

# Summary and Checklist: Audio, Video, Animations, and Motion

## Summary

### Four main methods of making video and audio accessible:

- **Captions:** Text synchronized with the media, for people who cannot hear the spoken words.
- **Transcripts:** The full text of the spoken words and important visual information in the media file, to read as an alternative to watching or listening to the media file.
- **Audio descriptions:** A version of the media file with a narrator explaining important visual information (such as unspoken actions and events) for the benefit of people who cannot see what's happening.
- **Sign language interpretation:** Video of an interpreter, synchronized with the media file (or in the same video frame), for the benefit of people who know sign language.

### Additional considerations:

- **Clear audio:** Prominent background sounds make it difficult for people who are hard of hearing to understand the spoken words.
- **Preventing seizures:** Flashing visual effects in videos can cause seizures in some people.
- **Preventing dizziness:** Some people with vestibular disorders can become dizzy, nauseous, or disoriented when there is excessive motion in video or animated content.
- **Accessible media player:** The media player itself has to be keyboard-accessible and needs to communicate the names, roles, and values of the controls and the states of the controls.
- **Preventing auto-play audio interference with screen readers:** Media players that auto-play can interfere with screen reader users' ability to hear their screen reader.

## Checklist

### Captions and Transcripts

#### When Captions Are Needed

- Prerecorded multimedia (video plus audio) files **MUST** include synchronized captions.
- All live multimedia (video plus audio) events that contain dialog and/or narration **MUST** be accompanied by synchronized captions.
- All live audio-only content that contains dialog and/or narration **SHOULD** be accompanied by text-based synchronized captions.

#### When Transcripts Are Needed

- A text transcript **SHOULD** be provided for multimedia (video plus audio) content.
- Prerecorded audio-only content **MUST** be accompanied by an easily-reachable text transcript.
- A text transcript describing the visual aspects of the video **SHOULD** be provided for video-only content.

#### What to Include in Captions and Transcripts

- **Captions** **MUST** be verbatim for scripted content (except when intentionally creating simplified captioning for a relevant target audience, e.g. people with cognitive disabilities).
- **Transcripts** **MUST** be verbatim for scripted content.
- **Captions** and **transcripts** **SHOULD** be verbatim for unscripted or live content (with the optional exception of stuttering or filler words -- like "um" -- when captioning the filler words reduces reading comprehension of the **captions** or **transcript**).
- Important visual events **MUST** be described in the **transcript**.
- Important background sounds **MUST** be conveyed in **captions** and **transcripts**, preferably in [brackets] or (parentheses).
- Speech that is spoken off-screen **MUST** be captured in **captions** and **transcripts**.
- The identity of the person speaking **MUST** be identified in **captions** and **transcripts**.
- **Captions** and **transcripts** **SHOULD** use punctuation to convey emphasis whenever possible, rather than write extra text to explain the emphasis.
- **Captions** **SHOULD** use conventional spelling, rather than attempt to write words phonetically, except when phonetic spelling is essential to the meaning of the content.
- **Captions** and **transcripts** **MUST NOT** reveal intentionally-withheld information in the content before the appropriate time.
- Music **SHOULD** be identified by title and artist whenever possible in **captions** and **transcripts**, unless doing so would be inappropriate to the content.
- Important music lyrics **SHOULD** be included in **captions** and **transcripts**, if relevant to the meaning of the content.
- When speech is inaudible or difficult to perceive clearly, **captions** and **transcripts** **SHOULD** say so, using neutral language.

# Summary and Checklist: Audio, Video, Animations, and Motion

## What to Include in Captions and Transcripts (contd.)

- Strong language **SHOULD** be retained and not edited out of **captions** and **transcripts**, whenever possible, OR **SHOULD** be bleeped or muted to match style or content requirements of the intended audience or context.
- **Captions** and **transcripts** **SHOULD** indicate when speech is whispered or mouthed.
- **Captions** **SHOULD** describe sounds in terms of the sounds themselves, not the actions causing the sounds.

## Visual Presentation of Captions

- Captions **SHOULD NOT** exceed three lines on the screen at a time.
- Caption line breaks, when necessary, **SHOULD** be inserted at logical points between phrases, rather than in the middle of a phrase.
- Captions **SHOULD** be typed in mixed case.
- The default font for captions **SHOULD** be a sans-serif font.
- The maximum number of characters per line of captions **SHOULD NOT** exceed 32 characters.
- Captions **SHOULD** remain on the screen for a minimum of 1 second under all circumstances, AND **SHOULD** take into account the number of words, at 0.3 seconds per word, if possible.
- Captions **SHOULD** be positioned to not obscure on-screen text, people's faces, or other important visual information.
- Captions **SHOULD** be precisely synchronized to the audio, except when doing so would make captions difficult to read.
- The default color combination **SHOULD** be white text on a black background.
- The default caption font color and background **MUST** be a minimum of 3:1 (assuming a minimum font size of 18 point).
- The default font size for captions **SHOULD** be at least 22pt.
- The default font weight for captions **SHOULD** be normal (not bold).
- Colors in captions **MUST** not be used as the only way to convey meaning.
- Italics or all caps **MAY** be used for emphasis in captions when punctuation alone does not convey the full meaning.

