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

**Part 21:
Guidance on audio descriptions**

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

Partie 21: Directives sur les descriptions audio

IECNORM.COM : Click to view the full PDF of ISO/IEC TS 20071-21:2015



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2015, Published in Switzerland

All rights reserved. Unless otherwise specified, 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
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword	v
Introduction	vi
1 Scope	1
2 Terms and definitions	1
2.1 General terms	2
2.2 Audio-specific terms	2
2.3 Language of presentation terms	3
2.4 Audio description production terms	4
3 Framework and process considerations for audio description	5
3.1 General	5
3.1.1 Alternate names for audio description	5
3.1.2 Motivation for audio description and spoken subtitles	5
3.2 Types of audio description	5
3.2.1 Live audio description	5
3.2.2 Pre-mixed audio description	5
3.2.3 Harmonious commentary	6
3.2.4 Extended audio description	6
3.3 Creating audio description	6
3.3.1 Narrator preparation	6
3.3.2 Content genre and programme category	7
3.3.3 Scripting	7
3.3.4 Writer considerations	7
3.3.5 Recording	7
3.3.6 Volume	7
3.3.7 Audio technician considerations	8
3.3.8 Inclusion of consumers in the process of creating audio description	8
3.3.9 Inclusion of consumers in evaluating audio description	8
3.4 Live audio description	8
3.4.1 Live narration considerations	8
3.4.2 Preparing for live narration	9
3.4.3 Narrator transitions in live audio	9
3.5 Pre-recorded descriptions	10
3.5.1 Creation of pre-mixed audio after main production process	10
3.5.2 Audio description team for pre-mixed audio	10
3.5.3 Evaluation of pre-mixed audio	10
3.5.4 Voices for pre-mixed audio	10
3.5.5 Sound engineering for pre-mixed audio	10
4 Guidance on audio description	10
4.1 General considerations	10
4.1.1 User considerations	10
4.1.2 Consistency across technologies	11
4.1.3 Consistency within a programme and series	11
4.1.4 Placement of descriptions	11
4.1.5 Use of pauses	11
4.2 Developing audio description	11
4.2.1 Clarity in audio description	11
4.2.2 Creativity in audio description	12
4.2.3 Narrating/reading audio description script	12
4.3 Styles of narration	13
4.3.1 Newsreader style	13
4.3.2 Commentator style	13
4.3.3 First person	13
4.3.4 Third person	14

4.4	Levels of importance.....	14
4.4.1	Using importance levels	14
4.4.2	Essential information.....	14
4.4.3	Significant information.....	14
4.4.4	Helpful information.....	15
4.4.5	Irrelevant information	15
4.5	General guidance on descriptions.....	16
4.5.1	Description of sounds.....	16
4.5.2	Logos/credits/titles	16
4.5.3	On-screen text.....	16
4.6	Guidance on identifying objects.....	16
4.6.1	Identifying objects.....	16
4.6.2	Describing colour.....	17
4.6.3	Describing visual effects.....	17
4.7	Guidance on identifying persons.....	17
4.7.1	Identifying characters or places by name	17
4.7.2	Physical appearance.....	18
4.7.3	Race/ethnic origin.....	19
4.7.4	Gender-related.....	20
4.7.5	Disabilities.....	20
4.7.6	Age.....	21
4.8	Guidance on relationships.....	22
4.8.1	General guidance on explicit content.....	22
4.8.2	General guidance on relationships.....	22
4.8.3	Establishing place/setting/time of day.....	22
4.8.4	Interaction between animated characters/objects and real actors.....	23
4.9	Guidance on parts of speech	23
4.9.1	Descriptive verbs.....	23
Bibliography		24

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).

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).

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 meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: [Foreword - Supplementary information](#)

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC 35, *User interfaces*.

ISO/IEC 20071 consists of the following parts, under the general title *Information technology — User interface component accessibility*:

- *Part 11: Guidance for alternative text for images* [Technical Specification]
- *Part 21: Guidance on audio description* [Technical Specification]

Introduction

This part of ISO/IEC 20071 provides audio description developers and practitioners with guidance in creating effective content describing audiovisual material in an auditory-only modality, the style or manner in which audio description is delivered, the audio description script and script time cues, in relation to the original content. The circumstances to which audio description applies include recorded video, broadcast and broadband television, cinema, live or recorded drama, museum and art gallery exhibits, heritage tours, news, and comedies. Some of this content is static and unchanging, some is dynamic, some is expressed in a visual modality alone, and some in a combination of visual and auditory modalities.

An example of mainstream use of audio description techniques can be found in sports commentary. It can be enjoyed by diverse audiences and uses personalities, emotional expression, and reaction to content to express qualities of the original content.

Audiovisual material is produced for people who can see and hear. It is also typically produced in a specific language for a particular audience. Translations of the original auditory content and associated audio description can be provided in different languages. Providing audio description and translating into other languages can be beneficial to diverse users in diverse contexts, including persons who are blind or persons with low vision, persons with learning difficulties or cognitive impairments and, in the case of translation, persons who need the content to be in another language. Translating subtitles and other on-screen text and providing those in auditory form can also be beneficial.

Producing audio description can be thought of in a similar way to language translation in that it involves several kinds of subjective decisions, for example, the use of vocabulary, tone, speed, intonation, and expression. The accuracy and quality that audio describers or script writers can achieve is determined by matching the factors above with the requirements of the original content.

The creation and delivery of audio description based on the guidance in this part of ISO/IEC 20071 is not intended to interfere with or change the meaning of the original content, or the freedom of speech, expression, or opinion of rights holders.

