an accessibility service is audio description;
- EPG: the user is browses the EPG;
- standby: the TV is on stand-by;
- timeshift: the user may start a time shift recording or play a time shift recording.

The following contexts are currently outside the scope of this specification and are to be considered for future versions of this specification:

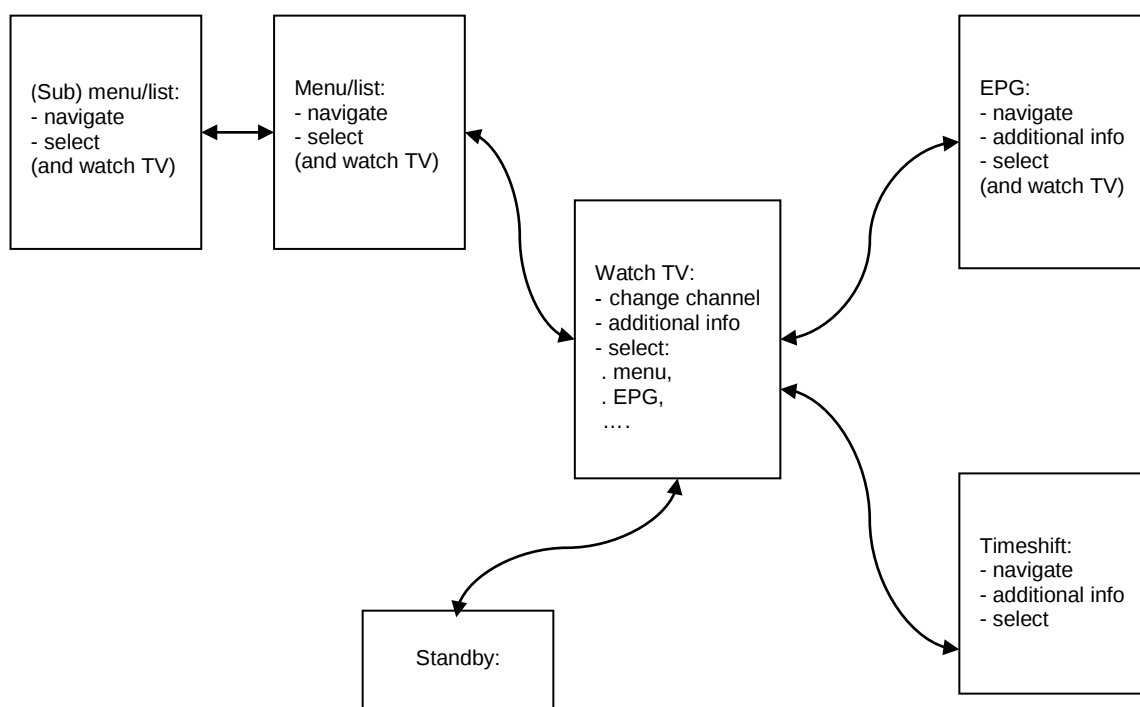
- interactive TV;
- pay-TV.

Within said contexts the following events may occur:

- channel change: the TV changes channels after the user initiated a change channel event, e.g. by pressing a button on the remote control;
- additional information: the TV presents additional information on the item shown. This item may be the event shown on the display, but may also be a highlighted menu item, etc. Additional information may be presented, e.g. after the user has pressed the “info” button on the remote control;
- navigation and selection: the TV highlights another item in e.g. a list or menu. The user selects an item;
- context switch: the TV switches to another context, e.g. from watching TV to menu;
- pop-up message: the TV puts a message on the screen for some purpose, e.g. PIN required, or gives status information.

The TV is expected to output data after an event to the TTS device using the TV-TTS interface. These data contain an indication of the event and the TTS data to be announced.

The different TV contexts and the events which may occur in a certain context are graphically shown in Figure 2. Figure 2 lists the events in a certain context and shows the possible transitions between the different contexts. It also shows that in the menu context it is possible to switch to sub (menus).



IEC 188/13

Figure 2 – Context event state diagram

7.2 TTS data per event

7.2.1 Details

Subclauses 7.2.2 to 7.2.6 specify which TTS data the TV shall output to the TTS device/engine given a specified event.

7.2.2 Channel change

7.2.2.1 General

A channel change event outputs data on the newly selected channel. Channel change typically only occurs in the watch TV context.

7.2.2.2 Watch TV context

A channel change event in the watch TV context shall at least output the following TTS data for the newly selected channel if present:

- <channel number>
- <service name>
- <present (DTV broadcast) event name>
- <audio description y/n>.

A channel change event in the watch TV context may also output the following other recommended TTS data if present:

- <present (DTV broadcast) event start time>
- <present (DTV broadcast) event end time or duration>

- g) <next (DTV broadcast) event name>
- h) <next (DTV Broadcast) event start time>
- i) <next (DTV Broadcast) event end time or duration>
- j) <audio description yes/no next (DTV broadcast) event>.

7.2.3 Additional information

7.2.3.1 General

An additional information event shall output the additional available information on the selected/highlighted item as shown on the TV screen as TTS data.

NOTE The TV channel shown on the TV is a “selected item”, i.e. it is the selected channel.

Additional information typically occurs in the watch TV, EPG, and timeshift contexts.

7.2.3.2 Watch TV context

An additional information event in watch TV context may also output the following other recommended information as TTS data:

- a) <(DTV broadcast) event synopsis>
- b) <parental rating>
- c) <next (DTV broadcast) event synopsis>
- d) <presence of subtitles y/n>
- e) <next (DTV broadcast) event parental rating>
- f) <encrypted y/n>
- g) <DTV broadcast event classification>.

7.2.3.3 EPG context

An additional information event in the EPG context may also output the following other recommended TTS data:

- a) <(DTV broadcast) event synopsis>
- b) <parental rating>
- c) <subtitle availability>
- d) <languages available>.

7.2.3.4 Timeshift context

An additional information event in the timeshift context may output the following data on an event or service to be recorded or to be played back:

- a) <(DTV broadcast) event synopsis>
- b) <parental rating>
- c) <subtitle availability>.

7.2.4 Navigation and selection

7.2.4.1 General

A navigation and selection event shall output as TTS data/audio the (new) highlighted item. The TV shall output as TTS data/audio a selected and/or changed item.

The highlighted item may only be fully understandable in context with other items on the screen. In this case these other items should also be output as TTS data. For example: in a matrix type of EPG the x,y position of the highlighted item also represents information. In this case the contents of the x-axis and y-axis at the position of the highlighted item should also be output as TTS data.

7.2.4.2 List context

The generic requirements for navigation and selection in 7.2.4.1 apply.

7.2.4.3 (Acc-)Menu context

The generic requirements for navigation and selection in 7.2.4.1 apply.

7.2.4.4 EPG context

The generic requirements for navigation and selection in 7.2.4.1 apply. The TV shall output as TTS data the individual elements listed below if also displayed on screen:

- a) channel number,
- b) service name,
- c) audio description (if available),
- d) subtitles (if available),
- e) event name,
- f) event start time,
- g) event end time or duration.

For usability reasons when moving within a service in the EPG, it is recommended that the TV does not repeat the <channel number> and <service name>.

This list does not mandate a specific order of enunciation of these items.

7.2.4.5 Timeshift context

The generic requirements for navigation and selection in 7.2.4.1 apply. Typically it may be expected that in this context the commands play, pause, rewind, forward, stop, and record are covered.

7.2.5 Context switch

7.2.5.1 General

A context switch event shall output as TTS data the context name and may output guidance on how to operate in the selected context. Guidance data may be on what type of buttons can be used and what actions they trigger.

