Figure 1: logo for Deafness Cognition and Language (DCAL) Research Centre

Review of different platforms for captioning videos

UCL DCAL – Deafness, Cognition and Language Research Centre

Captions are important for not only for accessibility for deaf and hard of hearing students but can also benefit all students. Not surprisingly, **non-native speakers** perform significantly better on video-related tasks when captions/subtitles are provided than when they are not (Kruger 2013; Kruger et al., 2014; Liao et al. 2020; Ryba et al. 2006). In addition, captions can improve engagement with and attention to online learning materials for all students**, even native language users**.

Many tools for recording lectures and videos provide automated captions via speech recognition. However, none are anywhere near perfect. **Auto-captions for recorded lectures always need to be checked and corrected.**

Teaching staff at DCAL have looked into various options both for creating and editing auto-captions (summarised below). We hope these can aid people and organisations in providing accessible video files. This review is non-exhaustive, and there are other captioning platforms that we have not had an opportunity to test and which may be more adequate for your specific needs.

Recommendation:

Staff at DCAL have different preferences based on their needs, how they record videos, and how easy they find editing and the quality of autocaptions in different platforms. Based on this, our two preferred methods of captioning are through Otter.ai and Echo360. For a non-cost option, we recommend editing captions in YouTube.

Factors to consider

With all options below, the amount of time it takes to edit captions of course depends on the quality of the auto-captions generated in the first place. Based on testing, the main factors to consider are:

- Microphone
    o Use the best available microphone you have. A headset or earbuds with mic will work better than the internal mic on your computer. The closer your mouth is to the mic, the better (so, the microphone on your phone can be another good option). This will help the sound quality on your recording and also therefore the auto-captions.
- Enunciation
    o Speak as clearly as possible. Don't mumble and be careful with speed.
- Turntaking
    o If you are captioning a video with multiple speakers, auto-captioning will be worse with overlapping speech than one person speaking at a time.

Further information

Otter.ai

Upload a video or audio file to Otter. Otter then produces auto-captions which can be edited and exported as a captions files (SRT file). The SRT file and original video file can then both be uploaded as linked files to the platform where video will be hosted (e.g. Moodle or other Learning Management System), and captions can be toggled on or off by viewer when video is played. The estimated amount of time for editing & processing (after initial upload and auto-captioning) is approximately 1.5-2 times real time.

Pros

-Flexible – can record original video with any platform (Zoom, Powerpoint with narration, etc).
-Easy to use interface for editing captions.
-Good accuracy of auto-captions (generally better than YouTube); few errors.
- Otter Premium can be integrated with Zoom Pro if you have subscriptions to both.
-Allows captioning of audio files

Cons

- Otter works with audio. Editing is of a transcript in a separate window, not overlaid onto video.
-The free version (Otter Basic) only allows a limited amount of files to be transcribed, and you need a paid version to download srt files. Therefore a paid version, Otter Premium or higher, is needed. A good alternative is to pay for an individual 1-month subscription.

Echo360

If your institution has access to Echo 360 (known as Lecturecast within UCL), and if you record your original video within Echo 360, it is worth trying the autocaption/autotranscipt correction within Echo360 itself. See Echo360 Accessing the Transcript Editor and Applying Updated Transcripts as Closed Captions for a step-by-step guidance on how to access a transcript and edit captions.

Pros

-Quick workflow if you record videos in this platform.

Cons

-Quality of autocaptions is variable. For some people, it is one of the best. For others, it contains many errors.
-Editing platform is line by line, similar to YouTube and MS Teams, which is more time consuming, especially if you have to correct many errors.

YouTube

Upload a video and select auto-captions. YouTube then produces auto-captions which can be edited directly (via transcript window right next to video showing captions). However, generation of the auto-captions takes an unpredictable amount of time - *up to 10 times real time*. So you will need to do this well

in advance. Estimated amount of time for editing & processing (after the initial upload and the generation of the auto-captions): ~2 times real time.

Pros
-Free
-Editing is of captions directly, overlaid on video.
-Accuracy of auto-captions is good - though generally worse than Otter.

Cons
-Unpredictable processing time for generation of auto-captions after uploading. This can be hours, and without indication of how long it will take, when it's ready, or even whether it's actually working/processing at all.

MS Teams
MS Teams captions will work for those with institutional MS accounts (e.g. University email through Microsoft). Record a video or one-person meeting with slides shared via screenshare. Turn on auto-captions before recording. The final recording will appear in your MS Streams account (log in with institutional account). Auto-captions can then be edited (and revised video + captions saved). Estimated amount of time for editing & processing (after recording and generation of auto-captions): ~3 times real time.

Pros
-No extra cost if you have an institutional account.
-All in one solution if you have not recorded yet.

Cons
-Slow and awkward interface for editing transcript/captions.
-Speaker can be visible but only in a tiny window in the corner (this doesn't seem to be changeable)

MS Powerpoint
Record audio narration over slides. PPT saves audio file for each slide. During playback, presentation can be viewed with auto-captions. Accuracy of auto-captions in our testing was roughly similar to Otter and YouTube. *However - auto-captions cannot be edited.*

Pros
-No extra cost
-Uses software many people already know.
-All in one solution if you have not recorded yet.

Cons
-*Auto-captions cannot be edited.*
-Speaker is not visible – there is no video.

Kapwing

Upload video and produce auto-captions. The accuracy of this seems pretty poor compared to the other options tested.

Pros

-Easy interface (but not really tested/used due to poor quality auto-captions)
-Might be worthwhile if using Kapwing already for video capture & editing (captioning is not its main function)

Cons

-Very poor quality of auto-captions

References

Kruger, J.-L. (2013). Subtitles in the classroom: Balancing the benefits of dual coding with the cost of increased cognitive load. *Journal for Language Teaching, 47*(1), 55–71.

Kruger, J.-L., Hefer, E., & Matthew, G. (2014). Attention distribution and cognitive load in a subtitled academic lecture: L1 vs. L2. *Journal of Eye Movement Research, 7*(5), 1–15.

Liao, S., Kruger, J.-L., & Doherty, S. (2020). The impact of monolingual and bilingual subtitles on visual attention, cognitive load, and comprehension. *JoSTrans, 33*, 70–98

Ryba, K., McIvor, T., Shakir, M., & Paez, D. (2006). Liberated learning: Analysis of university students' perceptions and experiences with continuous automated speech recognition. *E-J. Instructional Science and technology, 9*(1), 1–19.

Conflict of Interest

We have no conflict of interest and no personal financial link to any of the captioning platforms discussed above.