A further benefit of providing audio description can be the raising of awareness of barriers experienced by blind persons and persons with low vision and other beneficiaries in accessing audiovisual content. The use of this part of ISO/IEC 20071 can also help embed universal and inclusive designed media content production practices in educational programmes (such as film, theatre performances, museum and art gallery exhibits, and other media).

Information technology — User interface component accessibility —

Part 21: Guidance on audio descriptions

1 Scope

This part of ISO/IEC 20071 provides recommendations for describing audiovisual content in an auditory modality for use in recorded video presentations, broadcast television, cinema, live or recorded drama, museum and art gallery exhibits, heritage tours, news, and comedies, regardless of the language and technology being used to transmit and present the recorded or live audiovisual content.

NOTE 1 There are many secondary users of audio description, but the primary and intended users of audio description are blind persons or persons with low vision and their friends and family.

This part of ISO/IEC 20071 provides guidance on the subjective nature of creating audio description.

It also provides guidance for audio description developers (script writers, voice narrators, and organizations or groups responsible for delivering audio description) in faithfully and accurately representing audiovisual content.

NOTE 2 It might not always be possible to provide an equivalent experience due to limitations in the amount of time available for audio description depending on the content being described.

This part of ISO/IEC 20071 applies to describing audiovisual content and does not consider the devices or transmission mechanisms used to deliver the content or the audio description. These devices include, but are not limited to, televisions, computers, wireless devices, projection equipment, DVD and home cinema equipment and other forms of user interface technology. Therefore, this part of ISO/IEC 20071 does not consider transcoding files for the various video outputs.

NOTE 3 Technical matters of transmission and distribution are covered by other International Standards (e.g. MPEG standards and other technical international standards such as IEC 62731).

This part of ISO/IEC 20071 provides guidance on how to approach user preferences, available styles and flexibility in approaches to audio description.

It does not apply to static images contained in electronic documents (see ISO/IEC 20071-11 for guidance on text alternatives for images).

It applies to auditory presentations intended to be transmitted or delivered simultaneously to the original audiovisual content.

NOTE 4 Limitations experienced by broadcasters or people in synchronous environments might be reduced online or in asynchronous environments.

2 Terms and definitions

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

2.1 General terms

2.1.1

audiovisual content

<audio description> visual and/or auditory element of recorded video presentations, live performances, or static or dynamic events to which audio description applies

Note 1 to entry: This generic term is used where only the visual modality might apply, for example, in art gallery or museum exhibits.

Note 2 to entry: The range of circumstances to which this term applies includes recorded video presentations, broadcast and broadband television, cinema, live or recorded drama, museum and art gallery exhibits, heritage tours, news, and comedies.

2.1.2

video (recorded)

<audio description> combination of auditory and visual content intended to be presented together in a synchronized manner

2.1.3

programme

<audio description> complete unit of a recorded or live video broadcast

2.1.4

programme category

<audio description> classification of programmes

Note 1 to entry: Some programme categories include documentary, news and information, and drama.

Note 2 to entry: Programme categories are not necessarily mutually exclusive.

2.1.5

genre

<audio description> classification of a film, programme, drama or an event from an artistic perspective

Note 1 to entry: Genres include, but are not limited to, history, action, horror, romance, and comedy.

Note 2 to entry: Programme genres are not necessarily mutually exclusive.

2.1.6

creative

<audio description> creation and addition of new original content to existing content

2.1.7

harmony

<audio description> natural fit between the audio description script and its delivery and the original content

2.2 Audio-specific terms

2.2.1

audio track

auditory components of a video other than audio description components

2.2.2

audio description

descriptive audio

Note 1 to entry: Audiovisual content is described in an auditory modality.

Note 2 to entry: Audio description can be used to describe sound not easily identified or sound coming from an unknown source or location.

Note 3 to entry: Audio description can also be used to describe locations, directions, and objects.

2.2.3

basic audio description

audio description provided along with the original audiovisual content without lengthening the time of presentation.

2.2.4

extended audio description

audio description provided with the original audiovisual content that lengthens the total time of the presentation or event

2.2.5

audio introduction

complementary auditory content included before the original audiovisual content begins

2.3 Language of presentation terms

2.3.1

original language

native language in which audiovisual content is produced

2.3.2

re-voicing

dubbing

voice-over

secondary audio version of a film or video produced by voice talent, translators, and dialogue writers in a language other than the original language of the film or video

Note 1 to entry: When dubbing, this is timed to match the voicing or mouth movements of animated characters or the original actors.

2.3.3

subtitles

transcription or translation of the dialogue, suitable for when the sound is available but not understood

Note 1 to entry: Specific guidance on delivering subtitles, on-screen text, and embedded text will be provided in a forthcoming addition to ISO/IEC 20071.

2.3.4

spoken subtitles

audio subtitles

spoken captions

subtitles on audiovisual content that are read aloud and spoken over the audio in subtitled audiovisual content

2.3.5

spoken subtitles with audio description

audiovisual content which is both subtitled and audio described through the narrator reading the translation after stating that a subtitle appears

Note 1 to entry: Subtitles can also be identified through alternative voicing, synthetic speech, or sound indicator.

Note 2 to entry: Spoken subtitles often use a different voice to the voice of the audio describer to avoid confusion between the dialogue and the description.

Note 3 to entry: Spoken subtitles and audio description can be available to users separately or together.

Note 4 to entry: Spoken subtitles are synchronized with the text on screen.

2.4 Audio description production terms

2.4.1

live production

production process that involves creating audio description in real time along with the creation and delivery of the original audiovisual content

Note 1 to entry: Live productions are typically experienced initially in real time, but can also be recorded for later redistribution.

2.4.2

post script/pre-production audio description

audio description considered after the script has been developed, but before the event takes place

Note 1 to entry: An example of this is the text of a play.