Guidance data on how to operate in a certain context is up to the manufacturer. Examples of guidance data:

- “You are in the EPG context. Use navigation buttons. Up/down navigation will run you through all channels. Left/right navigation will run you through the time schedule.”
- “You are in the TV menu. Use up/down and OK to operate the menu.”

7.2.5.2 Watching TV context

The same TTS audio for a channel change event in the watching TV context shall be output as TTS audio when entering the watching TV context from the standby context.

7.2.5.3 List context

When entering the list context the TV shall output as TTS data:

- <list title>
- <number of list items>.

7.2.5.4 (Acc-) Menu Context

When entering the (Acc-) Menu Context the TV shall output as TTS data:

- <menu title>
- <number of menu items>.

7.2.5.5 EPG Context

When entering the EPG context the TV shall output as TTS data:

- <highlighted item>.

The TV may also output as TTS data:

- <EPG overview>: shortly describe EPG
- <EPG filter setting applied>.

The TV shall output as TTS data if the <highlighted item> is an (DTV Broadcast) event and is shown on screen:

- a) <channel number>
- b) <service name>
- c) <audio description y/n>
- d) <subtitles y/n>
- e) <present (DTV broadcast) event name>
- f) <present (DTV broadcast) event start time>
- g) <present (DTV broadcast) event end time or duration>.

7.2.5.6 Standby context

When entering standby the user should be notified that the receiver is switching to standby.

7.2.5.7 Timeshift context

When entering timeshift the user should be notified that the receiver is switching to a context for time shifting.

7.2.5.8 Currently not supported contexts

Contexts interactive TV and pay-TV are currently outside the scope of this specification. It is recommended to inform the user of entering an unsupported context and to inform the user on how to leave said context again.

7.2.6 Pop-up message

For a visually impaired user pop-up messages often constitute an insurmountable access barrier, unless fully supported by the TTS device. The visually impaired user must be made aware of the occurrence of the pop-up, its contents in full and what action is required from the user in order to proceed. Pop-up messages are often modal, thus preventing the user from exercising normal control and navigation until the pop-up has been dismissed.

The user shall be informed of any warnings and notifications, such as tuning issues or PIN control. When a pop-up message appears on the screen it should output as TTS data:

- <notification of pop-up>
- <description of pop-up>.

TTS audio on changing status information shall be provided regularly and after completion. It is not required to provide for TTS audio at every update.

8 TTS profiles

8.1 Basic profile

A TV with integrated TTS or a TV-TTS device combination shall at least support the Basic profile. The basic profile shall support TTS audio as specified in the event and context clauses for the following events and contexts:

- channel change event:
 - watch TV context;
- navigation and selection event:
 - list context,
 - acc-menu context,
 - EPG context;
- additional information event:
 - watch TV context;
- context switch event:
 - list context,
 - standby context,
 - acc-menu context,
 - EPG context;
- pop-up message event.

NOTE A channel change event can be seen as a special case of a navigation event.

8.2 Main profile

The main profile shall support TTS audio as specified in the event and context clauses for the following events and contexts:

- channel change event:
 - watch TV context;
- navigation and selection event:
 - list context,
 - acc-menu context,
 - menu context,
 - EPG context;
- additional information event:
 - watch TV context;
- context switch event:
 - list context,

- standby context,
- acc-menu context,
- menu context,
- EPG context;
- pop-up message event.

8.3 Enhanced profile

The enhanced profile shall support TTS audio as specified in the event and context clauses for the following events and contexts:

- channel change event:
 - watch TV context;
- navigation and selection event:
 - list context,
 - acc-menu context,
 - menu context,
 - EPG context,
 - timeshift context;
- additional information event;
 - watch TV context,
 - EPG context,
 - timeshift context;
- context switch event:
 - list context,
 - standby context,
 - acc-menu context,
 - menu context,
 - EPG context,
 - timeshift context;
- pop-up message event.

8.4 Summary

Table 1 gives an overview of the different profiles.

Table 1 – Overview of profiles

Event	Context	Basic profile	Main profile	Enhanced profile
channel change	watch TV context	X	X	X
navigation and selection event	list context	X	X	X
	acc-Menu context	X	X	X
	menu context		X	X
	EPG context	X	X	X
	timeshift context			X
additional information event	watch TV context	X	X	X
	EPG context			X
	timeshift context			X
context switch event	list context	X	X	X
	standby context	X	X	X
	acc-Menu context	X	X	X
	menu context		X	X
	EPG context	X	X	X
	timeshift context			X
pop-up message event	“all contexts in profile”	X	X	X

Bibliography

IEC/TR 62678:2010, *Audio, video and multimedia systems and equipment activities and considerations related to accessibility and usability*

ISO/IEC 13818-1, *Information technology – Generic coding of moving pictures and associated audio information: Systems*

ISO/IEC TR 29138-3:2009, *Information technology – Accessibility considerations for people with disabilities – Part 3: Guidance on user needs mapping*

DTG-UG024, *Digital TV group – UK digital TV usability and accessibility guidelines, including text to speech (u-book)*, version 1.0, September 2011

ETSI EN 300 468, *Digital video broadcasting (DVB) – Specification for service information (SI) in DVB systems*

ETSI TR 101 211, *Digital video broadcasting (DVB); Guidelines on implementation and usage of service information (SI)*

ETSI TS 101 154, *Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream*

ATSC A/53 part 5: *Digital television standard – AC-3 audio system characteristics: 2010*

ATSC A/53 part 6: *Digital television standard – Enhanced AC-3 audio system characteristics: 2010*

ARIB STD-B10: *Service Information for Digital Broadcasting System*

ARIB STD-B32: *Video Coding, Audio Coding and Multiplexing Specifications for Digital Broadcasting*

SELFCE, *Industry Self-Commitment, to improve the accessibility of digital TV receiving equipment sold in the European Union*, Brussels, November 30, 2007. Available at <<http://www.digitaleurope.org/IndustryGroups/TechnicalandRegulatoryGroup/eInclusion/Documents.aspx>>

SOMMAIRE

AVANT-PROPOS	21
1 Domaine d'application	23
2 Références normatives	23
3 Termes, définitions et abréviations	23
3.1 Termes et définitions	23
3.2 Abréviations	25
4 Principes directeurs et conventions	25
5 Exigences d'utilisateur des personnes ayant une déficience visuelle	26
5.1 Besoins des utilisateurs	26
5.2 Navigation entre les chaînes	26
5.3 Navigation entre les ports d'entrée de la télévision	26
5.4 Services de données supplémentaires	26
5.5 Utilisation du téléviseur	27
5.6 Utilisation du téléviseur	27
6 Exigences de fonctionnement	28
6.1 Fonctionnalité d'une combinaison téléviseur-dispositif de synthèse vocale	28
6.2 Fonctionnalité: dispositif/moteur de synthèse vocale	29
6.3 Fonctionnalité: téléviseur	29
6.4 Mise en service: combinaison téléviseur-dispositif de synthèse vocale	29
7 Événements de télévision et données de synthèse vocale	30
7.1 Contexte de télévision et événements	30
7.2 Données de synthèse vocale par événement	31
7.2.1 Détails	31
7.2.2 Changement de chaîne	31
7.2.3 Informations supplémentaires	32
7.2.4 Navigation et sélection	33
7.2.5 Changement de contexte	33
7.2.6 Message contextuel	35
8 Profils de synthèse vocale	35
8.1 Profil de base	35
8.2 Profil principal	36
8.3 Profil amélioré	36
8.4 Résumé	37
Bibliographie	38
Figure 1 – Schéma d'un système téléviseur-dispositif de synthèse vocale	28
Figure 2 – Diagramme d'état des événements de contexte	31
Tableau 1 – Vue d'ensemble des profils	37

