Captioning quality guidelines
Broadcasting Services (Television Captioning) Standard 2023

MARCH 2024

Canberra
Level 3
40 Cameron Avenue
Belconnen ACT

PO Box 78
Belconnen ACT 2616

T +61 2 6219 5555
F +61 2 6219 5353

Melbourne
Level 32
Melbourne Central Tower
360 Elizabeth Street
Melbourne VIC

PO Box 13112
Law Courts
Melbourne VIC 8010

T +61 3 9963 6800
F +61 3 9963 6899

Sydney
Level 5
The Bay Centre
65 Pirrama Road
Pyrmont NSW

PO Box Q500
Queen Victoria Building
NSW 1230

T +61 2 9334 7700
F +61 2 9334 7799

Contents

Captioning quality guidelines

Introduction

Captioning is the text-based service on free-to-air and subscription television programs that translates the soundtrack of an English language television program into on-screen English text. Captioning can include representations of all audio on a television program, including spoken dialogue, descriptions of music, sound effects and other sounds such as laughter or clapping. The captions that are provided for a television program, or for a distinct program segment within a television program, are referred to as a 'captioning service'.

Captioning is specifically designed to assist viewers who are deaf or hard-of-hearing to better access television content. Captioning differs from 'subtitling', which translates spoken dialogue from another language, or one-off incidences of inaudible English dialogue, into English text on screen.

In Australia, all commercial, national and subscription television services are required to provide minimum levels of captioning under Part 9D of the *Broadcasting Services Act 1992* (the BSA).[1] Collectively, these are described as 'broadcasters' in these guidelines. The providers of those television services are also required to ensure that captioning meets the minimum levels of quality set out in the *Broadcasting Services (Television Captioning) Standard 2023* (the Standard), that has been made by the ACMA under Division 4 of Part 9D of the BSA.

These guidelines are intended to provide guidance to broadcasters, captioning providers and audiences about the ACMA's interpretation of key elements in the Standard to:

> increase general understanding of how the requirements relating to quality in the Standard are interpreted by the ACMA

> assist broadcasters to achieve compliance with the requirements relating to quality in the Standard.

These guidelines provide general guidance about the Standard and do not vary or supplement the terms of the Standard. Examples used in these guidelines are illustrative only. The terms of the Standard should be relied upon when determining the requirements relating to the quality of captioning.

Note that these guidelines do not provide guidance about enforcement action that the ACMA will take if a broadcaster is found to be non-compliant with the Standard. The ACMA's approach to compliance and enforcement is set out in our compliance and enforcement policy.

The guidelines have been developed in response to feedback received from broadcasters in the course of ACMA investigations, during ACMA audits of captioning quality, and in response to consultation about the making of the Standard. The matters covered in these guidelines have been identified by the ACMA from those sources as matters where clarification might be useful.

It is intended that the guidelines will be reviewed and updated from time to time, subject to consultation with stakeholders. Amendment of the guidelines may be prompted if additional matters requiring clarification are raised by stakeholders.

---

[1] This includes all commercial television broadcasters, the national broadcasters (the Australian Broadcasting Corporation and the Special Broadcasting Service) and subscription television broadcasters and narrowcasters, but does not include subscription video on demand services provided over broadband internet.

What does the Standard do?

The ACMA is required to have a Standard in place by which the quality of captions can be assessed.[2]

The object of the Standard is to ensure that captions provided for television programs are meaningful to deaf and hard-of-hearing viewers. The Standard does this by establishing mandatory requirements for the quality of captioning provided on television programs. The Standard requires captions to be:

> readable (paragraph 10(a))

> accurate (paragraph 11(a))

> comprehensible (paragraph 12(a)).

The Standard also sets out factors that must be considered when determining whether captions are readable, accurate and comprehensible. These are at paragraphs 10(b) for readability, 11(b) for accuracy and 12(b) for comprehensibility.

The factors listed for readability, accuracy and comprehensibility are not exhaustive. They comprise factors that the ACMA must consider in the context of a program, or a distinct program segment, as a whole when assessing captions. However, it is up to the ACMA to decide, in the context of each assessment, what the relevant factors are in each case.