Note 2 to entry: This allows the script to be modified to include redundancies in audio and visual content (e.g. "Pass it to me" could be modified to "Pass my coffee, Jane").

2.4.3

pre-mixed production

content including audio description that is created and recorded prior to transmission

2.4.4

co-produced

<audio description> produced at the same time as the rest of the audiovisual content

2.4.5

post-produced

<audio description> produced after the audiovisual content

2.4.6

audio describing

delivering a script designed to meet the needs of persons wishing to have access to audiovisual content

2.4.7

voice narrator

narrator

describer

audio describer

<audio description> person(s) and/or technology by which audio description is voiced

Note 1 to entry: A human voice narrator can also be referred to as a narrator, describer, or an audio describer. For the purpose of this part of ISO/IEC 20071, these terms are synonymous and referred to as voice talent.

2.4.8

machine narrator

<audio description> technology which delivers text-based audio description

Note 1 to entry: To date, machine-based narration has not been as acceptable to persons who are blind or persons with low vision as voiced audio description.

2.4.9

writer

<audio description> person(s) who writes the audio description script

3 Framework and process considerations for audio description

3.1 General

3.1.1 Alternate names for audio description

Terms used in this part of ISO/IEC 20071 vary according to country, language, region, and type of audiovisual content. They include audio description, described video, described narrative, captioning for the blind, and others. For the purpose of this part of ISO/IEC 20071, these terms are synonymous.

NOTE In some jurisdictions, there are precise usages defined for one or more of these terms. Individuals can consult their own country's regulations for the locally appropriate terminology.

There are four different types of video language presentation: original language, dubbed, voice-over, and subtitled.

3.1.2 Motivation for audio description and spoken subtitles

Regardless of style, implementation, or quality, audio description is not only an access strategy; it is ultimately a creative process. Not all visual content can be described and decisions will be made regarding what is important to describe, the vocabulary used, voicing techniques, timing, and method of delivery. These decisions can help ensure that users of audio description have optimum access to audiovisual content thus minimizing the extent to which they are excluded from the experience.

Audio description can benefit a wider audience of diverse users beyond its primary target group, persons who are blind or who have low vision.

Reading is a complex cognitive process. Thus having subtitles read aloud can reduce access barriers and benefit various groups of people, including persons with learning and reading disabilities, persons with cognitive disabilities older people, children, and non-native language speakers, as well as persons who are blind or have low vision.

3.2 Types of audio description

3.2.1 Live audio description

Live audio description is suitable for events taking place in real time.

NOTE Live audiovisual content and associated audio description can also be recorded and later presented in recorded video form.

3.2.2 Pre-mixed audio description

Pre-mixed audio description allows careful planning and evaluation of the audio description before the delivery of audiovisual content.

- a) Where audiovisual is pre-recorded, the accompanying audio description should be pre-mixed.
- b) Pre-mixed audio description should be created as part of the production process rather than external to it.

NOTE This can involve the following:

- Beginning the description process during the script writing phase even though details might need to be added during production and post-production.
- Establishing timing parameters during script writing and rehearsal processes but then recording without strict attention to timing, as small adjustments can be made to timing during post-production.

- c) Where pre-mixed audio description was not created as part of the production process, it should be overseen or at least approved by a member of the original production team.

3.2.3 Harmonious commentary

Harmonious commentary enhances auditory descriptions with voicing techniques and aligns emotional and expressive reactions to script in conjunction with auditory descriptions.

- a) Harmonious descriptions should be used when character expression and actions are conveyed.
- b) Unlike announcers, audio describers should provide harmonious commentary that is in-line with the spirit and intention of the original audiovisual content.
- c) In blending traditional video narration with harmonious commentary, a narrator should work with and/or interview sources close to the production with the goal of understanding what it is about the audiovisual content that is most important to describe.

NOTE Harmonious commentary is frequently used for live events such as sports events, a New Year's Eve concert and countdown, as well as elections, parades, festivals, Olympic ceremonies, and theatre performances.

3.2.4 Extended audio description

- a) Extended audio description techniques may be used to allow the presentation of more information than could be presented using basic audio description.

EXAMPLE 1 A video is identified as being audio described in advance of its commencement.

EXAMPLE 2 The characters who will participate in the narration of a video that is audio described are introduced.

NOTE 1 Levels of importance, including recommendations for the types of content that are suitable to be placed in extended audio, are discussed in [4.4](#).

- b) Extended audio description may involve one or more techniques for presenting additional information to users including the following:
 - 1) stop motion: stopping a video while continuing description to provide additional time for audio description;
 - 2) video slowdown: slowing down the presentation of a video to provide additional time to insert audio description;
 - 3) audio introduction: added information is included before the event starts, offering complementary information.

NOTE 2 Extended audio description is only possible with content that is recorded because its use extends the playing time of a video presentation not just before or after the presentation (as during live presentations), but during the presentation itself (by pausing or lengthening the video to include more audio). This could include videos that are broadcast and recorded by the user for later playback.

3.3 Creating audio description

3.3.1 Narrator preparation

The narrator should have the following:

- a) good native language skills;
- b) the ability to articulate;
- c) a prior knowledge of radio drama, story-telling, and timing techniques (including emphasis, pauses, and sentence phrasing or shaping);

- d) a prior knowledge of when and where musical interludes occur and the order of speaking within live performances;
- e) a sense of the rhythm and basic knowledge of music (e.g. rock, jazz, opera) that occurs within the original audiovisual content;
- f) a sense of timing as it relates to scene pace and scene transitions;
- g) the ability to describe fast paced action without over lapping dialogue, music, lyrics, and intentional pauses in audiovisual content.

3.3.2 Content genre and programme category

The audio description team (of narrators, writers, producers) should make use of (and develop where necessary) genre-based cohesiveness across the programmes or series that are within the same programme category.

