

---

---

**Information technology — User  
interface component accessibility —**

**Part 23:  
Visual presentation of audio  
information (including captions and  
subtitles)**

*Technologies de l'information — Accessibilité du composant interface  
utilisateur —*

*Partie 23: Présentation visuelle d'informations sonores*





**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2018

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
CP 401 • Ch. de Blandonnet 8  
CH-1214 Vernier, Geneva  
Phone: +41 22 749 01 11  
Fax: +41 22 749 09 47  
Email: [copyright@iso.org](mailto:copyright@iso.org)  
Website: [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

	Page
<b>Foreword</b> .....	<b>vi</b>
<b>Introduction</b> .....	<b>vii</b>
<b>1 Scope</b> .....	<b>1</b>
<b>2 Normative references</b> .....	<b>1</b>
<b>3 Terms and definitions</b> .....	<b>1</b>
<b>4 Framework for the visual presentation of audio information</b> .....	<b>3</b>
4.1 Purpose.....	3
4.2 Motivation.....	4
4.3 Locations of presentations.....	4
4.4 Modes of presentations.....	4
4.5 Modes of access.....	5
4.6 Modes of display.....	5
4.7 Levels of importance.....	5
4.7.1 General.....	5
4.7.2 Essential information.....	5
4.7.3 Significant information.....	6
4.7.4 Helpful information.....	6
4.7.5 Unhelpful information.....	6
<b>5 Applicability of requirements and recommendations</b> .....	<b>6</b>
5.1 Predictable audio contents.....	6
5.2 Unpredictable audio contents.....	6
<b>6 Production of visual presentations of audio information</b> .....	<b>7</b>
6.1 Consideration to match the intended meaning.....	7
6.2 Ease of understanding of visual presentations of audio information.....	7
6.3 Consideration of output devices.....	7
6.4 Verification of visual presentations of audio information with the intended output devices.....	7
6.5 Connecting visual presentations of audio information data with content data.....	7
6.6 Combining multiple visual presentations of audio information.....	8
6.7 Update of visual presentations of audio information data.....	8
6.8 Evaluation.....	8
6.9 Evaluations including contribution of typical users.....	8
<b>7 Visual design</b> .....	<b>8</b>
7.1 General.....	8
7.2 Personalization.....	8
7.3 Engagement.....	8
7.4 Synchronization of presentations.....	9
7.5 Avoidance of information obstruction.....	9
7.6 Font size.....	9
7.7 Font type.....	10
7.8 Font face.....	10
7.9 Upper, lower, and mixed case letters.....	10
7.10 Contrast and use of colour.....	10
7.11 Speed.....	11
7.12 Number of lines.....	11
7.13 Spacing between characters and lines (kerning and leading).....	11
7.14 Correct punctuation.....	11
7.15 Spacing between words and phrases.....	12
7.16 Transitions between presentations.....	12
7.17 Sentence segmentation.....	12
7.18 Indication of sentence breaks over multiple visual presentations.....	12
7.19 Additional duration for location change.....	13

7.20	Modes of display.....	13
<b>8</b>	<b>Visual alternative container (VAC).....</b>	<b>13</b>
8.1	General.....	13
8.2	VAC position and area.....	13
<b>9</b>	<b>Describing speech.....</b>	<b>13</b>
9.1	Describing verbal content.....	13
9.2	Grammar.....	14
9.3	Vulgar verbal content and slang.....	14
9.4	Language variation.....	14
9.5	Foreign accents.....	14
9.6	Indiscernible audio content.....	14
9.7	Spelling.....	14
9.8	Abbreviations.....	15
9.9	Homophones, homonyms, homographs, heteronyms, and heterographs.....	15
9.10	Long speech.....	15
9.11	Describing multiple simultaneous information.....	15
9.12	Confirmation by content producers when producing visual presentations of speech.....	15
9.13	Sources of information.....	15
<b>10</b>	<b>Non-speech information (NSI).....</b>	<b>16</b>
10.1	General.....	16
10.2	Describing NSI.....	16
10.3	Correct description of NSI.....	16
10.4	Well-known sound descriptions.....	16
10.5	Onomatopoeia.....	16
10.6	Sound effects in speech.....	16
10.7	Censored language.....	17
10.8	Paralinguistic sound effects.....	17
10.9	Discrete and sustained sound effects.....	17
10.10	Confirmation by content producers when producing visual presentations of NSI.....	17
10.11	Sources of information.....	17
<b>11</b>	<b>Music.....</b>	<b>18</b>
11.1	General.....	18
11.2	Describing presence of music.....	18
11.3	Describing the reason or purpose for the music.....	18
11.4	Provide information that identifies the music.....	18
11.5	Clarification of music descriptions.....	19
11.6	Presentation of lyrics.....	19
11.7	Distinction of lyrics from speech.....	19
11.8	Confirmation by content producers when producing visual presentations of music.....	20
11.9	Sources of lyrics.....	20
<b>12</b>	<b>Emotions.....</b>	<b>20</b>
12.1	General.....	20
12.2	Describing intended emotional nuance.....	20
12.3	Describing the reason or purpose for the emotional nuance.....	20
12.4	Confirmation by content producers when producing visual presentations of audio information of emotions.....	21
<b>13</b>	<b>Silence.....</b>	<b>21</b>
13.1	General.....	21
13.2	Describing intentional silence.....	21
13.3	Describing prolonged silence.....	21
13.4	Describing the reason or purpose for the silence.....	21
<b>14</b>	<b>Identifying speakers.....</b>	<b>21</b>
14.1	General.....	21
14.2	Means of identifying speakers.....	22
14.3	Identifying speakers by word.....	22