COMMISSION ÉLECTROTECHNIQUE INTERNATIONALE

SYNTHÈSE VOCALE POUR TÉLÉVISION – EXIGENCES GÉNÉRALES

AVANT-PROPOS

- 1) La Commission Electrotechnique Internationale (CEI) est une organisation mondiale de normalisation composée de l'ensemble des comités électrotechniques nationaux (Comités nationaux de la CEI). La CEI a pour objet de favoriser la coopération internationale pour toutes les questions de normalisation dans les domaines de l'électricité et de l'électronique. A cet effet, la CEI – entre autres activités – publie des Normes internationales, des Spécifications techniques, des Rapports techniques, des Spécifications accessibles au public (PAS) et des Guides (ci-après dénommés "Publication(s) de la CEI"). Leur élaboration est confiée à des comités d'études, aux travaux desquels tout Comité national intéressé par le sujet traité peut participer. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec la CEI, participent également aux travaux. La CEI collabore étroitement avec l'Organisation Internationale de Normalisation (ISO), selon des conditions fixées par accord entre les deux organisations.
- 2) Les décisions ou accords officiels de la CEI concernant les questions techniques représentent, dans la mesure du possible, un accord international sur les sujets étudiés, étant donné que les Comités nationaux de la CEI intéressés sont représentés dans chaque comité d'études.
- 3) Les Publications de la CEI se présentent sous la forme de recommandations internationales et sont agréées comme telles par les Comités nationaux de la CEI. Tous les efforts raisonnables sont entrepris afin que la CEI s'assure de l'exactitude du contenu technique de ses publications; la CEI ne peut pas être tenue responsable de l'éventuelle mauvaise utilisation ou interprétation qui en est faite par un quelconque utilisateur final.
- 4) Dans le but d'encourager l'uniformité internationale, les Comités nationaux de la CEI s'engagent, dans toute la mesure possible, à appliquer de façon transparente les Publications de la CEI dans leurs publications nationales et régionales. Toutes divergences entre toutes Publications de la CEI et toutes publications nationales ou régionales correspondantes doivent être indiquées en termes clairs dans ces dernières.
- 5) La CEI elle-même ne fournit aucune attestation de conformité. Des organismes de certification indépendants fournissent des services d'évaluation de conformité et, dans certains secteurs, accèdent aux marques de conformité de la CEI. La CEI n'est responsable d'aucun des services effectués par les organismes de certification indépendants.
- 6) Tous les utilisateurs doivent s'assurer qu'ils sont en possession de la dernière édition de cette publication.
- 7) Aucune responsabilité ne doit être imputée à la CEI, à ses administrateurs, employés, auxiliaires ou mandataires, y compris ses experts particuliers et les membres de ses comités d'études et des Comités nationaux de la CEI, pour tout préjudice causé en cas de dommages corporels et matériels, ou de tout autre dommage de quelque nature que ce soit, directe ou indirecte, ou pour supporter les coûts (y compris les frais de justice) et les dépenses découlant de la publication ou de l'utilisation de cette Publication de la CEI ou de toute autre Publication de la CEI, ou au crédit qui lui est accordé.
- 8) L'attention est attirée sur les références normatives citées dans cette publication. L'utilisation de publications référencées est obligatoire pour une application correcte de la présente publication.
- 9) L'attention est attirée sur le fait que certains des éléments de la présente Publication de la CEI peuvent faire l'objet de droits de brevet. La CEI ne saurait être tenue pour responsable de ne pas avoir identifié de tels droits de brevets et de ne pas avoir signalé leur existence.

La Norme internationale CEI 62731 a été établie par le comité d'étude 100 de la CEI: Systèmes et appareils audio, vidéo et multimédia.

Le texte de cette norme est issu des documents suivants:

FDIS	Rapport de vote
100/2070/FDIS	100/2109/RVD

Le rapport de vote indiqué dans le tableau ci-dessus donne toute information sur le vote ayant abouti à l'approbation de cette norme.

Cette publication a été rédigée selon les Directives ISO/CEI, Partie 2.

Le comité a décidé que le contenu de cette publication ne sera pas modifié avant la date de stabilité indiquée sur le site web de la CEI sous "<http://webstore.iec.ch>" dans les données relatives à la publication recherchée. A cette date, la publication sera

- reconduite,
- supprimée,
- remplacée par une édition révisée, ou
- amendée.

SYNTHÈSE VOCALE POUR TÉLÉVISION – EXIGENCES GÉNÉRALES

1 Domaine d'application

La présente Norme Internationale définit la fonctionnalité de synthèse vocale pour un récepteur (de radiodiffusion) avec système de synthèse vocale. Un tel système peut être constitué d'un dispositif, à savoir, un récepteur avec un générateur de synthèse vocale intégré, ou être constitué de deux dispositifs, à savoir, un récepteur interfacé avec un dispositif de synthèse vocale extérieur. La présente Norme Internationale s'applique uniquement à des récepteurs de télévision numérique fixes (ou semi-fixes) entièrement fonctionnels, tels que des boîtiers décodeurs, des téléviseurs numériques intégrés, des enregistreurs et d'autres produits dont la fonction principale est de recevoir un contenu de télévision.

La présente Norme ne s'applique pas à des produits permettant de recevoir la télévision en tant que fonction secondaire (par exemple, des PC ou des consoles de jeux avec récepteurs de télévision numérique). Il ne s'applique pas non plus aux sous-ensembles (par exemple, cartes de syntoniseur pour PC).

2 Références normatives

Le document suivant, en totalité ou en partie, est référencé de façon normative dans le présent document et est indispensable pour son application. Pour les références datées, seule l'édition citée s'applique. Pour les références non datées, la dernière édition du document de référence s'applique (y compris les éventuels amendements).

(vide)

3 Termes, définitions et abréviations

3.1 Termes et définitions

Pour les besoins du présent document, les termes et définitions suivants s'appliquent.

3.1.1

contexte

une fonction spécifique du téléviseur

EXEMPLE Regarder la télévision, guide électronique de programme, etc.

3.1.2

événement de diffusion de télévision numérique

ensemble de flux de radiodiffusion avec heures de début et de fin définies, appelé généralement "Émission de télévision"

Note 1 à l'article: Les propriétés suivantes sont généralement associées aux événements de télévision numérique: heure de début, heure de fin ou durée, synopsis du contenu, informations supplémentaires sur le contenu, classement parental et disponibilité de sous-titres ou d'une description audio.

3.1.3

informations de service de télévision numérique

métadonnées décrivant le contenu de radiodiffusion ainsi que ses détails de programmation et d'horaires

Note 1 à l'article: Ces informations ont pour but d'aider l'utilisateur final à choisir et à programmer la visualisation et l'enregistrement ainsi que la configuration du matériel.

3.1.4

classification des événements de diffusion de télévision numérique

catégorie générale de contenu de programme/d'événement ou sa classification

EXEMPLE film (dramatique), informations/questions d'actualité, émission-débat, sport (football), etc.

3.1.5

filtre du guide électronique de programme

filtre qui organise ou réduit la liste d'éléments du guide électronique de programme affichés selon certains critères

EXEMPLES de critère pour ne présenter que:

- des émissions ayant un certain type de contenu;
- des émissions préférées;
- des émissions avec une description audio;
- des émissions pendant une période de temps donnée (par exemple "aujourd'hui", "demain", "dans les 7 prochains jours").