3.3.3 Scripting

Scripts can be created for both pre-mixed and live productions. It is expected that changes/additions to the scripts might need to be made for live productions. Script for audio description should consider spoken subtitles.

3.3.4 Writer considerations

The writer should have the following:

- a) a familiarity with the communication needs of people who are blind or low vision;
- b) a prior knowledge of or research into character names, locations, and other specific terminology related to the original content (e.g. uniforms, monuments, architectural elements, and fashion styles);
- c) story writing skills (e.g. skills in building a climax, developing characters, setting up the plot, etc.);
- d) the ability to write concisely and anticipate quality of spoken words per specified time constraints;
- e) the ability to create rhythm within phrases through word choices, order, and syllable counts.

3.3.5 Recording

It should be recognized that any audio description might be recorded (by users if not also by producers), regardless of whether the video programme was live or pre-mixed.

3.3.6 Volume

- a) The volume of audio description and spoken subtitles should have a good balance with the volume of the original audiovisual content audio.

NOTE 1 In some instances, audio description is delivered as a separate stream from the original content. In this case, it is important for both volume levels to be user controlled, where possible.

- b) Audio description and spoken subtitles should be prepared so that users can adjust the volume of audiovisual content and the volume of audio description separately.

NOTE 2 This allows for flexibility in the way the content is used.

NOTE 3 Regulatory rules on volume can exist and can vary from country to country.

3.3.7 Audio technician considerations

An audio technician should assist with the noise cancellation and audio signal processing levels to ensure that cut off, microphone noise, background noise, levels between the audio description and the live sound track, and other audio factors are mixed and filtered appropriately.

3.3.8 Inclusion of consumers in the process of creating audio description

The full range of consumers of audio description should be included in the process of creating audio description where possible.

NOTE 1 This can take the form of focus groups, employment of a consumer as a marketer or community consultant, use of a blind or low vision narrator. Consumers might also be included where stakeholder groups are actively involved in creating audio description and marketing specific audio-described performances.

NOTE 2 It is important to consider balance between artistic expression and consumer needs.

3.3.9 Inclusion of consumers in evaluating audio description

If evaluation is part of the process of developing audio description, representative consumers of audio description should be involved.

NOTE 1 It is important to evaluate pre-recorded and live audio description wherever possible.

NOTE 2 It is important for evaluations to focus on the understandability of the content of the audio description and to avoid dealing with stylistic or other artistic issues.

NOTE 3 It is also important to understand the enjoyment for the target audience; however, the entertainment of the audio description needs to be considered in relation to the general entertainment of the genre. It can be difficult to separate these two things.

3.4 Live audio description

3.4.1 Live narration considerations

- a) Live narrators should work in teams of two.
- b) For purposes of continuity, the narrator who is not performing should remain in the booth/location to assist by suggesting what needs describing as it comes up, or in identifying people and places in real time (such as audiences reactions to live and interactive performances through applause, laughing, standing ovations, etc.).
- c) For rehearsed events, e.g. a play, the narrator who is not performing should be stationed in the auditorium to monitor reception of the narration and to intervene if transmission fails.
- d) In the case of gallery/museum/heritage sites, the live tour should be conducted by the describer with a curator, who will lead the group around the venue.

NOTE 1 An audio technician could also play a key role in pointing out description or providing names of those appearing on camera during the live event.

- e) The narrator(s) should be at the broadcast facilities during a live event.

NOTE 2 Broadcast feed can lag time in receiving video from a remote location; audio mixing and sound controls can impact the quality of audio description if the video narrator is not on location.

- f) The audio description narrator should be isolated in a separate space.

NOTE 3 The location of the narrator has an impact on the recording technologies used and noise cancellation requirements.

NOTE 4 Using an isolated space allows the narrator to be able to speak without disrupting the audience at the live event. It also prevents noise from external sources being transmitted by the narrators' equipment to the listener.

EXAMPLE 1 Newscasts are normally straightforward because they are done from the same location each day. The narrator can be set up in a sound-proof booth somewhere in the studio facility with a video monitor and headphones.

EXAMPLE 2 Many sports facilities have media areas where the media (radio and TV) announcers are located.

EXAMPLE 3 The narrator is located away from the audience so that their voice does not disturb them at a live event such as a telethon, a play, or a sports event.

- g) If the description is prepared for a recorded audio guide, then this should be of studio quality.
- h) With allowances for translation and the possible need to describe unseen elements or unforeseen events, narrators should view and describe only what appears on the stage, or video, or visual art, or object/artwork/building.

NOTE 5 If the narrator is viewing the entire stage/arena and not looking at a video monitor, then the description might not match what is being broadcast due to camera angles, shot changes, zoom, etc.

3.4.2 Preparing for live narration

- a) The role of a script should be determined in advance of the narration and provided as guidance to the narrator.

NOTE 1 This will let the narrator know how much the script needs to be followed and how much the narrator can add to or modify the script.

- b) Where possible, a script should include cues for timing.
- c) A script should be prepared and reviewed by the narrators in advance of the narration. For drama, opera, etc., the provision of a script in advance of the narration is essential.

NOTE 2 A script can vary in its detail and contents from describing notes to fully prepared lines for delivery. However, it is important to recognize that spontaneous and unplanned events can occur during live broadcasts and events and to leave room for these to be allowed for within any script.

NOTE 3 A script can be produced with reference to a pre-recorded live performance of, for example, a play or live show. For drama, opera, etc., the provision of a recording of the performance is essential to aid production of the narrators' script, particularly with regard to timing the narration.

NOTE 4 In the visual arts/heritage sector, the describer needs to prepare by meeting the relevant curatorial/education staff and having access to catalogues and other research material.

