

INTERNATIONAL STANDARD

Text-to-speech for television – General requirements

(standards.iteh.ai)

IEC 62731:2018

<https://standards.iteh.ai/catalog/standards/sist/d411fba1-a97b-4a59-b1f9-4cebb3232bd5/iec-62731-2018>



THIS PUBLICATION IS COPYRIGHT PROTECTED

Copyright © 2018 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.

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

Tel.: +41 22 919 02 11
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.

IEC Catalogue - webstore.iec.ch/catalogue

The stand-alone application for consulting the entire bibliographical information on IEC International Standards, Technical Specifications, Technical Reports and other documents. Available for PC, Mac OS, Android Tablets and iPad.

IEC publications search - webstore.iec.ch/advsearchform

The advanced search enables 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 It also gives information on projects, replaced and withdrawn publications.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 online and also once a month by email.

Electropedia - www.electropedia.org

The world's leading online dictionary of electronic and electrical terms containing 21 000 terms and definitions in English and French, with equivalent terms in 16 additional languages. Also known as the International Electrotechnical Vocabulary (IEV) online.

IEC Glossary - std.iec.ch/glossary

67 000 electrotechnical terminology entries in English and French extracted from the Terms and Definitions clause of IEC publications issued since 2002. Some entries have been collected from earlier publications of IEC TC 37, 77, 86 and CISPR.

IEC 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: sales@iec.ch

<https://standards.iteh.ai/catalog/standards/sist/d411bc1-a97b-4a59-b1f9-4cebb3232bd5/iec-62731-2018>

IEC STANDARDS PREVIEW
(standards.iteh.ai)

IEC 62731

INTERNATIONAL STANDARD

Text-to-speech for television – General requirements

(standards.iteh.ai)

IEC 62731:2018

<https://standards.iteh.ai/catalog/standards/sist/d411fba1-a97b-4a59-b1f9-4cebb3232bd5/iec-62731-2018>

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

ICS 33.160.25; 33.160.99

ISBN 978-2-8322-5125-6

Warning! Make sure that you obtained this publication from an authorized distributor.

CONTENTS

FOREWORD.....	3
1 Scope.....	5
2 Normative references	5
3 Terms, definitions and abbreviated terms	5
3.1 Terms and definitions.....	5
3.2 Abbreviated terms.....	6
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.....	8
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.....	11
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.....	16
8 TTS profiles.....	16
8.1 Basic profile.....	16
8.2 Main profile.....	16
8.3 Enhanced profile.....	17
8.4 Summary	18
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.

This second edition cancels and replaces the first edition published in 2013. This edition constitutes a technical revision.

This edition includes the following significant technical changes with respect to the previous edition:

- a) in 6.2, the levels of announcement quality were revised as well as considerations for ways in which device users can provide service providers with feedback on incorrectly announced terms.
- b) in 6.3, the following TV functionality was added: the TV can receive updated words, associated conversions and updated conversion rules for the TTS engine via a network connection.

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

FDIS	Report on voting
100/2989/FDIS	100/3013/RVD

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

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

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

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

A bilingual version of this publication may be issued at a later date.

iTeh STANDARD PREVIEW **(standards.iteh.ai)**

[IEC 62731:2018](#)

<https://standards.iteh.ai/catalog/standards/sist/d411fba1-a97b-4a59-b1f9-4cebb3232bd5/iec-62731-2018>

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 document 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 document refers to TV, this will be shorthand for all such receivers.

This document 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

There are no normative references in this document.

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

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

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at <http://www.electropedia.org/>
- ISO Online browsing platform: available at <http://www.iso.org/obp>

3.1.1 context

one specific function of a TV

EXAMPLE: Watching TV, EPG.

3.1.2 DTV broadcast event classification

general category of programme/event content, or its classification

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

3.1.3 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").

3.1.4

event

trigger to start an action

3.1.5

list

collection of items

3.1.6

menu

subsequent order of items

3.1.7

receiver

device capable of receiving or handling digital television signals

3.1.8

service

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

3.1.9

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 document uses the term subtitles throughout.

3.1.10

TTS audio

audio generated by the TTS engine based on the TTS data

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

3.1.11

priority audio information

audio information for immediate output typically related to emergency information

3.1.12

TTS data

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

3.2 Abbreviated terms

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 document 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, for example a TV or an 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 document, 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 tasks similar to those they would with a graphical UI.

The main features of this document 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 document, mandatory requirements are specified; optional and informative features are also included.

A claim of conformity with this document 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 each feature of a product. While options for user configurations may be provided, the product shall meet the mandatory requirements.

5 User requirements of visually impaired people

5.1 Users' needs

This subclause 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 document. Item c) further relates to the fact that, in some countries, foreign language programmes are 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 document 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 that 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 of 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 one-time 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 such as 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, which are 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 document. 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.

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

- f) 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

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.

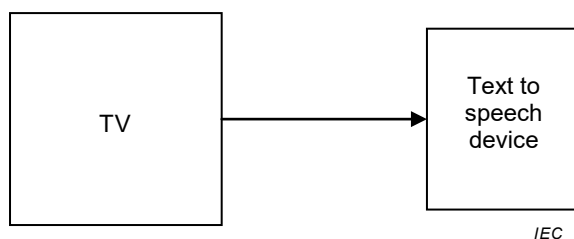


Figure 1 – TV – TTS device system diagram

The functionality requirements for a TV with TTS combination are:

- the TV, in principle, only outputs text strings towards the TTS device;
- 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;