

Transcripts Best Practices

Transcripts are written document that are created to be read aloud by screen readers. Since this is their purpose, transcribers not only need to record spoken word but also include any other audio that is happening, as well as any visuals on screen.

A transcript is different from a script or screenplay. While they both document audio and visuals, transcripts are specifically verbatim and must follow consistent formatting and guidelines.

There are three parts to creating a transcript. The first is to generate the base transcript, then we must edit thoroughly before finally implementing the transcript on a web page.

Generating the Base Transcript

There are two ways to generate a transcript, one is using YouTube and the other is using Adobe Premiere Pro. There are pros and cons to each. The main difference is in Premiere, you need the video to be downloaded on your computer and it is a longer process. However, it is a much more accurate auto-generated transcript.

YouTube

1. Open the video you would like to create a transcript for on YouTube. It doesn't matter if you're signed in or not. Underneath the video but above the description, there are three dots.
2. Click the dots and choose "Show transcript" from the options. Sometimes it loads weird, and you have to click "Show transcript" again in the transcript section in the description.
3. The auto-generated transcript will pop up with timestamps. Click the three little dots above the transcript and choose "toggle timestamps." Unfortunately, there is no way to export the transcript, so after the timestamps are turned off, copy and paste the entire transcript into a word document.
4. Unlike Premiere, there is no auto-capitalization or punctuation, and it is not in paragraph form. Before editing the transcript further, you must fix the spelling and capitalization mistakes and find where sentences end/begin.

Premiere

1. Using your YouTube Editor, go to the YouTube Studio and Webster's channel dashboard. Use the search bar to find the correct video in the Webster University channel.
2. Click the title and when the page opens, click the three dots next to the save button. From there, download the video to your computer.
3. Open Adobe Premiere Pro and choose to create a new project. It is going to ask for footage in order to start, so choose the video from wherever you have it saved. Click "create" which imports the video into the program.

4. Depending on how you've used Premiere in the past, you will probably have to change the workspace in order to access some of the features. In the top bar, click Window, hover over Workspaces when the dropdown opens, and choose the workspace "Captions and Graphics."
5. Now you should have 3 open boxes with nothing in them, and one box with the video in the lower left corner. Click and drag the video clip into the timeline, where the box says, "Drop media here to create sequence."
6. The video should have opened in the right box with the left box titled "Transcribe source clips" and offers a blue Transcribe button.
 - a. If you do not see the title or the button, make sure "Text=" is chosen at the top of the box, with "Transcript" chosen underneath.
7. Click the blue Transcribe button and let Premiere do its thing. Once the transcript loads, click open the three dots and make sure "Show speakers" under Preferences is checked.
8. You can also export the transcript in the same three-dot dropdown. There are a couple export options:
 - a. Export to CVS file which opens the transcript into a version of excel, but it looks a little weird and it is hard to edit.
 - b. My preference is to export to text file and open it through the Notepad app.
 - i. Everyone should already have the Notepad app, as it is considered a Windows Accessory.
9. You can edit in the Notepad app but at some point it will have to be moved over to a Word document.
 - a. I suggest, before any editing, Ctrl+A to highlight everything and then copy and paste the transcript into a new Word document.
10. Unlike YouTube, the document will retain timestamps. You will need to delete those and go through the transcript to edit any obvious punctuation or capitalization errors.

Editing the Transcript

All of Webster University transcripts follow the same guidelines, including the formatting and the use of AP style. However, there are several edits to make before we format the transcript.

First, regardless of the method used to generate the transcript, the basics must be fixed. This includes removing timestamps, basic capitalization, punctuation and adding or removing unnecessary space.

At this point, you need to listen to the video *at least* once, all the way through, while following the transcript. This makes it easier to identify and separate speakers and allows you to listen for mistakes in the generated dialogue.

Once the basic edits and changes to the dialogue are complete and each speaker is identified by name, you can begin adding the rest of the transcript. The accessibility and the formatting begin at this step.