- d) The audio technician should visit the location ahead of time to identify problems.
- e) Sound tests, with equipment to be used, on location should be conducted well in advance of a performance to allow for sufficient time to make modifications to equipment or equipment placement.

3.4.3 Narrator transitions in live audio

- a) Depending on the length of the programme, narrators should alternate every half hour to maintain their energy levels. This requirement may be relaxed for stage performances when long passages of dialogue provide narrators with natural breaks and the interval is logistically the most appropriate place to switch narrators. For live tours, it is possible for one describer to provide the description for the whole tour.
- b) Narrator transition should feel natural.
- c) There should not be any interruption in timing of audio description as narrator transition occurs.

3.5 Pre-recorded descriptions

3.5.1 Creation of pre-mixed audio after main production process

Where pre-mixed audio description was not created as part of the production process, it should be overseen or at least approved by a member of the production team, if available.

3.5.2 Audio description team for pre-mixed audio

The audio description team should involve a blind or low vision individual.

3.5.3 Evaluation of pre-mixed audio

An evaluation of the audio description should be done mid-production rather than leaving it until the end.

NOTE 1 The necessity for approval depends on the content and on the expertise of the audio describer.

NOTE 2 The writer compiling an audio description script can view the programme without the images to determine the challenges of the programme as part of developing the script.

3.5.4 Voices for pre-mixed audio

The audio description should be well timed to harmonize with the sounds and voices of the original audiovisual content.

3.5.5 Sound engineering for pre-mixed audio

Sound engineers should supervise recording and mixing to avoid unbalanced sound reception.

4 Guidance on audio description

4.1 General considerations

4.1.1 User considerations

- a) At the beginning of each programme, the user should hear an auditory notification such as "This programme is audio described."
- b) The auditory notice should assume that the listener is using audio description.

EXAMPLE "This programme is audio described" is preferable to "This programme is available...".

- c) The availability of audio description should be identified by a standardized logo, both on the screen (in the case of recorded video content) and in any media where it is advertised.

NOTE It is recognized that logos might vary from country to country. It is important that logos are as consistent across as many jurisdictions as possible.

- d) Users (including blind or low-vision persons) should be easily able to access information (both in real time and in advance) that identifies when audio description is available and for which programmes or events.
- e) The user should be provided with an easy-to-use means of selecting and changing between presenting such as the following:
 - 1) basic audio description;
 - 2) extended audio description;

- 3) no audio description.
- f) When multimedia content which is translated with subtitles has an intense dialogue, it might not be possible to translate and describe everything. If it is necessary to shorten the audio description, providers of audio description should make sure that there is an introduction where characters, scenery, etc. are well presented. If the spoken subtitles are summarized, the user also needs pre-information.

4.1.2 Consistency across technologies

Audio description should be consistent regardless of the technology or method being used to transmit and/or present the original content.

NOTE 1 Some technologies used for transmission include cable, satellite, internet, and DVD.

NOTE 2 Some technologies used for presentation include television, computer, and smart phone.

4.1.3 Consistency within a programme and series

- a) The style of audio description should be consistent throughout a programme.
- b) The style of audio description should be consistent throughout the programmes in a series.

4.1.4 Placement of descriptions

- a) Audio description should only occur during non-dialogue pauses.
- b) Audio description should never occur over dialogue.
- c) Audio description should not obscure music that is necessary to understand the story line.

NOTE Music (diegetic or non-diegetic), sound effects and/or ambient noise might be used to increase emotional resonance within a scene or to help enhance the story line. It is important to be sensitive to the inclusion of audio description in relation to music, sound effects and ambient noise and to understand the purpose of the music, sound effect or ambient noise before deciding whether or not to overlay descriptions on a song, musical piece, sound effect or ambient noise.

4.1.5 Use of pauses

- a) Audio description should not fill every moment.
- b) Where possible, descriptions should allow time for the listener to pause and enjoy/experience quiet moments, music (that is used to increase emotional resonance) and/or background noise.

4.2 Developing audio description

4.2.1 Clarity in audio description

Audio describers should present their information in a manner that can be easily understood by their intended users.

NOTE 1 Proper pronunciation/articulation is the primary objective of the narration and supersedes all other considerations for narration.

NOTE 2 The preparation of a script is useful for lexical choice and selecting which details to describe.

4.2.2 Creativity in audio description

- a) As a creative process, audio description should consider the history, logic and tangible rules that govern genre and sub-genres.

NOTE This is encouraged because the diverse nature of film, television, recorded, theatrical and online content require detailed consideration of the genre involved to appreciate its nuances.

EXAMPLE 1 A narrative genre has audio description that elaborates more on the characters whereas an action genre has audio description that focuses more on plot descriptions.

EXAMPLE 2 A video which has repeating parts has repeating audio description to provide cohesion within the content theme.

- b) The words selected and used for description should be appropriate to the era and genre of the original content.

EXAMPLE 3 A film portraying the Western States during the early 1900's might refer to black people as "coloured people" in dialogue.

EXAMPLE 4 The audio description of a music video makes use of a rhythmic diction.

4.2.3 Narrating/reading audio description script

- a) A narration style for description should be identified and followed consistently within the same content. There are different styles that could be adopted and which have different implications for how audio description is developed and delivered.

EXAMPLE 1 For a third person style of description, the voice of the narrator is clearly distinguishable from the other voices in the video. Where the cast is predominantly male, the director chooses to have a female narrator to clearly distinguish the narrator from the production itself or to have a male narrator to integrate better with the style of the audiovisual content.

NOTE 1 Speech synthesis can replace the human narrator in audio description; however, research indicates a preference for the human voice over voice synthesis.

