

Video Captioning Using YouTube

Andi Dietrich

May 8th, 2018

What is captioning?

For this presentation, I'm going to use "captions/captioning" to refer to all kinds of captioning and subtitling. But there are differences.

What's the Difference?

There are many different terms for captioning, and they all mean similar things, but they are not all the same thing.

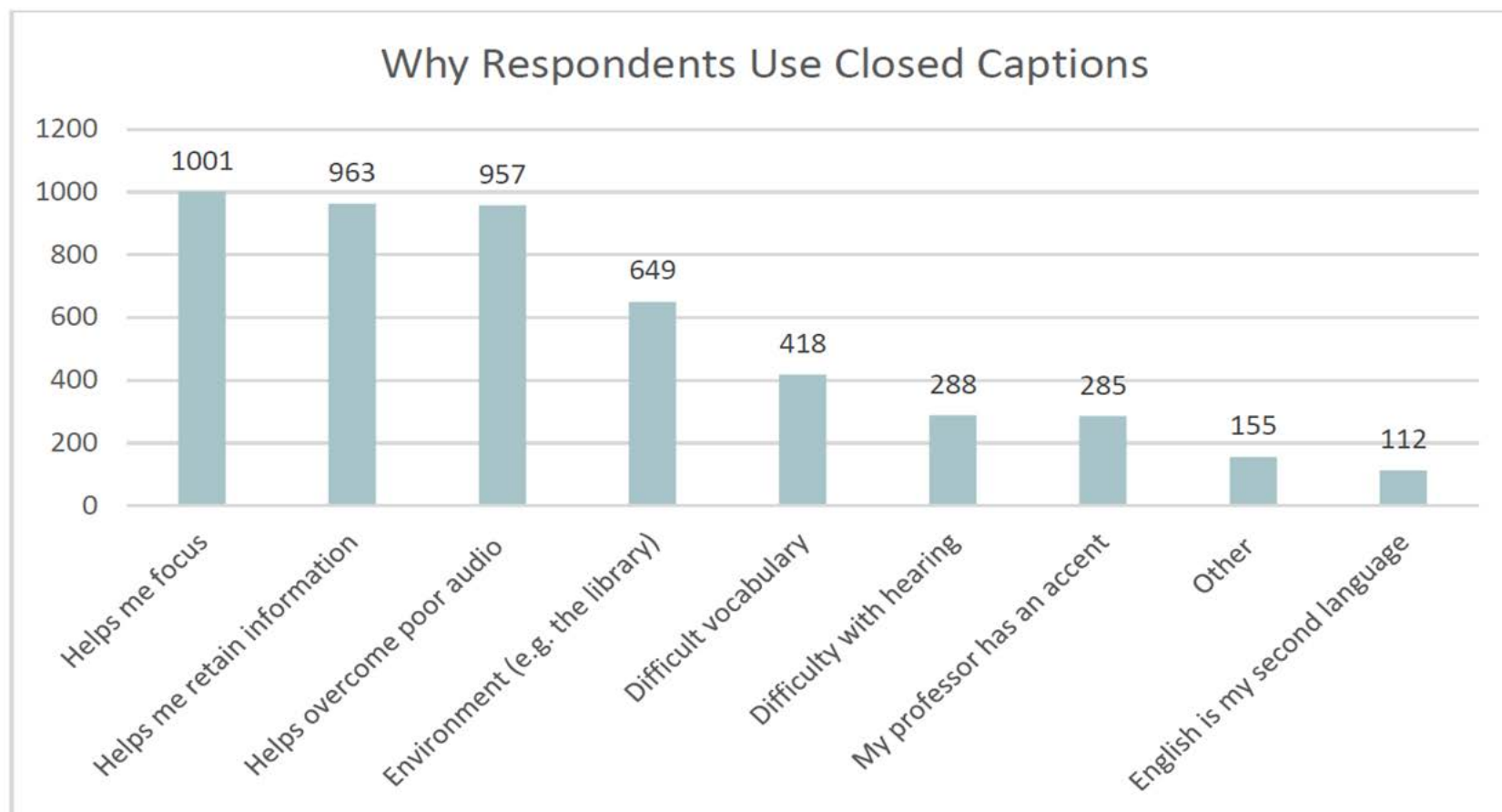
- **Transcripts** are the full, accurate, complete text of a video (or sometimes an audio production, such as a podcast). Some transcripts include only spoken dialogue, others include descriptions of non-dialog audio and song lyrics.
- **Subtitles** are text derived from either a transcript or screenplay of the dialog or commentary in films, television programs, video games, and the like, usually displayed at the bottom of the screen, but can also be at the top of the screen if there is already text at the bottom of the screen. They can either be a form of written translation of a dialog in a foreign language, or a written rendering of the dialog in the same language, with or without added information to help viewers who are deaf or hard of hearing to follow the dialog, or people who cannot understand the spoken dialogue or who have accent recognition problems.