Accessibility

Editing for accessibility is the most time intensive piece of creating a transcript. There are four components used in Webster's format of a transcript for accessibility: visual description, audio description, verbal description, speaker identification and text on screen.

YouTube and Premiere both do not add any of the components into the autogenerated transcript. *You* are the one who creates these additions – and thoroughly.

Here are some things to remember when creating an accessible transcript:

- Any time a component appears, you must describe it. Every time there is a graphic or visual change, new sounds or text on screen, it must be transcribed into written words. No cherry picking.
- Timing is important. If the sound of the whistle happens after someone is done speaking, make sure the brackets go after the dialogue. The order of the transcript follows the flow of the video almost perfectly.
 - The only time it can be changed slightly is when someone is giving a long monologue while a visual changes. See “*Video Types*” for more information.
- Transcripts are for screen readers. If something doesn't make sense being read aloud, or something is out of order, change it.

Visual Description

Visual description, which uses written words, depicts anything that is visually happening on screen. For every graphic or major visual change on screen, you must describe what it is, or what has changed.

The formatting of a visual description is that the description is in brackets and is **not** bolded. For example:

[Footage of an orange cat playing with a toy mouse.]

[Cut to...]

[Graphic of...]

When it is an educational video, as most we work with are, you need to be specific and fully describe any graphics on screen. Saying there's a graphic of a Ven diagram of

something is not enough. The visual's meaning must come across in written word as well.

Audio Description

Adjacent to visual description, audio description describes whatever happens in the audio, excluding spoken words. This includes music, other sounds or even the lack of audio. The formatting is the same as visual description, so for example:

[Music] or [Music begins playing slowly, getting louder as seconds pass.]

[A bird squawks loudly.]

[Airhorn]

A few things to note for both visual and audio descriptions:

1. The amount of detail needed in these description depends on the type of video. See the “*Video Types*” section.
2. Regardless of the number of details, break the descriptions up into smaller sections that are easier to digest.
3. Also, brackets that have complete sentences should end with a period. It is not necessary for singular words and/or phrases.

Verbal Description

Verbal description, also known as captioning, is the translation of speech into written word. Captioning is relatively self-explanatory, but when editing the transcript it is important that the captions are word-for-word. This means if the speaker repeats themselves or they don't make sense, you cannot change it. For example:

Speaker

I, well, we – we tried our best.

As for formatting, there really isn't any. The verbal description is just written in normal sentence case without font stylization. However, you should use bold to emphasize the written speech whenever the speaker emphasizes vocally.

Speaker Identification

A commonly left-out piece of a transcript is the identifier of whoever is speaking. Whenever a person speaks, their name must *precede* captioning. If the person speaks again, their name must still be written down before anything they say.

The formatting is simple. Speaker identification is given its own line, separate from any other words, and the name is always bolded. Additionally, the name does not need to be added every time the visual changes, or when the same person continues speaking. Such as:

Person A

[Description]

Person A – this is *not* needed.

Or:

Person A

[Description]

Person B

Person A – this *is* needed.

If you do not know the person's name, just use something generic such as Person A or Person B. Speaker identification is a little more complicated for interviews and story-driven videos. See "*Video Types*" for more information.

The formatting is simple. Speaker identification is given its own line, separate from any other words, and the name is always bolded.

Text On Screen

This one is pretty self-explanatory. Any text on the screen must be written out in the transcript in its own line, bolded, with no hyphen in between on screen. It can often be combined with audio and visual descriptions, where "text" is used generally in the description. For example:

[Beginning with a black display, text appears on screen using a glitch transition.]

