

I n t e r n a t i o n a l T e l e c o m m u n i c a t i o n U n i o n

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

T.701.21

(03/2022)

SERIES T: TERMINALS FOR TELEMATIC SERVICES
User interfaces - Accessibility and human factors

Guidance on audio description

Recommendation ITU-T T.701.21

ITU-T

ITU-T T-SERIES RECOMMENDATIONS
TERMINALS FOR TELEMATIC SERVICES

Facsimile – Framework	T.0–T.19
Still-image compression – Test charts	T.20–T.29
Facsimile – Group 3 protocols	T.30–T.39
Colour representation	T.40–T.49
Character coding	T.50–T.59
Facsimile – Group 4 protocols	T.60–T.69
Telematic services – Framework	T.70–T.79
Still-image compression – JPEG-1, Bi-level and JBIG	T.80–T.89
Telematic services – ISDN Terminals and protocols	T.90–T.99
Videotext – Framework	T.100–T.109
Data protocols for multimedia conferencing	T.120–T.149
Telewriting	T.150–T.159
Multimedia and hypermedia framework	T.170–T.189
Cooperative document handling	T.190–T.199
Telematic services – Interworking	T.300–T.399
Open document architecture	T.400–T.429
Document transfer and manipulation	T.430–T.449
Document application profile	T.500–T.509
Communication application profile	T.510–T.559
Telematic services – Equipment characteristics	T.560–T.619
General multimedia application frameworks	T.620–T.649
User interfaces - Accessibility and human factors	T.700–T.799
Still-image compression – JPEG 2000	T.800–T.829
Still-image compression JPEG XR	T.830–T.849
Still-image compression – JPEG-1 extensions	T.850–T.899

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T T.701.21

Guidance on audio description

Summary

Audio description is a service or feature of the auditory modality that contains an oral narration with details or context involving situations and scenes found in an audiovisual content (such as recorded video presentations, broadcast television, cinema, live or recorded drama). The primary users of audio description are blind persons or persons with low vision and their friends and family.

Recommendation ITU-T T.701.21 provides guidance to audio description developers and practitioners 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.

Recommendation ITU-T T.701.21 is twin with the published ISO/IEC TS 20071-21:2015 "Information Technology – User interface component accessibility – Part 21: Guidance on audio description" developed by ISO/IEC JTC1 SC35.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T T.701.21	2022-03-29	16	11.1002/1000/14972

Keywords

Accessibility; audio description.

* To access the Recommendation, type the URL <http://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID. For example, <http://handle.itu.int/11.1002/1000/11830-en>.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2022

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

		Page
1	Scope.....	1
2	References.....	1
3	Definitions	2
	3.1 Terms defined elsewhere	2
	3.2 Terms defined in this Recommendation.....	2
4	Abbreviations and acronyms	4
5	Conventions	4
6	Framework and process considerations for audio descriptions	4
	6.1 General	4
	6.2 Types of audio description	5
	6.3 Creating audio description.....	6
	6.4 Live audio description	7
	6.5 Pre-recorded descriptions	9
7	Guidance on audio description	9
	7.1 General considerations	9
	7.2 Developing an audio description	10
	7.3 Styles of narration.....	12
	7.4 Levels of importance	13
	7.5 General guidance on descriptions.....	15
	7.6 Guidance on identifying objects.....	15
	7.7 Guidance on identifying persons	16
	7.8 Guidance on relationships	20
	7.9 Guidance on parts of speech.....	21
	Bibliography.....	22

Introduction

Recommendation ITU-T.701.21 provides audio description (AD) 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 AD applies include recorded videos, broadcasts and broadband television, cinema, live or recorded dramas, museums and art gallery exhibits, heritage tours, news, and comedies. Some of this content is static and unchanging, some are dynamic, some are expressed in a visual modality alone, and some are in a combination of visual and auditory modalities.

An example of mainstream use of audio description techniques can be found in sports commentaries. It can be enjoyed by diverse audiences and uses personalities, emotional expressions and reactions 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 descriptions can be provided in different languages. Providing audio descriptions and translating them into other languages can be beneficial to diverse users within 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 texts and providing them in auditory forms can also be beneficial.

Producing audio descriptions can be thought of in a similar way to language translation where 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 descriptions based on the guidance in this Recommendation is not intended to interfere with or change the meaning of the original content, freedom of speech and expression, or opinion of rights holders.

Further benefits of providing audio descriptions can be attributed towards raising awareness of barriers experienced by blind persons and persons with low vision and other beneficiaries in accessing audiovisual content. The use of this Recommendation can also help embed universal and inclusive designed media content production practices in educational programmes (such as films, theatre performances, museums and art gallery exhibits and other forms of media).