3.1.6

événement

déclencheur pour démarrer une action

3.1.7

liste

collection d'éléments

3.1.8

menu

ordre d'éléments consécutifs

3.1.9

récepteur

dispositif capable de recevoir ou de traiter des signaux de télévision numérique

3.1.10

service

séquence d'émissions contrôlée par un diffuseur, pouvant être diffusées en tant que partie d'un horaire

3.1.11

sous-titre

représentation textuelle du dialogue (et fréquemment, d'information audibles supplémentaires), affichée généralement en bas de l'écran

Note 1 à l'article: Les sous-titres peuvent être constitués d'une restitution textuelle dans la même langue que le dialogue parlé ou peuvent fournir une traduction écrite dans une langue différente.

Note 2 à l'article: Dans certaines régions du monde, les sous-titres sont appelés "légendes (fermées)" et le sous-titrage est appelé "sous-titrage (fermé)".

Note 3 à l'article: La présente norme utilise partout le terme sous-titres.

3.1.12

audio par synthèse vocale

sortie audio du moteur de synthèse vocale en correspondance avec les données de synthèse vocale

Note 1 à l'article: Si le téléviseur utilise un convertisseur de synthèse vocale externe, l'audio par synthèse vocale est interprété comme des données de synthèse vocale.

3.1.13

données de synthèse vocale

données (texte) converties en informations audio par synthèse vocale par le moteur de synthèse vocale

3.2 Abréviations

Pour les besoins du présent document, les abréviations suivantes s'appliquent:

Abréviation	Anglais	Français
DTV	digital television	télévision numérique
EPG	electronic programme guide	guide électronique de programme
STB	set top box	boîtier décodeur
TTS	text-to-speech	synthèse vocale
TV	television	télévision
UI	user interface	interface utilisateur

4 Principes directeurs et conventions

La présente norme décrit le comportement de base exigé d'une combinaison téléviseur-synthèse vocale dans un profil de base, mais indique également des profils améliorés. Elle fournit également une courte introduction aux problèmes fondamentaux rencontrés par les personnes ayant une déficience visuelle: Quels sont les problèmes rencontrés réellement par les personnes ayant une déficience visuelle lorsqu'ils utilisent et qu'ils regardent la télévision?

La fourniture d'une fonctionnalité de synthèse vocale pour un récepteur de radiodiffusion, par exemple un téléviseur ou un boîtier décodeur, peut être d'un grand secours aux personnes ayant une déficience (visuelle). Une telle fonctionnalité peut être incorporée au récepteur ou peut être extérieure au récepteur, dans un dispositif séparé.

En règle générale comme principe directeur, lors de l'élaboration d'une interface de synthèse vocale dans le contexte de la présente norme, il convient que les responsables de la mise en œuvre visent à obtenir une équivalence fonctionnelle de l'expérience de l'utilisateur. Ceci signifie qu'il convient qu'une personne faisant fonctionner le dispositif utilisant l'interface vocale ait accès à des informations similaires et soit capable d'exécuter des tâches similaires à celles d'une interface utilisateur graphique.

Les principaux points traités dans la présente Norme internationale sont les suivants:

- description du fonctionnement de base d'une combinaison téléviseur-dispositif de synthèse vocale ou d'un téléviseur avec dispositif de synthèse vocale incorporé;
- profils pour des niveaux différents de fonctionnalité du téléviseur-dispositif de synthèse vocale;
- adaptation à l'application de télévision numérique.

Des exigences obligatoires sont spécifiées dans la présente norme, des propriétés facultatives et informatives sont également incluses.

Une déclaration de conformité avec la présente norme exige la conformité avec toutes les exigences obligatoires. Une combinaison téléviseur-dispositif de synthèse vocale ou un téléviseur avec dispositif de synthèse vocale incorporé peut proposer des options à un utilisateur pour activer ou désactiver les fonctions du produit.

5 Exigences d'utilisateur des personnes ayant une déficience visuelle

5.1 Besoins des utilisateurs

Ce paragraphe 5.1 explique les besoins des personnes ayant une déficience visuelle en tant que principaux utilisateurs cibles d'un téléviseur avec synthèse vocale. Sauf si ces besoins sont satisfaits, le système n'est pas accessible à ce groupe d'utilisateurs. Les personnes ayant une déficience visuelle sont confrontées à des obstacles au cours des activités suivantes lorsqu'ils regardent la télévision:

- a) Suivi de la programmation de télévision, par exemple, une série télévisée;
- b) Utilisation d'une télécommande;
- c) Impossibilité de voir les sous-titres;
- d) Navigation entre les chaînes;
- e) Navigation entre les ports d'entrée de la télévision;
- f) Utilisation de services de données (textuelles) supplémentaires proposés par le radiodiffuseur, par exemple un guide électronique de programme;
- g) Utilisation quotidienne du téléviseur et mise en service initiale du téléviseur pour son utilisation.

Les points a), b) et c) ne font pas partie du domaine d'application de la présente norme. Le point c) concerne en outre le fait que dans certains pays, les émissions en langue étrangère sont traduites par l'intermédiaire de sous-titres. Pour les utilisateurs qui ne peuvent pas voir les sous-titres, des services audio supplémentaires sont parfois utilisés pour fournir une version audio des sous-titres. La présente norme donne des détails sur les quatre points restants, à savoir, d), e), f) et g) en 5.2 à 5.6.

NOTE 1 Pour les systèmes de diffusion vidéo numérique, le point a) est déjà résolu par la Description Audio. D'autre part, le cas d'utilisation consistant à fournir des services audio supplémentaires pour délivrer une version audio des sous-titres est traité par la spécification DVB-SI ETSI EN 300 468.

NOTE 2 Pour les systèmes ATSC, le système audio contient un service associé pour les personnes ayant une déficience visuelle (VI) permettant d'obtenir un mélange complet de programmes contenant de la musique, des effets, un dialogue et de plus, une description narrative du contenu des images, voir ATSC A/53 partie 5 et partie 6.

5.2 Navigation entre les chaînes

Le problème est qu'un utilisateur ne sait pas quelle est la chaîne affichée par le téléviseur, c'est-à-dire que l'utilisateur "se perd pendant la navigation". Le téléviseur affiche sur l'écran des données de navigation mais l'utilisateur est incapable de les voir. Ces données sont par exemple:

- le numéro de chaîne,
- le nom du service,
- le nom d'un événement (diffusion de télévision numérique).

5.3 Navigation entre les ports d'entrée de la télévision

Le problème est celui d'un utilisateur qui n'est pas capable de choisir le port d'entrée nécessaire sur la télévision, par exemple l'utilisateur souhaite choisir télévision numérique ou un port d'entrée externe spécifique raccordé à un appareil d'enregistrement ou autre. Le choix est présenté sur l'écran mais l'utilisateur ne peut pas le voir.

5.4 Services de données supplémentaires

Avec la télévision numérique, un radiodiffuseur peut transmettre des services de données supplémentaires (textuelles) pour compléter la programmation de télévision, fournir des informations supplémentaires sur la programmation ou diffuser des informations. Ces données supplémentaires sont:

- des informations concernant la disponibilité d'une description audio, d'un sous-titrage,
- le nom d'événement (diffusion de télévision numérique) (suivant),
- des informations concernant un événement (diffusion de télévision numérique) (description améliorée de l'événement (de diffusion vidéo numérique)),
- des données du guide électronique de programme.

