

# INTERNATIONAL STANDARD

## NORME INTERNATIONALE

**Text-to-speech for television – General requirements**

**Synthèse vocale pour télévision – Exigences générales**

[IEC 62731-1:2013](https://standards.iteh.org/standards/sis/re/99caed-51c4-40aa-a2a0-7e454c1dae7c/iec-62731-2013)

<https://standards.iteh.org/standards/sis/re/99caed-51c4-40aa-a2a0-7e454c1dae7c/iec-62731-2013>



**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 Varembe  
CH-1211 Geneva 20  
Switzerland

Tel.: +41 22 919 02 11  
Fax: +41 22 919 03 00  
[info@iec.ch](mailto:info@iec.ch)  
[www.iec.ch](http://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](http://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](http://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](http://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/iec-](http://webstore.iec.ch/csc/iec-)

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

# 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: <https://www.iec.ch/standards-shop>
- 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.

Withhold

iTeh STANDARD PREVIEW  
(standards.iteh.ai)

<https://standards.iteh.ai/catalog/standards/sist/99caed-51c4-40aa-a2a0-7e454c1dae7c/iec-62731-2013>

# 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:

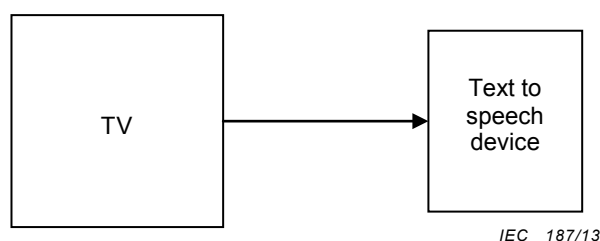
- a) channel number, name and event information – key for a user to identify which service has been selected;
- b) availability of audio description – key for a user to know about the availability of this service feature;
- c) availability of subtitles – key for a user to know about the availability 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).