Recommendation ITU-T T.701.21

Guidance on audio description

1 Scope

This Recommendation provides recommendations for describing audiovisual content in an auditory modality for use in recorded video presentations, broadcast television, cinema, live or recorded dramas, museums 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 descriptions are blind persons or persons with low vision and their friends and family.

This Recommendation 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 consistently 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 an audio description depending on the content being described.

This Recommendation 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, digital versatile disk (DVD) and home cinema equipment and other forms of user interface technology. Therefore, this Recommendation does not consider transcoding files for the various video outputs.

NOTE 3 – Technical matters of transmission and distribution are covered by other standards (e.g., motion picture experts group (MPEG) standards and other technical international standards such as [b-IEC 62731]).

This Recommendation 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 [b-ITU-T T.701.11] for guidance on text alternatives for images).

It applies mainly 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.

ITU-T T.701.21 is a twin text publication with ISO/IEC TS 20071-21:2015, "Information technology – User interface component accessibility – Part 21: Guidance on audio descriptions".

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

None.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 General terms

3.2.1.1 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 – This generic term is used where only the visual modality might apply, for example, in an art gallery or museum exhibits.

NOTE 2 – 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.

3.2.1.2 Creative: <audio description> Creation and addition of new original content to existing content.

3.2.1.3 genre: <audio description> Classification of a film, programme, drama, or an event from an artistic perspective.

NOTE 1 – Genres include, but are not limited to, history, action, horror, romance, and comedy.

NOTE 2 – Programme genres are not necessarily mutually exclusive.

3.2.1.4 harmony: <audio description> Natural fit between the audio description script and its delivery and the original content.

3.2.1.5 programme category: <audio description> Classification of programmes

NOTE 1 – Some programme categories include documentary, news and information, and drama.

NOTE 2 – Programme categories are not necessarily mutually exclusive.

3.2.1.6 programme: <audio description> Complete unit of a recorded or live video broadcast.

3.2.1.7 video (recorded): <audio description> Combination of auditory and visual content intended to be presented together in a synchronized manner.

3.2.2 Audio-specific terms

3.2.2.1 audio track: Auditory components of a video other than audio description components.

3.2.2.2 audio description: Descriptive audio.

NOTE 1 – Audiovisual content is described in an auditory modality.

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

NOTE 3 – Audio description can also be used to describe locations, directions, and objects.

3.2.2.3 audio introduction: Complementary auditory content included before the original audiovisual content begins.

3.2.2.4 basic audio description: Audio description provided along with the original audiovisual content without lengthening the time of presentation.

3.2.2.5 extended audio description: Audio description provided with the original audiovisual content that lengthens the total time of the presentation or event.

3.2.3 Language of presentation terms

3.2.3.1 original language: Native language in which audiovisual content is produced.

3.2.3.2 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 – When dubbing, this is timed to match the voicing or mouth movements of animated characters or the original actors.

3.2.3.3 subtitles: Transcription or translation of the dialogue, suitable for when the sound is available but not understood.

NOTE – Specific guidance on delivering subtitles, on-screen text, and embedded text will be provided in a forthcoming revision to this Recommendation.

3.2.3.4 spoken subtitles/audio subtitles/spoken captions: Subtitles on audiovisual content that are read aloud and spoken over the audio in subtitled audiovisual content.

3.2.3.5 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 – Subtitles can also be identified through alternative voicing, synthetic speech, or sound indicators.

NOTE 2 – 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 – Spoken subtitles and audio descriptions can be available to users separately or together.

NOTE 4 – Spoken subtitles are synchronized with the text on screen.

3.2.4 Audio description production terms

3.2.4.1 audio describing: Delivering a script designed to meet the needs of persons wishing to have access to audiovisual content.

3.2.4.2 co-produced: <audio description> Produced at the same time as the rest of the audiovisual content.

3.2.4.3 live production: Production process that involves creating audio descriptions in real time along with the creation and delivery of the original audiovisual content.

NOTE – Live productions are typically experienced initially in real time but can also be recorded for later redistribution.

3.2.4.4 machine narrator: <audio description> Technology which delivers text-based audio description.

NOTE – To date, machine-based narrations have not been as acceptable to persons who are blind or persons with low vision as voiced audio description.

3.2.4.5 post-produced: <audio description> Produced after the audiovisual content.

3.2.4.6 postscript/pre-production audio description: Audio description considered after the script has been developed, but before the event takes place.

NOTE 1 – An example of this is the text of a play.

NOTE 2 – 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").

3.2.4.7 pre-mixed production: Content that includes an audio description that is created and recorded prior to transmission.

3.2.4.8 voice narrator/narrator/describer/audio describer: <audio description> Person(s) and/or technology by which audio description is voiced.