Paragraph 9(a) of the Standard stipulates that the quality of a captioning service for 'a program' must be considered in the context of the program as a whole. However, if a program is made up of distinct program segments that are unrelated to each other (such as in most news, current affairs, and magazine-style programs), the quality (and therefore, meaningfulness) of the captioning for each individual segment will be considered separately, in accordance with paragraph 9(b) of the Standard.

In general, when assessing the quality of captioning under the Standard, the ACMA uses an outcomes-based approach and will consider a range of factors associated with readability, accuracy and comprehensibility of captions to assess whether the captioning service, overall, enabled viewers to meaningfully access the program or distinct program segment.

The circumstances of a broadcast and the nature of a program are also relevant considerations. We recognise that broadcasters may use different methods of captioning, such as live captioning and pre-prepared captioning. We take the view that it is important to consider whether the captioning service provided with a program is meaningful in the context of the program, or distinct program segment, as a whole.[3]

This approach allows us to be flexible in determining what factors are most relevant in all of the circumstances of any particular television program or distinct program segment. For example, this allows for the captions for a live news segment to include minor spelling mistakes if those mistakes have no material impact on conveying the meaning of the segment to a deaf or hard-of-hearing viewer.

So, while the factors in paragraphs 10(b), 11(b) and 12(b) of the Standard must be considered when determining the readability, accuracy and comprehensibility of captions, those factors do not comprise exhaustive lists for readability, accuracy, or comprehensibility. In assessing the quality of captions, the weight given to each factor, if any, will depend on the circumstances of each case, including the context of the program. This means that in assessing the quality of captioning, there is no requirement that all or even some of those factors must be evident for the captioning service of a program to satisfy the obligations set out in paragraphs 10(a), 11(a) and 12(a) of the Standard.

---

[2] See subsection 130ZZA(8) of the BSA.
[3] See the Explanatory Statement in relation to section 9 of the Standard.

General approach

When assessing the quality of a captioning service against the Standard, the ACMA will have regard to the object of the Standard:

> The object of this standard is to specify mandatory requirements for broadcasters and narrowcasters that relate to the quality of captioning services, to ensure that captioning services are meaningful to deaf and hard-of-hearing viewers.[4]

Consistent with this object, we assess compliance with the Standard from the point of view of a deaf or hard-of-hearing viewer to determine whether the captions would be meaningful to those viewers. In practice, this is achieved by first reviewing a television program with audio enabled and captions disabled, and then a second time with audio disabled and captions enabled (as a deaf or hard-of-hearing viewer would experience the program).

What is 'meaningful'?

The ACMA considers that, to be 'meaningful', a captioning service must convey the same meaning to a viewer who is deaf or hard-of-hearing as would be understood by a viewer who is not deaf or hard-of-hearing, from watching the program and listening to the soundtrack. It is not sufficient for the captioning service to have some meaning, or to be partly meaningful.

How is quality determined?

The Standard also sets out how the quality of captioning services is to be determined and requires that a captioning service for a television program must be considered either:

> in the context of the television program as a whole[5]; or

> if the content broadcast is a distinct program segment within a television program, in the context of that distinct program segment considered as a whole segment.[6]

When determining the quality of a captioning service, the ACMA is required to consider the cumulative effect of the readability, accuracy, and comprehensibility of the captions.[7] Accordingly, the ACMA does not prioritise the importance of any of these 3 requirements.

The Standard does not set different requirements for the quality of captioning between live and pre-recorded content. However, we will consider the differences between providing captioning services for live and pre-recorded television programs; and wholly live or wholly pre-recorded television programs and television programs that include both live and pre-recorded program material. For example, this will include the fact that time constraints apply to captioning of live and partially live television programs and that latency may be more likely to occur between the broadcast of a soundtrack, and the corresponding captions, when a program is broadcast live.

The requirement to determine the quality of a captioning service for a television program (including a distinct program segment) in the context of the program (or distinct program segment) as a whole, while considering the cumulative effect of the readability, accuracy, and comprehensibility of the captions, means that the ACMA will not necessarily conclude that every captioning error will result in non-compliance with the Standard.

---

[4] Section 5 of the Standard.
[5] Paragraph 9(a) of the Standard.
[6] Paragraph 9(b) of the Standard.
[7] Paragraph 9(c) of the Standard.