NOTE 2 Research indicates that, if speech synthesis is used, there is a preference for quality speech synthesis since the synthesized speech is clearly identified as audio-description and avoids confusing the human voice of an audio describer with the characters' voices.

NOTE 3 The quality of speech synthesis is determined by its similarity to the human voice and by its ability to be understood. People familiar with using speech synthesis can be used to judge the quality of speech synthesis.

EXAMPLE 2 The reading of spoken subtitles or other unique visual elements (such as point displays or thought bubbles) is voiced using speech synthesis or by another narrator to indicate a change in content delivery.

EXAMPLE 3 Two voices, male and female, are used. One is used to read spoken subtitles and one to read audio description script.

- b) The voice of the narrator should complement the content being described.

NOTE 4 Often, trained actors are employed as narrators and use their talent to infuse the description with appropriate emotive characteristics.

EXAMPLE 4 A young voice is used to narrate a video intended for children to make the narration seem that it is coming from another child. In some countries, a female voice is used for audio describing children's programmes.

- c) The narrator should be identified at the beginning of the programme to eliminate confusion unless, to advance his or her artistic vision, the director chooses to have the narrators remain anonymous.

NOTE 5 If the main character is anonymous, to coincide with the story, the director might choose to have the narrator remain anonymous too.

- d) The narrator should be consistent throughout a series of programmes unless, to advance his or her artistic vision, the director chooses to have multiple narrators.

EXAMPLE 5 If the main character dies half way through a film and a new character continues the story from a different perspective and gender, the director could decide to switch to a new narrator to reinforce the break in perspectives.

- e) The scriptwriter and narrator of the audio description should not alter, filter or exclude content or images that disturb or offend them.

NOTE 6 The narrator's role is to describe what is seen, not to censor it.

EXAMPLE 6 If the original script uses coarse language, the description script adopts similar coarse language.

- f) The scriptwriter and narrator should make use of a background in or previous knowledge of the culture from which the content originated.
- g) Where it is not possible to make use of knowledge of the culture in advance, pre-recorded audio description should be evaluated by persons with a background in or previous knowledge of the culture from which the content originated, before being released.

NOTE 7 This also includes knowledge of how to voice spoken subtitles on films which include them due to language changes.

EXAMPLE 7 In a Chinese horror film, a character selects number 4 in an elevator. This is significant and indicates potential threat. Knowing the cultural significance of the number four in Chinese culture, a Chinese narrator might describe the number, whereas an Italian describer might not.

4.3 Styles of narration

4.3.1 Newsreader style

Newsreader style focuses on relaying information in a serious manner.

NOTE This style is appropriate when the goal is to educate viewers, as is frequently demonstrated in documentary films or news programme categories.

4.3.2 Commentator style

Commentator style focuses on providing entertainment.

NOTE This style is appropriate when the goal is to include commentary/affective content along with factual content, as is frequently utilized for content that is primarily for entertainment purposes.

EXAMPLE 1 When describing visual actions and events for a situation comedy, the narrator takes on a light-hearted tone, strategically choosing to communicate comedic visual events in a tone that is similar to the show.

EXAMPLE 2 For a courtroom drama, the narrator aims to capture and maintain the intensity of each scene.

4.3.3 First person

The narrator can take on the role of a character describing his/her perspective/reflections on the events.

When first person narration is being used, it should be made clear to the audience that the narrator is taking on a first person role.

NOTE 1 Past tense is the common verb form used for this style.

EXAMPLE 1 "He moved his queen diagonally to take my rook. Then I clobbered his queen with my bishop."

EXAMPLE 2 "There was an 18 wheeler to the left of me and a police car chasing me. I put the pedal to the metal..."

EXAMPLE 3 While “Ann wears only one glove.” is in third person, “I notice that Ann is wearing a single glove.” is in first person.

NOTE 2 The nature of first person is that it is one perspective that might or might not be trustworthy.

4.3.4 Third person

The narrator speaks from an external point of view.

In third person narration, the narrator should not include or reference themselves in the narration.

EXAMPLE 1 “The sun rises over the city skyline.”

EXAMPLE 2 “Don picks up a book and begins to read.”

EXAMPLE 3 While “We now see Bruce entering the living room.” is in first person; “Bruce enters the living room.” is in third person.

4.4 Levels of importance

4.4.1 Using importance levels

Too much information can interfere with a user’s ability to understand the programme. Additionally, depending on the programme genre, there may be very little opportunity to place information.

Levels of importance are determined by the persons responsible for developing the audio description (e.g. content provider, script writer, and narrator) and will vary between different contexts.

NOTE It is useful to include consumers of audio description in making this decision where possible.

4.4.2 Essential information

Essential information is information necessary for the listener to understand the visual and auditory content within the video or performance in which it occurs.

- a) Elements essential to the narrative should be described before additional details are provided.
- b) Audio description should describe primarily the most essential visual details for the comprehension of context, plot and characters.

NOTE 1 Essential information can have some or all of the following properties:

- It is aimed at the target audience.
- It must be known in order to comprehend the content.
- Most people want/need it most of the time.
- The listener would be confused as to what is happening without this information.
- For the content provider, this is the information that they absolutely want the listener to know.
- It identifies visual information that intentionally conflicts with the dialog or background sound.

As more of these properties apply, it is more likely that the information is essential.

NOTE 2 If time allows, the audio description can include additional descriptions about the setting, a character’s physical appearance and/or clothing to enhance the experience.

4.4.3 Significant information

Significant information is secondary to essential information. It provides listeners with a more complete understanding of the visual content and sounds within the audiovisual content in which it appears.

Significant information should be presented, if time permits, in the audio description.