Les points ci-dessus sont énumérés par ordre d'importance, le point le plus important apparaissant en premier. On notera que ces données apportent une commodité supplémentaire pour utiliser le téléviseur mais qu'elles ne sont pas essentielles à la fonction principale consistant à regarder la télévision et au choix des chaînes.

5.5 Utilisation du téléviseur

Outre la navigation, les réglages de l'utilisateur sont une autre fonction nécessaire. Ceux-ci peuvent être effectués en utilisant les touches de la télécommande (ne faisant pas partie du domaine d'application de cette spécification), mais également par l'intermédiaire de menus sur l'écran. Pour les personnes ayant une déficience visuelle, les menus sur l'écran ont généralement peu d'utilité.

Il existe une distinction entre la mise en service initiale et le fonctionnement quotidien du téléviseur. La mise en service initiale est généralement une opération effectuée une seule fois au cours de la durée de vie du téléviseur. Le fonctionnement quotidien est plus fréquent et plus important. Il existe donc une distinction entre les éléments de menu pour le fonctionnement quotidien, ceux qui concernent des fonctions d'accessibilité spécifiques et les éléments de menus de réglage du téléviseur. On notera toutefois que les touches les plus souvent utilisés sont "volume", "chaîne suivante/précédente" et les touches numériques.

5.6 Utilisation du téléviseur

La caractérisation de l'utilisation d'un téléviseur facilite la détermination des profils de mise en œuvre. La navigation parmi les chaînes, par exemple, est effectuée le plus souvent en regardant la télévision, de même que des commandes telles que l'augmentation et la diminution du volume. Ceci peut être pris en charge par des services de données supplémentaires mais n'a pas d'influence sur les fonctions principales du téléviseur. La modification des réglages du système du téléviseur n'est pas effectuée souvent, sauf peut-être pour modifier les réglages du son ou de la vidéo ou pour activer et désactiver la description audio. Ces réglages peuvent comporter un mode d'accès rapide par un menu spécial. L'installation d'un téléviseur est généralement effectuée une seule fois pendant la durée de vie du téléviseur. Les personnes ayant une déficience visuelle peuvent souvent utiliser une assistance spécialisée pour installer le téléviseur, c'est-à-dire que celle-ci fait partie du service lors de l'achat d'un nouveau téléviseur. La compréhension de ce cycle de durée de vie et d'utilisation d'un téléviseur facilite la définition des solutions les plus efficaces et les plus rationnelles et il est représenté dans les profils. Dans les alinéas suivants, on se réfère à des profils "de base", "principal" et "amélioré" comme cela est défini plus en détail à l'Article 8.

L'utilisation des touches pour la mise en œuvre minimale d'une synthèse vocale sur un récepteur pour l'utilisation du téléviseur est présentée dans le profil de base défini dans la présente norme. Ce profil de base doit inclure:

- a) le numéro et le nom de la chaîne et des informations relatives aux événements – touche pour qu'un utilisateur identifie le service qui a été choisi;
- b) la disponibilité d'une description audio – touche pour qu'un utilisateur soit informé de la disponibilité de cette possibilité de service;
- c) la disponibilité de sous-titres – touche pour qu'un utilisateur soit informé de la disponibilité de cette possibilité de service;

- d) le guide électronique de programme de base – permet à un utilisateur de naviguer dans le guide électronique de programme, si ces données sont présentes dans la radiodiffusion, pour identifier les futurs événements auxquels il aura accès;
- e) modifications du contexte – touche pour qu'un utilisateur comprenne si le téléviseur est passé dans un autre état ou si un message contextuel est apparu;
- f) en plus de tous les points du profil de base, le profil principal doit inclure des fonctions de menu du récepteur (permettant à l'utilisateur de naviguer parmi les opérations et les fonctions du récepteur).

Les actions supplémentaires qui doivent être incluses dans le profil amélioré, en plus de toutes celles des profils de base et principal, sont les suivantes:

- g) informations concernant un événement – fournissent le synopsis de l'événement;
- h) données supplémentaires du guide électronique de programme – permettent à l'utilisateur d'obtenir plus d'information sur le service ou l'événement;
- i) fonctionnement d'un dispositif d'enregistrement – permet à l'utilisateur d'enregistrer des événements futurs, pouvant être éventuellement sélectionnés par l'intermédiaire du guide électronique de programme. Lecture/pause d'un événement enregistré.

6 Exigences de fonctionnement

6.1 Fonctionnalité d'une combinaison téléviseur-dispositif de synthèse vocale

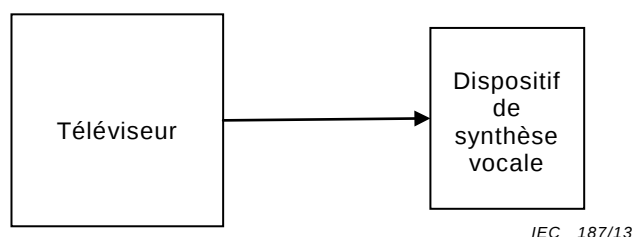


Figure 1 – Schéma d'un système téléviseur-dispositif de synthèse vocale

Le schéma d'un système téléviseur-dispositif de synthèse vocale est illustré à la Figure 1. Comme représenté sur la figure, le dispositif de synthèse vocale est une fonction séparée du téléviseur, pouvant être mise en œuvre sur un dispositif raccordé avec une interface téléviseur-dispositif de synthèse vocale ou pouvant également être incorporé dans le téléviseur.

Les fonctionnalités d'une combinaison de téléviseur avec synthèse vocale sont les suivantes:

- le retard entre un événement et l'audio par synthèse vocale résultant associé à cet événement doit être tel qu'ils soient perçus comme liés l'un à l'autre;
- l'audio par synthèse vocale prioritaire doit annuler la lecture des informations audio par synthèse vocale;
- il convient que l'utilisateur soit capable d'arrêter l'audio par synthèse vocale en cours de lecture;
- l'utilisateur doit être capable de répéter l'audio par synthèse vocale en cours ou précédent;
- l'utilisateur doit être capable de couper l'audio par synthèse vocale;
- utilisateur doit être capable d'activer/de désactiver la fonction de synthèse vocale;
- la langue de l'audio par synthèse vocale doit être la même que celle qui a été réglée pour l'interface utilisateur du téléviseur sauf si cela est signalé différemment. Le

dispositif/moteur de synthèse vocale peut choisir de prononcer le texte ou d'indiquer une défaillance dans le cas où il ne prend pas en charge la langue signalée;

- l'audio par synthèse vocale peut ne pas avoir besoin de représenter littéralement les informations visuelles associées sur l'écran dans la mesure où la signification des informations visuelles reste intacte.

6.2 Fonctionnalité: dispositif/moteur de synthèse vocale

En principe, le téléviseur ne fournit en sortie au dispositif de synthèse vocale que des chaînes textuelles.

Le dispositif de synthèse vocale doit se conformer à ces sorties:

- il convient qu'un dispositif de synthèse vocale externe soit conçu de manière à être entièrement accessible aux utilisateurs ayant une déficience visuelle sans être dépendant du téléviseur;
- le niveau de volume du dispositif/moteur de synthèse vocale doit pouvoir être modifié par l'utilisateur. Le dispositif de synthèse vocale doit annoncer le nouveau niveau de volume. Il convient que cela soit effectué indépendamment du volume du téléviseur;
- il convient que l'utilisateur soit capable de régler les caractéristiques de la parole, telles que la vitesse, la hauteur, le type de voix, le cas échéant;
- il convient que le moteur de synthèse vocale énonce les abréviations lettre par lettre plutôt que sous la forme d'un mot normal. Exemple: Il convient d'énoncer "TTS" sous la forme "T T S" ("té té esse") au lieu de "tts" ("têtsse") le cas échéant. Le moteur de synthèse vocale peut également prononcer en totalité les acronymes courants, par exemple, "sous" peut être prononcé "sous-titre" le cas échéant. La présente norme n'identifie pas ce qui est compris comme une abréviation, mais laisse plutôt cela à la discrétion des responsables de la mise en œuvre;
- il convient que le moteur de synthèse vocale énonce les nombres d'une manière appropriée au contexte, par exemple sous forme de nombre naturel, chiffre par chiffre, etc.;
- on considère que le moteur de synthèse vocale détermine le contexte.