Accuracy – 'nights' is misspelled as 'Knights', but the meaning of the statement is clear in the context of the segment.

However, a high incidence of errors can render a program, or distinct program segment, not meaningful to a deaf or hard-of-hearing viewer. This would lead to a finding of non-compliance with the Standard as the Standard requires an assessment of the quality of a captioning service based on the cumulative effect of those errors in the program, or distinct program segment, as a whole.

Example 1: Captioning to be meaningful – no breach

A complaint the ACMA received concerned the quality of a captioning service for a near-live current affairs program that was broadcast a short time after it was recorded in a studio. The program comprised a series of unscripted discussions, that were live-captioned, between the presenter and a group of panellists about issues of the day. The segments were separated by pre-recorded clips that were captioned ahead of time. The dialogue was fast paced, with the presenter and the panellists frequently interjecting and speaking over each other.

In one near-live segment, the ACMA found that the captions frequently fell behind the dialogue, and that on many occasions, brief interjections by a person speaking that were ignored by other panellists were not captioned. Occasionally, captions for the dialogue of a speaker were not verbatim, and fewer words were used to convey the meaning of what was being said. Some captions were misspelt – usually because a homonym was used for the word or words that were spoken. In order to mitigate the latency of the captioning, the presenter's closing remarks for the segment, and opening remarks for the next segment, were not captioned.

The ACMA noted that, although the captions did not always synchronise with the dialogue, some comments were not captioned, and there were some spelling errors, this was to be expected due to the near-live nature of the segment. Despite the latency of the captions and the omission of some captioning, the ACMA considered that the cumulative effect of these shortcomings did not materially detract from a meaningful understanding of the dialogue. The ACMA found in this case that the captioning service for the program was meaningful to deaf and hard-of-hearing viewers because they would have understood the dialogue in the segment in the same way as viewers who did not rely on the captions.

Distinct program segments

The Standard provides that a 'program' includes both a television program and 'a distinct program segment within a television program'. For the purposes of the Standard, a captioning service for either a television program, or a distinct program segment within a television program, must comply with the requirements relating to quality. The Standard defines a 'distinct program segment within a television program' to mean a distinct segment that is unrelated to other program segments within that same television program.

Television programs that comprise distinct program segments can include news and current affairs, breakfast and light entertainment shows, sports compilations or magazine style programs. These programs usually consist of discrete segments about a single item or issue that are self-contained in the sense that a viewer's capacity to understand the segment does not rely on whether they watch the preceding or following segments.

For example, news items or current affairs stories fall within the meaning of a distinct program segment in the Standard because one news item or current affairs story is usually unrelated to other news items or current affairs stories within a television program. Other examples of television programs that comprise distinct program segments include different guests appearing on a breakfast show to spruik their new book or appearance in a new movie, or a sports program comprising a panel speaking about recent matches in different sports.

In television programs made up of distinct program segments, the ACMA assesses each distinct program segment, such that, in the context of the segment as a whole, the captioning must be meaningful to a deaf and hard-of-hearing viewer.

The ACMA has previously considered and rejected alternative interpretations of distinct program segments suggested by broadcasters. For example, in a news program, it has been suggested that each section, that is, news, sport, finance and weather, is a distinct program segment. Another suggestion was that the section between each commercial break is a distinct program segment. Those models have not been adopted as the ACMA's standard approach because they do not necessarily accord with the intended meaning of a distinct program segment in a television program, since each section between commercial breaks, or the news, sport, finance or weather sections in a news program, can be made up of several items where the subject matter of each item is unrelated to or has no bearing on the other.

However, in practice, we have usually found that a special interest section within a news program, such as finance, sport or weather, constitutes a distinct program segment. This is because those sections either contain a single item, or multiple items that relate to the same subject. This is less likely for a news section, which is inherently more varied, being made up of numerous stories on a range of different topics.

Example 2: Distinct program segments – no breach

A complaint the ACMA received concerned the quality of a captioning service for a 30-minute news program broadcast live to air. The news program opened with 3 stories about flooding in a regional area. In the first story, a reporter on location was talking about the disruption and damage caused by the flood and referred to reports that a popular local identity was believed to have drowned in floodwaters. The broadcast then returned to the studio, where the presenter introduced the second story. A different reporter was shown standing outside the missing local identity's home in the regional town reporting on the circumstances surrounding his disappearance.