- **Closed captioning** is the American term for closed subtitles specifically intended for people who are deaf or hard of hearing. These are a transcription rather than a translation, and usually contain descriptions of important non-dialog audio as well such as “[sighs]”, “[wind blowing]”, “[“SONG TITLE” playing]”, “[kisses]” or “[door creaks]” and lyrics. From the expression “closed captions” the word “caption” has in recent years come to mean a subtitle intended for the deaf or hard of hearing, be it “open” or “closed”. In British English “subtitles” usually refers to subtitles for the deaf or hard of hearing (SDH); however, the term “SDH” is sometimes used when there is a need to make a distinction between the two.
- **Subtitles for the deaf or hard-of-hearing (SDH)** is an American term introduced by the DVD industry. It refers to regular subtitles in the original language where important non-dialog information has been added, as well as speaker identification, which may be useful when the viewer cannot otherwise visually tell who is saying what.

Subtitles exist in two forms; open subtitles are “open to all” and cannot be turned off by the viewer; closed subtitles are designed for a certain group of viewers, and can usually be turned on/off or selected by the viewer.

Who Uses Captions?

Many people use captions, even if they can hear. You might read captions while watching TV on mute at the gym, or if you need to watch a video while you're studying in a quiet library, for example. This graph shows some of the reasons students report using captions. Notice that hearing difficulties are far from the top of the list.



Some Notes on Quality

Quality captions are:

Accurate:

Errorless captions are the goal for each production.

Consistent:

Uniformity in style and presentation of all captioning features is crucial for viewer understanding.

Clear:

A complete textual representation of the audio, including speaker identification and non-speech information, provides clarity.

Readable:

Captions are displayed with enough time to be read completely, are in synchronization with the audio, and are not obscured by (nor do they obscure) the visual content.

Equal:

Equal access requires that the meaning and intention of the material is completely preserved.

Ways to Create Captions

YouTube has many different ways you can use their platform to create captions. These include:

- Creating and editing YouTube automatic captions
- Starting with a transcript
- Transcribing a video in YouTube

How to Upload a Video to YouTube

1. Sign in to YouTube.
2. Click the Upload button at the top of the page.
3. While the video is uploading, you can edit both the basic information and the advanced settings of the video. You can have a title up to 100 characters and a description up to 5,000 characters.
4. Click Publish to finish uploading a public video to YouTube.
 - If you set the video privacy setting to Private or Unlisted, click Done to finish the upload or click Share to privately share your video.
5. If you haven't clicked Publish, your video won't be viewable by other people. You can always publish your video at a later time in your Video Manager.

YouTube Privacy Settings

You can change the privacy settings on your videos so that they cannot be viewed by others, or so that they cannot be viewed by anyone who does not know the video link.

- Public videos and playlists can be seen by and shared with anyone.
- Private videos and playlists can only be seen by you and the users you choose.
- Unlisted videos and playlists can be seen and shared by anyone with the link.

About Automatic Captions

Once you've uploaded a video to YouTube, automatic captions will be created. Once automatic captions are available, they'll automatically be published on the video.

Problems With Automatic Captioning

Speech recognition technology is not perfect! Some things that can cause YouTube's speech recognition to fail include:

- Accents
- Dialects
- Background noise or poor quality audio
- Specialized terminology

Additionally, automatic captions will not have correct capitalization or punctuation, and lines will be broken in unusual ways that can make the captions harder to understand.

You should always review automatic captions and edit any parts that haven't been properly transcribed.

Click the image below to see an example of a video with unedited automatic captions:



These are the first 4 captions from that video. You can see that there are some obvious errors in content, capitalization and punctuation, and logical line breaks.

president pizza mr. vice president
governor congressman Thomas senator

Wylie and congressman nila mr. Webb Bell
scientists distinguished guests ladies

and gentlemen I appreciate your
president having made me an honorary

visiting professor and I will assure you
that my first lecture will be very brief

And here's the same video, with edited captions. Click the image below to see the video:



These are the same 4 captions, but edited for accuracy and clarity, with line breaks in appropriate places and punctuation and capitalization added in correctly.