6.3 Fonctionnalité: téléviseur

Le téléviseur détermine l'interface utilisateur, c'est-à-dire ce qui est affiché sur l'écran et la façon dont le téléviseur interagit avec l'utilisateur. Le téléviseur détermine donc quel texte est envoyé au moteur de synthèse vocale

Il convient que l'utilisateur soit capable de commander par l'intermédiaire du téléviseur les réglages de synthèse vocale tels que le volume, la vitesse, la hauteur, le type de voix.

NOTE En Europe, pour satisfaire aux besoins d'accessibilité de base, on s'attend à ce que le téléviseur soit conforme au SELFC.

6.4 Mise en service: combinaison téléviseur-dispositif de synthèse vocale

Un guide audio par synthèse vocale peut être fourni pendant l'installation d'un récepteur. Il convient que le guide audio par synthèse vocale pour l'installation d'un récepteur soit conforme aux exigences suivantes lorsqu'il est mis en œuvre:

- il convient que les exigences du 6.1 s'appliquent;
- il convient que l'utilisateur soit capable de passer rapidement en audio par synthèse vocale au début de l'installation;
- il convient que toutes les informations essentielles pour exécuter la procédure d'installation soient disponibles sous forme d'audio par synthèse vocale. Il convient de prêter une attention particulière à l'information de l'utilisateur de l'avancement des processus lents tels que les recherches de chaînes.

On peut trouver des suggestions supplémentaires concernant le guide audio par synthèse vocale pendant l'installation du récepteur dans [DTG-UG024].

7 Événements de télévision et données de synthèse vocale

7.1 Contexte de télévision et événements

Le téléviseur peut fournir en sortie des données en réponse à des événements. Des événements peuvent avoir lieu dans des contextes différents. Les contextes suivants peuvent exister dans le téléviseur (noter que cette liste n'est pas obligatoirement exhaustive):

- télévision: l'utilisateur regarde la télévision;
- liste: l'utilisateur explore une liste (de chaînes);
- menu (acc): l'utilisateur explore la structure du menu (d'accessibilité) du téléviseur. Si le menu contient des services d'accessibilité, celui-ci est appelé menu d'accessibilité ou menu acc. Un exemple de service d'accessibilité est la Description audio;
- EPG: l'utilisateur explore le guide électronique de programme;
- veille: le téléviseur est en veille;
- décalage temporel: l'utilisateur peut démarrer l'enregistrement d'un décalage temporel ou lire un enregistrement décalé dans le temps.

Les contextes qui suivent ne font habituellement pas partie du domaine d'application de cette spécification et devront être traités dans les versions futures de cette spécification:

- télévision interactive;
- télévision payante.

Les événements suivants peuvent survenir dans lesdits contextes:

- changement de chaîne: Le téléviseur change de chaîne après que l'utilisateur a lancé un événement de Changement de chaîne, par exemple en appuyant sur une touche de la télécommande;
- informations supplémentaires: Le téléviseur présente des informations supplémentaires sur l'élément représenté. Cet élément peut être l'événement affiché sur l'écran mais peut également être un élément de menu en surbrillance, etc. Des informations supplémentaires peuvent être présentées, par exemple après que l'utilisateur a appuyé sur la touche "info" de la télécommande;
- navigation et sélection: Le téléviseur met en surbrillance un autre élément, par exemple dans une liste ou un menu. L'utilisateur sélectionne un élément;
- changement de contexte: Le téléviseur passe à un autre contexte, par exemple entre télévision et menu;
- message contextuel: Le téléviseur affiche un message sur l'écran dans un certain but, il demande par exemple un numéro d'identification personnel ou il donne des informations d'état.

On s'attend à ce que le téléviseur fournisse en sortie des données après un événement du dispositif de synthèse vocale en utilisant l'interface téléviseur-synthèse vocale. Ces données contiennent une indication de l'événement et des données de synthèse vocale à annoncer.

Les différents contextes du téléviseur et les événements pouvant survenir dans un certain contexte sont représentés graphiquement à la Figure 2. Figure 2 donne la liste des événements dans un certain contexte et montre les transitions possibles entre les différents contextes. Elle montre également que dans le contexte d'un menu il est possible de passer à des sous-menus.

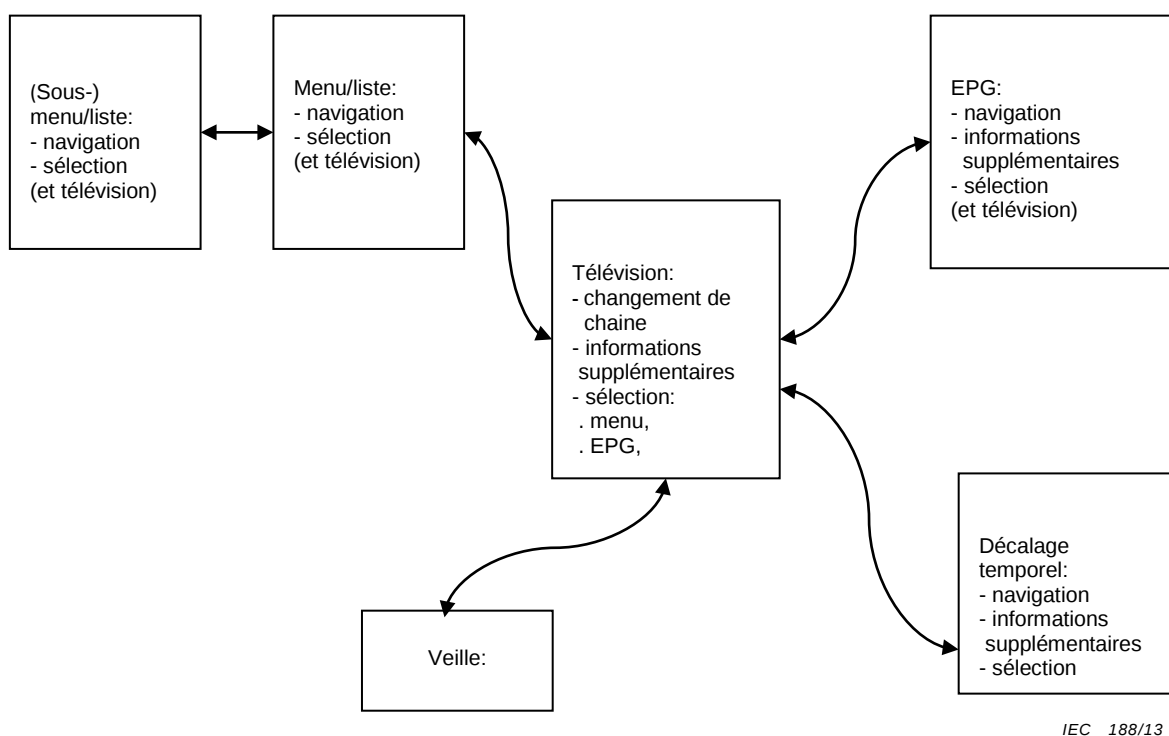


Figure 2 – Diagramme d'état des événements de contexte

7.2 Données de synthèse vocale par événement

7.2.1 Détails

Les paragraphes 7.2.2 à 7.2.6 spécifient les données de synthèse vocale que le téléviseur doit fournir en sortie au dispositif/moteur de synthèse vocale étant donné un événement spécifié.