The broadcast then switched back to the studio for the third story, a live cross to a spokesperson from the State Emergency Service, with the presenter interviewing him for an update on the flooding situation. Later in the news program, during the weather section, it was predicted that heavy rains would ease over the coming days and the floods in the affected regional Queensland towns would subside.

The ACMA was of the view that the 3 stories comprised 3 items in the news program, but as they were sequential and each was about the floods in the same region, those 3 stories comprised a single distinct program segment within the meaning of the Standard. However, the ACMA was of the view that, although the weather section of the news bulletin contained reference to the floods in the region, that was an aside made by the weather presenter in providing a report on state, territory and national weather. In addition, the weather report was a separate distinct program segment because it was separated from the earlier 3 stories by intervening news items, sports and finance.

Latency and timing of captions

Latency in a captioning service refers to the delay between when the soundtrack of a program is broadcast (for example, dialogue) and when the corresponding captions appear on screen. Captioning delays usually occur when captioners cannot keep up with the pace of the soundtrack.

The Standard does not specify a maximum tolerable duration of latency. Paragraph 12(b) of the Standard requires that the following factors relating to the latency and timing of captions must be considered when determining whether captions are comprehensible:

> whether the captions are displayed for a sufficient length of time to allow the viewer to read them and follow the action of the program (sub-paragraph (ii))

> the extent to which the appearance of the caption coincides with the onset of speech of the corresponding speaker, sound effect or music (sub-paragraph (iii))

> the extent to which the disappearance of the caption coincides with the end of the speech of the corresponding speaker, sound effect or music (sub-paragraph (iv))

> the extent to which a caption over-runs a shot or scene change (sub-paragraph (viii))

> the extent to which the appearance or disappearance of the caption, as the case may be, coincides with the relevant shot or scene change (sub-paragraph (ix)).

The ACMA understands that captions do not always synchronise with dialogue, especially in live broadcasts, and that some latency is to be expected. The ACMA has found that, while captioning delays will not inevitably make captioned content less accessible to deaf or hard-of-hearing viewers, the longer the delay between the audio content of a program and the corresponding captions, the more likely that latency will lead to issues with comprehensibility. This is because the further captions fall behind the audio content, the more difficult it is for deaf and hard-of-hearing viewers to relate them to the images or speakers appearing on screen. In some cases, a delay between the on-screen visual content and the corresponding captions can affect comprehensibility by detracting from the humour or emotional impact of a scene.

The ACMA has observed that captioning delays are more likely to lead to non-compliance with the Standard in live-captioned programs.

Delays in captioning are also likely to cause secondary issues. For example, we have observed that when captions fall behind, there is a tendency for abbreviations to be used or captions to be summarised, sometimes causing comprehensibility or accuracy issues. In other instances, the captions are displayed for shorter durations, affecting the readability of the captions as deaf or hard-of-hearing viewers have insufficient time to read them.

Generally, we have found that continuous delays of 6 seconds or more are likely to make it difficult for a viewer relying on captions to follow a program, although this may depend on other factors. In some programs, captioning delays can affect program comprehension, while in others, even long delays can have minimal impact, particularly where visual cues make up for long pauses in captions.

For example, we have found that comprehensibility issues caused by latency in captions are more likely to occur when there are several people talking and/or there are frequent scene changes. Latency can also be a critical factor in programs featuring

comedy, as comedy is dependent on the timing of delivery relative to visual cues or the interplay between multiple speakers.

Conversely, latency in captions is less likely to affect comprehensibility in a program that is slow-paced, with minimal dialogue or with few scene or topic changes. This includes programs where a single person is speaking to camera (for example, a public figure delivering a speech or a news anchor giving editorial comment). In those cases, latency is less likely to affect comprehensibility because deaf or hard-of-hearing viewers will be more likely to know who is talking and to understand the substance of what is being said.