- Quotation marks (or optionally italics or underline, if supported in the caption format) and mixed case capitalization **SHOULD** be used to designate titles (e.g. of books or movies) when appropriate in captions.
- The last caption frame **SHOULD** be removed from the screen during long silent intervals.
- Any gaps in captions **SHOULD** be a minimum of 1.5 seconds.
- Periods of silence **SHOULD** be noted as such when the visual content gives the impression that there may be important sounds or speech.
- Additional on-screen time for a caption frame **SHOULD** be added when the caption contains unfamiliar or uncommon words or phrases.
- Additional on-screen time for a caption frame **SHOULD** be added when there is a lot happening visually on the screen.

## Visual Customizability of Captions

- Users **SHOULD** be able to customize the visual appearance of captions.

## Caption File Format

- Multiple caption file formats **SHOULD** be provided.
- At least one of the file formats **SHOULD** be in a WebVTT file format.

## Audio Descriptions

### When Audio Descriptions Are Needed

- Prerecorded multimedia (video plus audio) **MUST** include audio descriptions.
- Prerecorded video-only files **MUST** include audio descriptions.
- Audio descriptions **MAY** be provided for live multimedia (audio plus video) content.
- Audio descriptions **MAY** be provided for live video-only content.

### Extended Audio Descriptions

- Where pauses in foreground audio are insufficient to allow audio descriptions to convey the sense of the video, extended audio description **SHOULD** be provided for all prerecorded multimedia (video plus audio) content.

# Summary and Checklist: Audio, Video, Animations, and Motion

## Sign Language Interpretation

- Determining When Sign Language Interpretation is Needed
- Sign language interpretation MAY be provided for prerecorded multimedia (video plus audio).
- Sign language interpretation MAY be provided for prerecorded audio-only content.
- Sign language interpretation MAY be provided for live multimedia (video plus audio) content.
- Sign language interpretation MAY be provided for live audio-only content.

## Media Player Accessibility

### Keyboard Accessibility

- All functionality of a media player MUST be available to keyboard users

### Screen Reader Accessibility

- All controls of a media player MUST present the correct names, roles, and values to screen reader users.

### Captions, Transcripts, and Audio Descriptions

- Media players SHOULD allow users to access captions, transcripts, and audio descriptions.

### Customizability

- Media players SHOULD allow visual customization of captions.
- Media players SHOULD remember user preferences.
- Media players SHOULD allow full screen video.

## Background Sounds

### Background Sounds in Media

- Background sounds in prerecorded audio-only and prerecorded multimedia content SHOULD be minimized (20dB lower than foreground sounds, except for occasional sounds of no more than 2 seconds) or eliminated during narration or dialog, or a method must be available to turn off background sounds.
- Background sounds in live audio-only and live multimedia content SHOULD be minimized (20dB lower than foreground sounds, except for occasional sounds of no more than 2 seconds) or eliminated during narration or dialog, or a method must be available to turn off background sounds.

### Background Audio on Web Pages

- A mechanism MUST be provided to stop, pause, mute, or adjust volume for audio that automatically plays on a page for more than 3 seconds.

## Flashing Content

- A page MUST NOT contain content that flashes more than 3 times per second unless that flashing content is sufficiently small, the flashes are of low contrast, and do not violate general flash thresholds.

## Animations and Motion

### Parallax Effects

- Parallax effects SHOULD be kept to a minimum, in terms of the total number of parallax effects, the amount of parallax within each individual effect, and the size of the area affected.
- All content and features within parallax scrolling content MUST be accessible by keyboard.
- The contrast of the text against all parts of a moving background MUST be a minimum of 4.5 to 1 for small text or 3 to 1 for large or bold text.

### Background Videos and Animations

- Important content MUST NOT be conveyed through background videos and animations, unless users have full control over playback, and full access to any required captions, transcripts, and audio descriptions.
- A method MUST be provided to pause, stop, or hide any background videos or animations that begin playing automatically and which last 5 seconds or more.
- A method SHOULD be provided to pause and restart background videos and animations.
- The contrast of the text against background videos and animations MUST be a minimum of 4.5 to 1 for small text or 3 to 1 for large or bold text.
- Movement within background videos and animations SHOULD be minimal.
- Background videos and animations SHOULD NOT contain sounds.

### Auto-Play

- A method MUST be provided to pause, stop, or hide any media content that begins playing automatically and which lasts 5 seconds or more.