7.2.2 Changement de chaîne

7.2.2.1 Généralités

Un événement de changement de chaîne fournit en sortie des données concernant la chaîne nouvellement choisie. Un changement de chaîne ne survient généralement que dans le contexte de télévision.

7.2.2.2 Contexte de télévision

Un événement de changement de chaîne dans le contexte de télévision doit au moins fournir en sortie les données de synthèse vocale suivantes pour la chaîne nouvellement sélectionnée, le cas échéant:

- <numéro de chaîne>
- <nom du service>
- <nom d'événement présent (diffusion de télévision numérique)>
- <description audio o/n>.

Un événement de changement de chaîne dans le contexte de télévision peut également fournir en sortie les autres données de synthèse vocale recommandées suivantes le cas échéant:

- e) <heure de début d'événement présent (diffusion de télévision numérique)>
- f) <heure de fin d'événement présent ou durée (diffusion de télévision numérique)>
- g) <nom d'événement suivant (diffusion de télévision numérique)>
- h) <heure de début d'événement suivant (diffusion de télévision numérique)>
- i) <heure de fin d'événement suivant ou durée (diffusion de télévision numérique)>
- j) <description audio oui/non-événement suivant (diffusion de télévision numérique)>.

7.2.3 Informations supplémentaires

7.2.3.1 Généralités

Un événement d'informations supplémentaires doit fournir en sortie les informations supplémentaires disponibles concernant l'élément sélectionné/en surbrillance affiché sur l'écran du téléviseur en tant que données de synthèse vocale.

NOTE: La chaîne de télévision affichée sur le téléviseur est un "élément sélectionné", c'est-à-dire qu'il s'agit de la chaîne sélectionnée.

Des informations supplémentaires apparaissent généralement dans les contextes de télévision, de guide électronique de programme et de décalage temporel.

7.2.3.2 Contexte de télévision

Un événement d'information supplémentaire dans le contexte de télévision peut également fournir en sortie les autres informations recommandées suivantes, en tant que données de synthèse vocale:

- a) <synopsis d'événement (diffusion de télévision numérique)>
- b) <classement parental>
- c) <synopsis d'événement suivant (diffusion de télévision numérique)>
- d) <présence de sous-titres o/n>
- e) <classement parental d'événement suivant (diffusion de télévision numérique)>
- f) <chiffré o/n>
- g) <classification des événements de diffusion de télévision numérique>.

7.2.3.3 Contexte de guide électronique de programme

Un événement d'information supplémentaire dans le contexte de guide électronique de programme peut également fournir en sortie les autres informations recommandées suivantes, en tant que données de synthèse vocale:

- a) <synopsis d'événement (diffusion de télévision numérique)>
- b) <classement parental>
- c) <disponibilité de sous-titres>
- d) <langues disponibles>.

7.2.3.4 Contexte de décalage temporel

Un événement d'information supplémentaire dans le contexte de décalage temporel peut fournir en sortie les données suivantes concernant un événement ou un service à enregistrer ou à lire:

- a) <Synopsis d'événement (diffusion de télévision numérique)>
- b) <Classement parental>
- c) <Disponibilité de sous-titres>

7.2.4 Navigation et sélection

7.2.4.1 Généralités

Un événement de navigation et de sélection doit fournir en sortie le (nouvel) élément en surbrillance en tant que données/audio de synthèse vocale. Le téléviseur doit fournir en sortie un élément sélectionné et/ou modifié en tant que données/audio de synthèse vocale.

L'élément en surbrillance ne peut être entièrement compréhensible dans le contexte qu'avec les autres éléments présents sur l'écran. Dans ce cas, il convient également de fournir en sortie ces autres éléments en tant que données de synthèse vocale. Par exemple: Dans un type de guide électronique de programme matriciel, la position x, y de l'élément en surbrillance représente également une information. Dans ce cas, il convient également de fournir en sortie le contenu de l'axe x et de l'axe y dans la position de l'élément en surbrillance en tant que données de synthèse vocale.

7.2.4.2 Contexte de liste

Les exigences générales en 7.2.4.1 concernant la navigation et la sélection s'appliquent.

7.2.4.3 Contexte de menu (acc)

Les exigences générales en 7.2.4.1 concernant la navigation et la sélection s'appliquent.

7.2.4.4 Contexte de guide électronique de programme

Les exigences générales en 7.2.4.1 concernant la navigation et la sélection s'appliquent. Le téléviseur doit fournir en sortie les éléments individuels énumérés ci-dessous en tant que données de synthèse vocale s'ils sont également affichés sur l'écran:

- a) numéro de chaîne,
- b) nom du service,
- c) description audio (le cas échéant),
- d) sous-titres (le cas échéant),
- e) nom d'événement,
- f) heure de début de l'événement,
- g) heure de fin ou durée de l'événement.

Pour des raisons de convivialité, lorsqu'on entre dans un service du guide électronique de programme, il est recommandé que le téléviseur ne répète pas <numéro de chaîne> et <nom de service>.

Cette liste n'impose pas un ordre spécifique pour énoncer ces éléments.

7.2.4.5 Contexte de décalage temporel

Les exigences générales en 7.2.4.1 concernant la navigation et la sélection s'appliquent. On peut généralement s'attendre à ce que dans ce contexte les commandes lecture, pause, retour arrière, avance rapide, arrêt et enregistrement soient traitées.

7.2.5 Changement de contexte

7.2.5.1 Généralités

Un événement de changement de contexte doit fournir en sortie le nom du contexte en tant que données de synthèse vocale et peut fournir en sortie un guide concernant la façon d'opérer dans le contexte sélectionné. Les données du guide peuvent concerner le type de touches qui peuvent être utilisées et les actions qu'elles déclenchent.

Les données du guide concernant la façon d'opérer dans un certain contexte sont de la responsabilité du fabricant. Exemples de données de guide:

- "Vous êtes dans le contexte de guide électronique de programme. Utilisez les touches de navigation. Une navigation vers le haut/vers le bas vous fera parcourir toutes les chaînes. Une navigation à gauche/à droite vous fera parcourir la programmation horaire."
- "Vous êtes dans le menu télévision. Utilisez haut/bas et OK pour activer le menu."

7.2.5.2 Contexte de télévision

Le même son de synthèse vocale pour un événement de changement de chaîne dans le contexte de télévision doit être fourni en sortie comme audio par synthèse vocale lorsqu'on entre dans le contexte de télévision depuis le contexte de veille.

7.2.5.3 Contexte de liste

Lorsqu'on entre dans le contexte de liste, le téléviseur doit fournir en sortie en tant que données TTS

- <le titre de la liste>
- <le nombre d'éléments de la liste>

7.2.5.4 Contexte de menu (acc)

Lorsqu'on entre dans le contexte de menu (acc), le téléviseur doit fournir en sortie en tant que données TTS:

- <le titre du menu>
- <le nombre d'éléments du menu>

7.2.5.5 Contexte de guide électronique de programme

Lorsqu'on entre dans le contexte de guide électronique de programme, le téléviseur doit fournir en sortie en tant que données TTS:

- <l'élément en surbrillance>.