NOTE Significant information can have some or all of the following properties:

- It is aimed at the target audience.
- It gives a more detailed and thorough understanding of the audiovisual information.
- It is information that can be obtained by more than, for example, a quick glance at a video.
- It is important for the user to know about it as they are listening to the main audio content in order to understand the video or event.
- Without this information, the user has an idea of what the video or event is about, but does not have a detailed understanding of it.
- For the content provider, this is information that further explains and gives more details on what they want to tell the listeners.

As more of these properties apply, it is more likely that the information is significant.

4.4.4 Helpful information

Helpful information provides a more thorough understanding of the visual content within the video or event in which it occurs.

Helpful information is information which might be of interest to some users some of the time and if created, should be placed in an extended audio description file.

NOTE Helpful information can have some or all of the following properties:

- It might be of interest to some who are the target audience of the audiovisual content.
- It is the “director’s cut” version of the audio description which further defines cinematic and background details.
- It can be used to reassure the listener that they have not missed something of greater importance.
- Without this information, listeners have a fairly complete understanding of what the audiovisual content is about but have some things that they still want to know.
- It includes different or other possible interpretations of the information being expressed by the visual image.
- For the content provider, this is information that could clarify some things for some people, for example, those with cognitive impairments, English as second language speakers or children.
- It includes optional extra information that is seldom wanted or needed, and which elaborates on what is already there.

As more of these properties apply, it is more likely that the information is helpful.

4.4.5 Irrelevant information

Information is irrelevant if it does not help to provide additional understanding of the visual content for any users to understand the audiovisual content in which it occurs.

Information that is irrelevant should not be presented to users either in basic audio description or in extended audio description.

NOTE Information that is irrelevant can have some or all of the following properties:

- Very few users will want to know this information.
- It is rarely helpful.
- It is not important enough to mention.

- Without this information, the listener knows everything they want or need to know to understand and enjoy the video or event.
- This is information that might result in unintended confusion or boredom and does not help users understand what the content provider wants to communicate.

As more of these properties apply, it is more likely that the information is irrelevant.

4.5 General guidance on descriptions

4.5.1 Description of sounds

- a) Common well-known sounds, such as a ringing telephone, barking dog, or a car horn, should not be identified/described unless the sound is out of context or is coming from an unknown source.

EXAMPLE Malcolm is in the driver seat of his car. Another car moves close behind and the driver honks his horn and shakes his middle finger angrily. The description of the honking horn need not be described as the honking can be heard and it would be an unnecessary description.

- b) An unknown sound should not be identified/described if such description will give away the plot.
- c) In the case of a subtitled programme, a description with the sound effects should be added at the beginning of the video if there are spoken subtitles.

NOTE This is considered to be “describing the screen real estate”.

4.5.2 Logos/credits/titles

- a) Titles and credits at the beginning and end of programmes should be read at a normal pace without rushing.
- b) In general, although it is unlikely that there will be sufficient time to allow for the verbal delivery of all credits, effort should be made to include as many as possible.

4.5.3 On-screen text

The narrator should acknowledge text elements, describe the content, and read the message.

NOTE Signs, speech bubbles, and other non-captioned on-screen text elements might appear in the background, which have some relevance to the setting, plot, or understanding of actions taking place.

EXAMPLE “A road sign reads: ‘Caution Construction Zone’”

4.6 Guidance on identifying objects

4.6.1 Identifying objects

- a) In languages that have definite and indefinite articles when an object is first introduced, the indefinite article (e.g. “a”, “an” in English) should be used.
- b) When the same object is referred to, or has already been known, named or identified, the definite article (e.g. “the” in English) should be used.

NOTE 1 This helps distinguish new from known objects and multiple objects.

EXAMPLE 1 “Sandra removes a chocolate from a heart shaped red box. She bites into the chocolate, closes her eyes and smiles.”

EXAMPLE 2 “I entered the cave bearing a sword. I saw the sword Excalibur embedded in a large stone and ran to it. I dropped my own sword on the ground and slowly withdrew the jewel-encrusted Excalibur from the stone.”

EXAMPLE 3 “Alan takes a broom out of the closet. He takes the broom with him and leaves the kitchen.”

NOTE 2 An object can be considered as known if its presence is understood as a given, or is common to the circumstance or setting.

EXAMPLE 4 “Ida and George are sitting on the couch.”

EXAMPLE 5 “I light the candles on the dining room table.”

EXAMPLE 6 “She waters the flowerbeds with the garden hose.”

4.6.2 Describing colour

- a) Colour should be included in descriptions as appropriate.
- b) It should not be assumed that colour holds no significance to blind persons or persons with low vision or that they have never seen colour or lack understanding of the concept of colour or its cultural symbolism.

NOTE Some poetic descriptions involving the symbolism of colours might not be easily translated into other languages or cultures.

- c) Word choices for colour may also contain a textural element.

EXAMPLE 1 “The ballroom is lit with fire.” The word “fire” in this example is intended to represent the colours of red and yellow but also has textural properties of heat or warmth.

EXAMPLE 2 “My gown trimmed in gold”. The word gold used in this example can be understood as a colour but also has textural properties of being heavy and soft.

4.6.3 Describing visual effects

- a) Camera-specific terminology should be avoided unless it is a commonly used term in the language.

NOTE 1 Visual distortions, such as blurring, focus, and mirror views, can be used to provide a character’s point of view or mood.

NOTE 2 Camera angles or camera point-of-view can be meaningful.

EXAMPLE 1 “I focused in on a small paper lying on the sidewalk.” The camera zooms in on the small paper on the sidewalk.

NOTE 3 Lighting and lighting effects can be used for mood and/or symbolism.

EXAMPLE 2 “The light cast a shadow of someone behind the curtains.”

EXAMPLE 3 “The operating room becomes awash in white light. A blue sky appears.”