14.4	Identifying speakers by pictogram .....	22
14.5	Identifying speakers by colour .....	22
14.6	Identifying speakers by position .....	22
14.7	Identifying change of speakers by changing position .....	22
14.8	Multiple visual presentations of audio information .....	23
<b>15</b>	<b>Evaluating quality of visual presentations of audio information .....</b>	<b>23</b>
15.1	General .....	23
15.2	Quality review process .....	23
<b>Annex A</b>	<b>(informative) Evaluation index and references of guidelines and reports for</b>	
	<b>accessibility .....</b>	<b>24</b>
<b>Bibliography</b>	.....	<b>26</b>

## Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see [www.iso.org/directives](http://www.iso.org/directives)).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see [www.iso.org/patents](http://www.iso.org/patents)).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL [www.iso.org/iso/foreword.html](http://www.iso.org/iso/foreword.html).

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 35, *User interfaces*.

A list of all parts in the ISO/IEC 20071 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at [www.iso.org/members.html](http://www.iso.org/members.html).

## Introduction

Captions and/or subtitles for audio content provide visual alternatives for audio information in audiovisual content. This document provides requirements and recommendations on the production and design of the visual presentation of audio information (including captions and subtitles) that supports users who cannot make use of the audio content.

The use of this document helps to support universal and inclusive media content production practices. It provides guidance for producers, exhibitors, or distributors of audio content (including the medium of distribution and the medium of delivery) to support the accessibility and usability of visual alternatives of audio content.

Different jurisdictions have different expectations of what belongs in a caption or a subtitle. From the point of view of the user, what is important is receiving the information in an accessible design, not the technological means of its delivery. This information can include text conveying speech, sound information, verbatim transcription of the spoken word content, translations of the spoken word content, etc. This document uses “visual presentations of audio information” to include all audio information needed to be made accessible for some users.

Standardized guidance for producing visual presentations of audio information is important to meet a variety of needs. For example, it is important to recognize acceptable values for specifying typography variables such as the letter size and/or number of characters in visual alternatives that rely on text.

Providing visual presentations of audio information (including captions and subtitles) can be beneficial to all, and in particular to diverse users who cannot hear or understand the audio content in diverse contexts, including: persons with hearing loss, persons who are deaf or hard of hearing, persons with learning difficulties or cognitive disabilities, persons watching a movie in a non-native language, persons who need the content to be in another language, persons who cannot hear the audio content due to environmental conditions, or circumstances where the sound is not accessible (e.g. noisy surroundings), the sound is not available (e.g. muted, no working speakers), or the sound is not appropriate (e.g. a quiet library). Although this guidance acknowledges the need of visual presentations of audio information to provide non-visual presentations for diverse users, it does not include guidance for producing non-visual presentations, such as spoken captions/subtitles (see ISO/IEC TS 20071-25 for further reference) and tactile displays (e.g. Braille). The production, delivery, and exhibition of visual presentations of audio information based on this standard are not intended to interfere with or change the meaning of the audio content.

The production, delivery and exhibition of visual presentations of audio information vary according to the time and methodology of production, the technology used for its production, the system of delivery, and the display (including the brand and model of the display).





# Information technology — User interface component accessibility —

## Part 23: Visual presentation of audio information (including captions and subtitles)

### 1 Scope

This document provides guidance for producers, exhibitors, and distributors on the visual presentation of alternatives to audio information in audiovisual content, such as captions/subtitles.

This document provides requirements and recommendations that are intended to support users who are not able to use the audio information, prefer to use a visual representation of audio information, or prefer both audio and visual presentations.

**NOTE** Many users do not have a choice, for instance, when in a noisy environment (e.g. bar, restaurant, etc.). In these situations, the user does not select a visual presentation of audio information but is offered the content with captions/subtitles.

This document acknowledges the various needs and preferences of viewers (end users) as well as the different approaches to visual presentation of audio information. It applies to all presentations of visual alternatives to audio information intended to be presented as captions/subtitles.

This document does not apply to the presentation devices or transmission mechanisms used to deliver the content or visual presentations of audio information. These devices could include, but are not limited to: televisions, computers, wireless devices, projection equipment, DVD and home cinema equipment, video game consoles, and other forms of user interfaces technology. This document does not apply to transcoding files and formats for the various video outputs.

This document gives guidance on visual presentations which are delivered in the same language as in the audio (i.e., intra-lingual captions/subtitles) and visual presentations which are translated into a different language (i.e., inter-lingual captions/subtitles). This document does not apply to the specific process of language translation.

This document helps to improve accessibility. This document does not establish requirements on specific industries (e.g. television broadcasting, motion pictures) nor is it intended to supersede specific international standards within their domain.

### 2 Normative references

There are no normative references in this document.