Le téléviseur peut également fournir en sortie en tant que données TTS:

- <une vue d'ensemble du guide électronique de programme>: brève description du guide électronique de programme
- <Réglage de filtre du guide électronique de programme appliqué>.

Le téléviseur doit fournir en sortie en tant que données de synthèse vocale si <l'élément en surbrillance> est un événement (de diffusion de télévision numérique) et s'il est affiché sur l'écran.

- a) <numéro de chaîne>
- b) <nom du service>
- c) <description audio o/n>
- d) <sous-titres o/n>
- e) <nom d'événement présent (diffusion de télévision numérique)>
- f) <heure de début d'événement présent (diffusion de télévision numérique)>
- g) <heure de fin d'événement présent ou durée (diffusion de télévision numérique)>

7.2.5.6 Contexte de veille

Lors de l'entrée en veille, il convient que l'utilisateur soit averti que le récepteur passe en veille.

7.2.5.7 Contexte de décalage temporel

Lors de l'entrée en décalage temporel, il convient que l'utilisateur soit averti que le récepteur passe dans un contexte de décalage temporel.

7.2.5.8 Contextes actuellement non pris en charge

Les contextes de télévision interactive et de télévision payante ne font actuellement pas partie du domaine d'application de cette spécification. Il est recommandé d'informer l'utilisateur de l'entrée dans un contexte non pris en charge et d'informer l'utilisateur de la façon de quitter ledit contexte.

7.2.6 Message contextuel

Pour un utilisateur ayant une déficience visuelle, les messages contextuels constituent souvent une barrière d'accès insurmontable, sauf s'ils sont entièrement pris en charge par le dispositif de synthèse vocale. L'utilisateur ayant une déficience visuelle doit être informé de l'apparition du menu contextuel, de son contenu en totalité et de l'action qui est demandée à l'utilisateur pour poursuivre. Les messages contextuels sont souvent modaux, empêchant ainsi l'utilisateur d'exécuter une commande et une navigation normales tant que le menu contextuel n'a pas été supprimé.

L'utilisateur doit être informé de tous les avertissements et notifications, par exemple les problèmes de syntonisation ou une commande par numéro d'identification personnel. Lorsqu'un message contextuel apparaît sur l'écran, il convient de le fournir en sortie en tant que données de synthèse vocale.

- <notification du message contextuel>
- <description du message contextuel>

L'audio par synthèse vocale concernant les informations de changement d'état doit être fourni régulièrement et après achèvement. Il n'est pas exigé de fournir l'audio par synthèse vocale à chaque mise à jour.

8 Profils de synthèse vocale

8.1 Profil de base

Un téléviseur avec synthèse vocale incorporée ou une combinaison téléviseur-synthèse vocale doit au moins prendre en charge le profil de base. Le profil de base doit prendre en charge l'audio par synthèse vocale comme spécifié dans les articles concernant l'événement et le contexte pour les événements et contextes suivants:

- événement de changement de chaîne:
 - contexte de télévision;
- événement de navigation et sélection:
 - contexte de liste,
 - contexte de menu acc,
 - contexte de guide électronique de programme;
- événement d'informations complémentaires:
 - contexte de télévision;
- événement de changement de contexte:
 - contexte de liste,
 - contexte de veille,
 - contexte de menu acc,

- contexte de guide électronique de programme;
- événement de message contextuel.

NOTE Un événement de changement de chaîne peut être considéré comme un cas particulier d'un événement de navigation.

8.2 Profil principal

Le profil principal doit prendre en charge l'audio par synthèse vocale comme spécifié dans les articles concernant l'événement et le contexte pour les événements et contextes suivants:

- événement de changement de chaîne:
 - contexte de télévision;
- événement de navigation et sélection:
 - contexte de liste,
 - contexte de menu acc,
 - contexte de menu,
 - contexte de guide électronique de programme;
- événement d'informations complémentaires:
 - contexte de télévision;
- événement de changement de contexte:
 - contexte de liste,
 - contexte de veille,
 - contexte de menu acc,
 - contexte de menu,
 - contexte de guide électronique de programme;
- événement de message contextuel.

8.3 Profil amélioré

Le profil amélioré doit prendre en charge l'audio par synthèse vocale comme spécifié dans les articles concernant l'événement et le contexte pour les événements et contextes suivants:

- événement de changement de chaîne:
 - contexte de télévision;
- événement de navigation et sélection:
 - contexte de liste,
 - contexte de menu acc,
 - contexte de menu,
 - contexte de guide électronique de programme,
 - contexte de décalage temporel;
- événement d'informations complémentaires:
 - contexte de télévision,
 - contexte de guide électronique de programme,
 - contexte de décalage temporel;
- événement de changement de contexte:
 - contexte de liste,
 - contexte de veille,
 - contexte de menu acc,

- contexte de menu,
- contexte de guide électronique de programme,
- contexte de décalage temporel;
- événement de message contextuel.

8.4 Résumé

Le Tableau 1 fournit une vue d'ensemble des différents profils.

Tableau 1 – Vue d'ensemble des profils

Événement	Contexte	Profil de base	Profil principal	Profil amélioré
changement de chaîne	contexte de télévision	X	X	X
événement de navigation et sélection	contexte de liste	X	X	X
	contexte de menu acc	X	X	X
	contexte de menu		X	X
	contexte de guide électronique de programme	X	X	X
	Contexte de décalage temporel			X
événement d'informations complémentaires	contexte de télévision	X	X	X
	contexte de guide électronique de programme			X
	Contexte de décalage temporel			X
événement de changement de contexte	contexte de liste	X	X	X
	contexte de veille	X	X	X
	contexte de menu acc	X	X	X
	contexte de menu		X	X
	contexte de guide électronique de programme	X	X	X
	contexte de décalage temporel			X
événements de message contextuel	"tous les contextes dans le profil"	X	X	X

Bibliographie

IEC/TR 62678:2010, *Audio, video and multimedia systems and equipment activities and considerations related to accessibility and usability*

ISO/IEC 13818-1, *Information technology – Generic coding of moving pictures and associated audio information: Systems*

ISO/IEC TR 29138-3:2009, *Information technology – Accessibility considerations for people with disabilities – Part 3: Guidance on user needs mapping*

DTG-UG024, *Digital TV group – UK digital TV usability and accessibility guidelines, including text to speech (u-book)*, version 1.0, September 2011

ETSI EN 300 468, *Digital video broadcasting (DVB) – Specification for service information (SI) in DVB systems*

ETSI TR 101 211, *Digital video broadcasting (DVB); Guidelines on implementation and usage of service information (SI)*

ETSI TS 101 154, *Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream*

ATSC A/53 part 5: *Digital television standard – AC-3 audio system characteristics: 2010*

ATSC A/53 part 6: *Digital television standard – Enhanced AC-3 audio system characteristics: 2010*

ARIB STD-B10: *Service Information for Digital Broadcasting System*

ARIB STD-B32: *Video Coding, Audio Coding and Multiplexing Specifications for Digital Broadcasting*

SELFCE, *Industry Self-Commitment, to improve the accessibility of digital TV receiving equipment sold in the European Union*, Brussels, November 30, 2007. Disponible à l'adresse <<http://www.digitaleurope.org/IndustryGroups/TechnicalandRegulatoryGroup/eInclusion/Documents.aspx>>

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

3, rue de Varembé
PO Box 131
CH-1211 Geneva 20
Switzerland

Tel: + 41 22 919 02 11
Fax: + 41 22 919 03 00
info@iec.ch
www.iec.ch