- b) Some dances, dance styles, or movements have specific names that should be used. When the dance is not identifiable by a specific name, a more general term can be used.

NOTE 4 Dancing and choreography are a combination of movements, music, and motions in addition to the dress and costume attire of the individuals appearing within the scene.

EXAMPLE 4 A specific type of dance could be “the Tango”, whereas a general term could be “Latin dancing”.

EXAMPLE 5 A specific martial art form could be “Karate”, whereas a general descriptor could be “martial arts”.

4.7 Guidance on identifying persons

4.7.1 Identifying characters or places by name

- a) Minor characters or places should be described by their prominent physical appearance or identity.

- b) Once an appellative is established for minor characters, it should be used consistently.
- c) Minor characters or places, or those whose names remain unknown, should be described in general terms.

EXAMPLE 1 “A tall well-dressed woman walks through the crowded sidewalk to the car and leans through the open window and asks.”

- d) When known characters or places first appear, they should be identified or introduced using their full name.
- e) After their first introduction, the narrator should consistently use the name that is most commonly used by the other characters to refer to well-known characters.

EXAMPLE 2 When a doctor first appears, the narrator says “Michael Hunt entered the room.” Since Michael is well known to the others in the room, the narrator later refers to him as “Michael”.

EXAMPLE 3 “Raj Binder enters the classroom and approaches Consuela Singh. Raj grasps Consuela’s hand and together they turn towards the children sitting at their desks.”

4.7.2 Physical appearance

4.7.2.1 Character’s physical appearance

- a) Third person descriptions should describe qualities that would allow the listener to draw their own conclusions, i.e. long shiny dark hair, large brown eyes, tall athletic build, flawless complexion.
- b) Third person descriptions should avoid using terms that evaluate the character’s appearance (such as beautiful, lovely, ugly, plain) unless relevant to the plot or genre.

NOTE In third person descriptions, the level of attractiveness of a character is a value judgment.

- c) First person descriptions may provide a personal assessment of attractiveness if it is relevant to that particular character.

EXAMPLE When the narrator takes on the character of a beauty pageant contestant, using language that describes the individual’s self-assessment of her beauty (or that of others) is appropriate.

4.7.2.2 Facial/physical expression

Terms should be used which describe emotion or reaction through facial/physical expression when they are important to the understanding of characters and plot.

EXAMPLE “John gives an evil smile as he escapes in his car” adds extra unnecessary words “gives an evil smile” where one word is sufficient as in “John sneers as he speeds away in his car.”

4.7.2.3 Attire

Describing the clothing worn by characters should be included if it helps establish characterization, role, plot development, setting, occasion, location, or the enjoyment of the genre.

NOTE Attire is often used to portray a stereotype. In these instances, a sighted viewer instantly identifies the character type by their clothing. The narrator could either describe the attire or the character type as appropriate for the situation.

EXAMPLE 1 A comedy series has a recurring character that is a nerd/geek. Describing this character’s unfashionable, outdated, and ill-fitting or inappropriate clothing could help support the characterization.

EXAMPLE 2 A scene in a drama series involves investigators entering a bar where many of the patrons are tattooed and wearing leather and chains. The narrator could elect to simply call them “bikers” or to describe their attire or both.

EXAMPLE 3 A person wearing a white coat over their regular clothing and carrying a stethoscope is a character stereotype which instantly identifies them as a doctor.

4.7.3 Race/ethnic origin

4.7.3.1 Identifying race/ethnic origin

The appearance of characters should not require that they be identified by race or ethnic origin, unless these types of identifications would provide meaningful information or insight to a sighted viewer which would not otherwise be available to persons who are blind or who have low vision in a similar timeframe, i.e. plot development, character motivation, or background.

NOTE 1 Consider the role of a casting director. Is a character's race or ethnicity relevant to the selection and casting of roles and if so why?

NOTE 2 If sighted viewers could use the race or ethnic origin of the characters as clues to solving the crime or providing a better understanding of the culture or behaviour of the characters, it would be appropriate to include race/ethnic origin in the description.

EXAMPLE 1 In a police drama series, investigators are interviewing witnesses/victims of a robbery. The actors playing the roles are Asian, Black, and Aboriginal. The narrator does not identify the subjects' race/ethnic origin unless it is relevant to the plot and cannot be recognized from the existing dialogue.

EXAMPLE 2 A medical series with a large multiracial ensemble cast has doctors treating patients of various races and ethnic origins. In these instances, the race or ethnic background of the characters might have no relevance. The rare exception might be if such information specifically relates to a plot development, i.e. the patient has a disease that is specific to a racial/ethnic group.

EXAMPLE 3 A reality show has contestants of various racial and ethnic backgrounds. If the narrator thinks or has been informed by the director that this type of information is useful or meaningful to the viewer, then it is important that all contestants be identified using these characteristics.

4.7.3.2 Terms to describe race/ethnicity

- a) Derogatory or disrespectful language should never be used unless it is part of the development of the character in a first person style.

EXAMPLE If the narrator took on the role of an Indian princess from the mid-nineteenth century, the use of derogatory language when talking about or referring to people of a "lower" class might be appropriate if that type of language is used in the original content.

NOTE It is important to remain sensitive and aware that some ethnic/racial descriptive terms accepted in Western culture might differ from other cultures and do change over time.

- b) Descriptions should use common and factual terms to describe individuals that would be understood by the majority of listeners such as the following:

- 1) light brown, brown, dark brown, tan, light tan;
- 2) light/fair skinned, dark skinned, olive-skinned;
- 3) tan, ruddy, olive, white, yellow.

- c) If the attire of a character suggests their cultural or ethnic origin, then this information should be described, i.e. kilt, burka, saris, kimono, parka, etc.