President Pitzer,
Mr. Vice President, Governor,

Congressman Thomas, Senator Wiley,
and Congressman Miller,

Mr. Webb, Bell, scientists,
distinguished guests

and ladies and gentlemen:

I appreciate your president having
made me an honorary visiting professor,

and I will assure you that
my first lecture will be very brief.

How to edit YouTube captions

1. Go to your Video Manager by clicking your account in the top right and choose Creator Studio from the drop- down menu.
2. Next to the video you want to edit, click the drop-down menu next to the Edit button.
3. Select Subtitles/CC.
4. If automatic captions are available, you'll see Language (Automatic) in the "Published" section to the right of the video.
5. Click on English (Automatic) to open the subtitles.
6. Click Edit to begin editing.

Transcribe and set timings: English

All changes saved in Drafts

Delete draft

Save changes

Actions ▾

Keyboard shortcuts | Help

Type subtitle here then press Enter +

0:03.0	President Pitzer, Mr. Vice President	A
0:11.2	Governor, Congressman Thomas,	
0:11.2	Senator Wylie and Congressman Miller,	B
0:17.0	Mr. Webb, Mr. Bell,	
0:17.0	scientists, distinguished guests ladies	C
0:20.9	and gentlemen;	
0:20.9	I appreciate your	D
0:24.6	president having made me an honorary	
0:24.6	visiting professor and I will assure you	E
0:30.2	that my first lecture will be very brief	
0:30.2	I am delighted to be here and I'm	F
0:34.9	particularly delighted to be here on	

Pause video while typing

- A. The blue-shaded area shows the current caption.
- B. The numbers show the start and end times of the caption, in minutes:seconds: milliseconds format
- C. The area below the video shows the caption timing.
- D. The red line shows the current position in the video.
- E. The lines below the current caption and the video are the audio waveform.

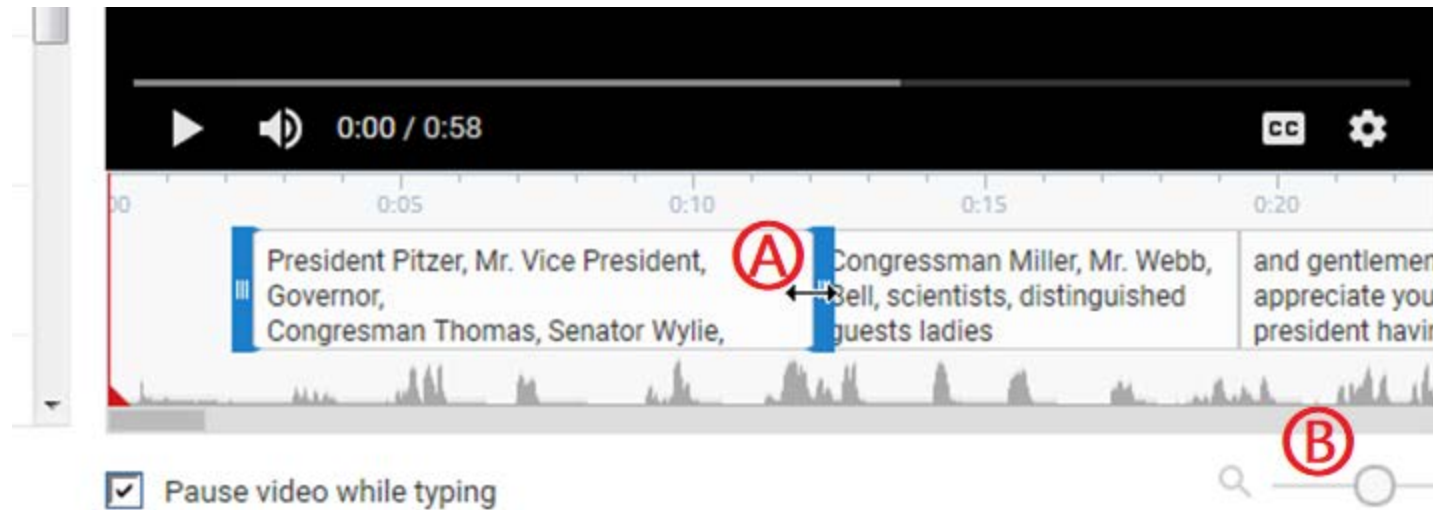
Transcribe and set timings: English

Actions ▾

Type subtitle here then press Enter +

0:02.4 0:12.0	President <u>Pitzer</u> , Mr. Vice President, Governor, <u>Congresman</u> Thomas, Senator Wylie,	✕ +
0:12.0 0:19.3	Congressman Miller, Mr. Webb, Bell, scientists, distinguished guests ladies	
0:19.3 0:24.6	and gentlemen I appreciate your president having made me an honorary	
0:24.6 0:30.2	visiting professor and I will assure you that my first lecture will be very brief	
0:30.2	I am delighted to be here and I'm	