Text on screen: WEBSTER GAME DESIGN

Video Types

Interviews

Most of the videos that Webster University works with are Program Spotlights. These videos are basically multi-person interviews that speak positively about a Webster program, which is why there is extra information to keep in mind:

1. Speaker Identification

2. Component Changes
3. Amount of Detail

Speaker Identification – Most times, when a video introduces a new interviewee, their name and position or who they are will be shown as text on screen and then they begin speaking. These are two different components, the text on screen and the speaker identification. For example:

Text on screen: Kathy Foy, School of Education Faculty Member

Kathy Foy

Once that person is introduced the first time, the next time they are the speaker, *only* their name needs to be written. This is regardless of if the text on screen, or their position, returns or not.

Component Changes – The speech of an interviewee can get extremely long, especially if they ramble. Sometimes, there will be an added graphic or a visual change. Do not cut to the visual description in the middle of the monologue. Wait until a sentence ends, cut and describe the visual or the changes, then immediately go back to documenting the speech. One example:

Sheila Hobson

And using this engineering design process starts off with a goal.

[A close up on “the goal” on the graphic.]

And how do we accomplish a task or find a product that will help mankind?

The best way to handle these situations is to read it to yourself, including the cuts, and decide: *does it make sense?* If not, you can move the description around, regardless of timing.

Amount of Detail – In a transcript, details can range from the absolute basic to extended, flowery prose. In an interview-type video, there are less details than other types because it’s mainly the person being interviewed. There isn’t much visually changing on screen unless there are graphics or extra visual shown.

Whatever detail is used must be consistent throughout the entire video. This mainly includes names and descriptions of recurring visuals. Also, even if there are graphics or visuals in an interview video, these still must be described accurately and the visual’s meaning must come across in written word. This doesn’t mean you have to spend several lines describing it or including all copy on a graphic, just enough to get the meaning across.

Videos with No Dialogue

Often times, there will be a video with no dialogue. This can have either all the text on screen or a video with only visuals. In most cases, there will be text on the screen with

faint music in the background. You will have to create the transcript from scratch with all the descriptions and word-for-word text on screen.

As for the amount of detail in the descriptions, go with what feels right to you. If the visuals are relatively generic, feel free to keep detail to a minimum. See the 2023 Welcome Week video on the [Student Life](#) page, for instance.

If the whole point of the video is specific visuals, detail should be on the higher end and be able to portray the visuals for screen readers. For an example, check out the transcript on the [Games and Game Design](#) webpage or the third-party audio description: [Immortals Fenyx Rising](#).

Story-Driven Videos

Webster does not often utilize story-driven videos with a beginning, middle and ending. Still, it is important to note if there is a narrator, use “Narrator” as the speaker identification that precedes any of their speech.

Also, as the person creating the transcript, you are an omniscient presence, while screen readers will only read from beginning to end. This is why if you know what the speaker identification is, but they have not been introduced to the audience, do not use it until that introduction. Instead, title them by a descriptor, such as Blonde Woman, or something like Person A until the name is revealed.

Editing Reminders

A transcript needs to start a new line every time there is a new component, when a description bracket gets too long or when there is a substantial change in visuals.

Whether the description is long or short, each description bracket must have an opening and a closing bracket.

Sentence case is used throughout the entire transcript, and for other style guidelines, AP style should be used. This includes: no hyphen in on-screen, no oxford comma, no ampersands, etc. See the AP Style Fun Tips document for more.

Implementing the Transcript

This part of the process is mainly for web team members, as they are the ones who will be putting the transcripts directly on the website.

1. When adding the transcript to a video in Omni, you will have to use the Video with Transcript snippet. If the video is currently there, open it on YouTube to save the URL, and then delete the video and any snippet needed in Omni.

2. Give the video an H2 title and a video title tag, which is basically a quick description of a video. It is equivalent to an alt tag for images. Pre-content is optional, but makes the snippet look better.
3. Make sure to include the video link by copying the end of the YouTube URL (after *watch?v=*) and paste it at the end of the video link in the snippet. It should look like this <https://www.youtube.com/embed/1Kdf9Zqi8dY> with the highlighted portion being what changes for each video.
4. Save and publish. The video should resize automatically, but it is a good idea to double check anyway.