Example 3: Latency – breach

A complaint the ACMA received concerned the quality of a captioning service for a late-night, light entertainment topical quiz show. The program consists of a host asking 5 celebrity guests a range of news-related questions to which the guests give humorous or satirical answers. The format of the program involves the display of visual clues, fast-paced dialogue, rapid scene changes with a heavy reliance on timing, including the interplay of images and sounds, to create humour. Although the program is not broadcast live, it relies on being topical, meaning the program is broadcast within 24 hours of being recorded and is live-captioned.

When assessing the episode referred to in the complaint, the ACMA identified captioning latency ranging from 8 to 11 seconds on 14 occasions during the program. On 2 occasions, the delays were followed by periods where the length of time the captions were displayed were so brief that it was difficult for a viewer to read them and keep pace with the dialogue.

The ACMA found that the captioning delays in the fast-paced program, combined with the short on-screen appearance of some captions, affected the humour and context of the jokes. While the ACMA accepted that some delay between the dialogue and the live captions is to be expected, in this case, the nature of the program meant that the number, duration and frequency of the captioning delays made it difficult for viewers relying on captions to readily follow or comprehend the program.

Example 4: Latency – no breach

A complaint the ACMA received concerned the quality of a captioning service for a 'breaking' news bulletin that interrupted a scheduled program. The bulletin consisted of a title card with a voice-over, both indicating that the scheduled program was being interrupted. The state Premier then appeared on screen and proceeded to deliver a prepared statement advising the public of a state-wide emergency response to widespread bushfires in several parts of the state. The bulletin ran for 20 minutes, comprising a 15-minute delivery by the Premier followed by questions from assembled journalists.

The ACMA found that captions were not displayed for several seconds after the bulletin started. Captions appeared 8 seconds after the Premier's first spoken words and, for the duration of the speech, the delays ran between 4 and 17 seconds. During the transition from the Premier's speech to questions from journalists, some of the transitional comments were not captioned, but all the questions and answers were subsequently captioned.

The ACMA noted the long periods of latency but found they did not affect comprehensibility. Most of the delayed captions occurred after the Premier had been introduced and there were no changes in footage of him speaking. The Premier delivered the statement in a matter-of-fact manner and did not often change his tone to emphasise specific points. In this case, the ACMA found that the delays in captioning did not affect the ability of viewers to comprehend the broadcast.

Style and format of captions

The Standard does not specify that captions should be in a particular style or format. For example, there is no requirement for captions to be a certain size, font or colour. This allows broadcasters some flexibility when deciding how captions are presented on screen in order to meet the requirements in the Standard.

The ACMA recognises providers of television programs, including broadcasters that are subject to the Standard and those providers in unregulated media, adopt a variety of methods of presenting captions. The Standard is not intended to constrain broadcasters from innovating with the presentation of captions, as long as the captioning is meaningful to viewers as required by the Standard.

However, as the Standard does require that the captioning service for a program must use captions that are readable and comprehensible, it is appropriate for the ACMA to have regard to the size, font, colour, and other features of captions when assessing a captioning service against the quality requirements in the Standard.
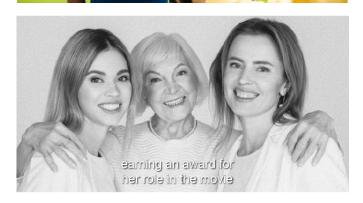
Readability of captions

The Standard states that, when determining whether captions are readable, the ACMA must consider 'whether colour and font is used in the captions in a way that makes them legible' (sub-paragraph 10(b)(i)). Therefore, when considering whether a captioning service is readable, we will have regard to the size, font, and colour of the captions. The following are examples of where we may find that the style or format of captions may mean that the captioning service is not readable:

> the captions are difficult to read because they are the same colour as the background over which they are placed

> the captions are too small to be read by the average viewer sitting a reasonable distance from the television screen

> the font used for the captions makes it hard for an average viewer to distinguish between letters or words.

Readability – Example of white captions over a white background being difficult to read on screen.