- Click to edit the caption text. The blue border indicates this is the “active” caption.
- As you change the text, you will see it also change on the video.
- To add a line break, press and hold Shift and press Enter. In this example, there is a manual line break between “Governor,” and “Congresman.”
- Red underlines indicate spelling errors (“Congresman”) in this case.



- A.** To change the start or end time of the caption, mouse over the caption until the mouse changes to a double-headed arrow, then click and drag to change the time.
- B.** Drag left and right on the slider bar below the timing window to see more or fewer seconds in the caption timing panel.

To insert a new caption, drag the end of the previous caption to the left, to make it shorter. This will add a gap between the two captions. Then click the plus sign next to the active caption to add a new caption.

Starting with a transcript

If you have a transcript of your video, you can upload it and YouTube will add timing.

Format your transcript file

- Your transcript must be in plain text format
 - You can convert other formats (like Microsoft Word, HTML, PDF) to plain text.

In order to get the best results:

- Use a blank line to force the start of a new caption.
- Use square brackets to designate background sounds. For example,
 - [music] or [laughter].
- Put speaker names on a separate line in brackets. For example, [Alice] or [John].
- Make sure your captions have line breaks in appropriate places.

For more information, see the Captioning Style Tip Sheet at the end of this presentation.

For a much more comprehensive overview of captioning styles, see the DCMP Captioning Key, at http://www.captioningkey.org/quality_captioning.html.

Upload a transcript for timing

Once you've prepared your transcript, you're ready to upload it:

1. Go to your Video Manager by clicking your account in the top right and choose Creator Studio from the drop-down menu.
2. Next to the video you want to edit, click the drop-down menu next to the Edit button.
3. Click "Add new subtitles or CC"
4. Select "English" and then click the button that says English
5. From the menu on the right-hand side of the screen, select "Upload a file."
6. Choose "Transcript", find your file, and press Upload.

Sample Transcript

This is an example of a transcript, with correct formatting, breaks, and speaker identification.

[ALICE]

Hi, my name is Alice Miller
and this is John Brown

[JOHN]

and we're the owners
of Miller Bakery.

[ALICE]

Today we'll be teaching you how to
make our famous chocolate chip cookies!

[intro music]

Okay, so we have all
the ingredients laid out here

Transcribing a video in YouTube

You can transcribe your video in YouTube, and then use YouTube to set the timing.

1. Select Add new subtitles or CC. Choose the language for the subtitles or closed captions you want to create.
You can use the search bar to find languages that don't automatically show in the list.
2. Select Add new subtitles or CC.
3. Type all of the spoken audio in the text field. If you're creating closed captions, make sure to incorporate sound cues like [music] or [applause] to identify background sounds.
4. Click Set timings to sync your transcript with the video.
5. Setting the timings can take a few minutes. While you wait, you'll be brought back to the video tracklist.
Once it's ready, your transcription will automatically be published on your video.

Captioning Style Cheat-sheet

Line Length and Caption Presentation Rate

- Captions should ideally be no more than 32 characters per line, but can be as long as 38 characters per line if necessary, to help retain grammar and allow for more reading time.
- Captions should be on-screen for at least 2 seconds per caption.

Editing

- Captions should be written verbatim except for the following exceptions:
 - Redundant and nonessential information can be removed ("err" or "um," and false starts)
 - Do not caption the same, or nearly the same, information that is already shown onscreen. Caption (reading from board) or similar instead.

Spelling / Capitalization

- Do not emphasize a word using all capital letters except to indicate screaming.
- Be consistent in the spelling of words throughout the media.
- If information is displayed on screen and spoken simultaneously, (for example, a professor talking about an equation written on a whiteboard), make sure the capitalization and formatting of the caption matches the displayed information.

Music

- When captioning music, use objective descriptions that indicate the mood (for example, "(celebratory music)" or "(sad music)"
 - Avoid subjective words, such as "delightful," "beautiful," or "melodic."
- If music contains lyrics, caption the lyrics verbatim.
 - The lyrics should be introduced with the name of the vocalist/vocal group and the title (in brackets) if known.
- Caption lyrics with music icons (♪). Use one music icon at the beginning and end of each caption within a song, but use two music icons at the end of the last line of a song.

