

# Caption requirements



**CAPTIONS** are defined as the displayed text of the speech within a video, in the language native to the original video.

- They are superimposed onto the video when it is played.
- They can also include descriptions of all significant audio content and non-speech information, such as audience laughter, background noises or a dramatic music.

**Captions must be word-for-word of the scripted content,**

except when intentionally creating simplified captioning for a relevant target audience, e.g. people with cognitive disabilities.

## HOW TO WRITE THEM

- **SPELLING:** They should use conventional spelling.
  - Do not write words phonetically, except when phonetic spelling is essential to understand the context.
  - When speech is inaudible or difficult to perceive clearly, captions should say “inaudible”.
- **SOUNDS** should be described by themselves, not the actions causing them.
- **MUSIC** should be identified by title and artist (whenever possible), unless that would be inappropriate to the content.
- **LENGTH:**
  - Captions should be limited to a maximum of 2 lines of text.
  - A 3-line format is only acceptable if the first line is used ~~specifically~~ for speaker identification and it does not block important visual information.
- **NUMBER OF CHARACTERS** should not exceed 32 per line.



## TIME ON SCREEN

- It should take into account the number of words: 0.3 seconds per word (if possible).
- A minimum of 1 second under all circumstances, regardless of the length of the caption.

## APPEARANCE

Users should be able to customize the typeface, size and colour of captions.

- Default font: a sans-serif font, not bold.
- Default font size: 22 pt.
- Default colour combination: white text on a black background.
- Default contrast ratio between the colours of the text and the background: at least 3:1.