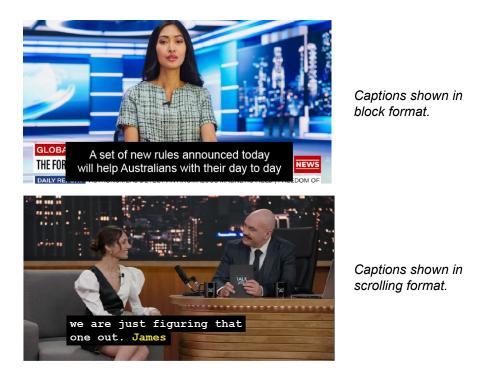*Readability – Example of white captions used over black-and-white on-screen pictures being difficult to read.*

Block or scrolling formats

Captions are also typically displayed in either 'block' or 'scrolling' format. Block captioning is where whole lines of captions appear all at once, while scrolling captions appear in a line, one word at a time.

As with style and format, the Standard does not require captions to be displayed in a block or scrolling format, provided they are readable.

Captions shown in block format.

Captions shown in scrolling format.

Example 5: Format of captions – no breach

A complaint the ACMA received concerned the quality of a captioning service for an hour-long daytime television drama. The complaint alleged that captioning for the program was disrupted and that when it resumed, the block captions were replaced by a scrolling format, which the complainant claimed was harder to read.

On seeking comments from the relevant broadcaster, the ACMA was advised that the program was usually broadcast with pre-prepared captioning presented in a block format. However, on this occasion, due to human error, captions for the final 15 minutes of the program were not included with the captioning service when it was uploaded for broadcast.

The broadcaster advised that its captioning staff monitoring the broadcast had noticed the error and promptly contacted the captioning provider to live-caption the final 15 minutes of the program. The broadcaster confirmed that this resulted in a short period when captioning was not provided, followed by captions being presented in scrolling format for the remainder of the program.

In considering the change of captioning from block format to scrolling format in the context of the final 15 minutes of an hour-long program, the ACMA formed the view that it did not affect the readability and comprehensibility of the captions. As the quality of the captions, regardless of their format, was meaningful to deaf and hard-of-hearing viewers, the ACMA made a no breach finding in this instance.

Comprehensibility of captions – identifying and distinguishing speakers

The Standard states that, when determining whether captions are comprehensible, the ACMA must consider 'whether the captions clearly identify and distinguish individual speakers, including off-screen and off-camera voices' (sub-paragraph

12(b)(i)). The ability for viewers using captions to distinguish between individual speakers is a necessary element of a comprehensible captioning service.

When considering whether a captioning service is comprehensible, the ACMA will have regard to the colour of the captions, if this is the method used by a broadcaster to identify and distinguish individual speakers. However, as noted above, the Standard is not intended to constrain broadcasters in this regard and the ACMA recognises that other methods may be used to identify and distinguish different speakers, including hyphens, chevrons or other punctuation marks.

The Standard does not specify that a particular method must be used to identify and distinguish individual speakers. Whichever method is chosen by a broadcaster to identify and distinguish different speakers (provided the chosen method does not otherwise affect the readability, accuracy or comprehensibility of the captions), it should be used consistently throughout a program to reduce any risks of non-compliance with the requirements relating to quality in the Standard.

Comprehensibility – different colours used to identify different speakers.

Comprehensibility – a new speaker is identified with a # symbol.

Comprehensibility – a new speaker is identified by name.

Effect of television reception equipment

The ACMA has received enquiries highlighting that, in some cases, television receivers may change or distort the style or format of captions for programs transmitted by broadcasters. The ACMA recognises that this is beyond the control of a broadcaster and, when enquiries of this nature are received, the ACMA generally works with the relevant broadcaster to assist the viewer to resolve the issue.

Varying these guidelines

It is intended that the guidelines will be reviewed and updated from time to time, and we invite comments from users of captions, members of the deaf and hard-of-hearing community, captioning providers and broadcasters about how these guidelines can be improved. This could include whether existing material in the guidelines requires clarification or whether the guidelines should cover additional topics.

Any substantial changes to the guidelines would be made subject to consultation with stakeholders.

Making a comment

Comments about the guidelines can be made via:

> email to captioning@acma.gov.au. Please include the words 'captioning guidelines' in the subject heading of your email

> the ACMA's online enquiry form. Please include the words 'captioning guidelines' in the body of your message.

> post – comments can be sent to:

The Manager
Media Diversity Section
Australian Communications and Media Authority
PO Box Q500
Queen Victoria Building NSW 1230