Grammar

- When a speaker is interrupted and another speaker finishes the sentence, the interruption should be conveyed by double hyphens.
- If a speaker stops speaking abruptly or has a false start, that should be conveyed by double hyphens.
- When a speaker stutters, caption what is said. (For example, "B-b-but I don't know!")
- Use an ellipsis when there is a significant pause within a caption.
- Use an ellipsis to lead into or out of audio relating to an onscreen graphic.
- When someone is quoting, use quotation marks as you would in other writing.
 - Beginning quotation marks should be used for each caption of quoted material except for the last caption, which should have a beginning and ending quotation mark.

Sound Effects

- Caption background sound effects only when they're essential to the plot.
- Use specific terms to describe sounds. "(cell phone rings)" rather than "(ringing)."

Speaker Identification

- When a speaker's name is known, include their name, in parentheses, on a separate line before the first time they speak.
- When a speaker's name is unknown, identify the speaker using the same information a hearing viewer has (e.g., "female #1," "male narrator").
- If there is only one narrator, identify as (male narrator) or (female narrator) at the beginning of the media. It is not necessary to identify gender for each caption thereafter.

Slang/Foreign Languages

- Caption the actual foreign words. If it is not possible to caption the words, use a description (e.g., [speaking French]). Never translate into English.
- Use accent marks, diacritical marks, and other indicators.
- Indicate regional accent at the beginning of the first caption if relevant. For example, "(French accent)."
- Keep the flavor of dialect and the speaker's language.
- When a word is spoken phonetically, caption it the way it is commonly written.
- If someone is spelling out a word, caption it with hyphens between letters (for example, "word is spelled W-O-R-D.")

Line Division

When a sentence is broken into two or more lines of captions, it should be broken at a logical point where speech normally pauses.

1. Do not break a modifier from the word it modifies. Example:

Inappropriate:

Mark pushed his black
truck.

Appropriate:

Mark pushed
his black truck.

2. Do not break a prepositional phrase. Example:

Inappropriate:

Mary scampered under
the table.

Appropriate:

Mary scampered
under the table.

3. Do not break a person's name nor a title from the name with which it is associated. Example:

Inappropriate:

Bob and Susan
Smythe are at the movies.

Appropriate:

Bob and Susan Smythe
are at the movies.

4. Do not break a line after a conjunction. Example:

Inappropriate:

In seconds she arrived, and
he ordered a drink.

Appropriate:

In seconds she arrived,
and he ordered a drink.

5. Do not break an auxiliary verb from the word it modifies. Example:

Inappropriate:

Mom said I could
have gone to the movies.

Appropriate:

Mom said I could have gone
to the movies.

6. Never end a sentence and begin a new sentence on the same line unless they are short, related sentences containing one or two words. Example:

Inappropriate:

He suspected that his face
turned pale. He knew he

wouldn't be able to speak
if spoken to. Running toward

the void, he halted..

Appropriate:

He suspected that his face
turned pale.

He knew he wouldn't be able
to speak if spoken to.

Running toward the void,
he halted..

Sources

1. [Wikipedia Page: Subtitle \(captioning\)](https://en.wikipedia.org/wiki/Subtitle_(captioning)),
[https://en.wikipedia.org/wiki/Subtitle_\(captioning\)](https://en.wikipedia.org/wiki/Subtitle_(captioning))
(last visited May 7, 2018).
2. [NATIONAL RESEARCH STUDY: Student Uses and Perceptions of Closed Captions & Transcripts](https://www.3playmedia.com/resources/industry-studies/student-uses-of-closed-captions-and-transcripts/)
<https://www.3playmedia.com/resources/industry-studies/student-uses-of-closed-captions-and-transcripts/>
(last visited May 7, 2018).
3. [DCMP Captioning Key](http://www.captioningkey.org/quality_captioning.html)
http://www.captioningkey.org/quality_captioning.html
(last visited May 7, 2018).
4. [YouTube Help, Translation tools > Do-it-yourself transcription & translations. Add your own subtitles & closed captions](https://support.google.com/youtube/answer/2734796)
<https://support.google.com/youtube/answer/2734796>
(last visited May 7, 2018).
5. [YouTube Help, Live streaming guide > YouTube live features, Use automatic captioning](https://support.google.com/youtube/answer/6373554)
<https://support.google.com/youtube/answer/6373554>
(last visited May 7, 2018).