# **Synote: Enhancing Inclusive Learning from Lecture Recordings**

**Mike Wald**

*ECS, University of Southampton, Southampton, United Kingdom*

**Abstract**

# This paper explains how speech recognition captioning with collaborative editing provides affordable transcription/captioning of lecture recordings, supports inclusive learning, retention & recruitment and enables universities to comply with law. It considers how lecture recordings can be inclusively enhanced and what features in a lecture recording system would be beneficial for disabled students. It proposes that all university students learn better when they make their own notes rather than use notes made by somebody else and that a notetaker is not necessary when a time synchronised transcript and slides are available apart from for hearing impaired students who can’t check a recording to correct transcription errors. The paper provides evidence that speech recognition can be more accurate than human transcribers and that we should use students to collaboratively correct caption errors as commercial manual captioning is too expensive for universities.**Introduction**

Cuts to the Disabled Students’ Allowance for notetaking (Johnson 2015) requires universities to fund support for disabled students and HEFCE doubled disability funding to Universities from 2016-18 to help universities move towards a more inclusive approach to learning and teaching to support disabled students (*Supporting disabled students* 2017). The Inclusive Teaching and Learning in Higher Education as a route to Excellence: Disabled Student Sector Leadership Group report (Layer 2017) stated:

“there are some very simple changes that can make a significant difference to student outcomes around inclusive practice … Allow or facilitate the recording of teaching”

This paper considers ways in which the recording of teaching can be enhanced to better support disabled students.

**Notetaking**

Piolat et al. (2005) identified how making notes during a lecture is extremely cognitively demanding requiring to “attend, store, and manipulate information selected from the lecture simultaneously, while also transcribing ideas just presented and processed”. Hanafin et al (2007) reported that note-taking remains a challenge for students in face to face teaching sessions while Boyle (2012) identified that students with Learning Disabilities such as dyslexia were likely to miss important points in multiple sections of a lecture. Burgstahler (2015) clarified how ‘universal access’ to video content was required for students with sensory impairments to make the most of these resources.

James et al (2016) reviewed the notetaking literature as well as surveying 60 disabled students about their confidence and effectiveness with notetaking: “45 % had dyslexia or other Specific Learning Difficulty, 25 % physical difficulties or chronic health conditions, 22 % had a mental health condition, 7 % had sensory impairments and 3 % had social and communication needs. Two students mentioned the fact that they had a hearing impairment but were able to use audio recordings.” The researchers concluded that “While transcripts and captions are often considered necessary for students with hearing impairments, the synchronisation of the text with audio and annotations enables students to use dual channels for processing information in order to increase processing capacity.”

A notetaker is not actually necessary when you have a transcript synchronised with slides and images from the lecture recording as you have access to all the information. Really only hearing impaired students who are unable to check the recording to correct transcription errors need the support of a notetaker or transcriber because all other students can check the transcript if they think there's a mistake because they can listen back to the recording. An advantage of a transcript is it's much quicker to read the transcript than it is to have to replay and pause the recording. University students also learn better when they make their own notes rather than use notes made by somebody else as trying to understand somebody else's notes is much harder than actually understanding your own notes.

# **Captioning**

In 2015 Harvard and Massachusetts Institute of Technology were sued by the National Association of the Deaf for not adequately captioning their videos (Lewin 2015) and the National Association of the Deaf stated: “the selection of such high-profile defendants will send a signal ... so we are suing them first and expect to ensure full online video access at all other universities and colleges across the country. “

In 2011 the National Association of the Deaf also sued Netflix (Whitney 2011) who as a result agreed to caption all their videos (National Association of the Deaf v. Netflix 2012)

The organisation TED has for example captioned their videos since 2009 (TED's Open Translation Project 2009) which has enabled hearing impaired people to follow the talks and allowed everyone to search the videos. It also allows the transcript to be printed which can be read anywhere allowing people to learn faster and in situations with low bandwidth or where silence is required. Captions also help the understanding of talks from non-native speakers and by non-native listeners. The TED Talks use commercial manual captioning which is too expensive for universities to caption and transcribe lectures.

Wald (2017) describes a 2012 study providing 18 lecture recordings of a variety of topics, lengths, recording qualities, and speaker accents to 4 captioning companies which found an average cost of $260 per hour with the most expensive being $407. The study also established that a university editing speech recognition produced transcripts themselves would require average editing effort of 4.10 hours / media hour and paying $15 to $30/hour to people (e.g. students) to edit would cost on average between $60 -$120 /hr without including the cost of the overheads of running the service and paying for the speech recognition software.

One of the arguments that some of the captioning companies use to persuade customers to use their manual captioning service is to say that speech recognition automatic captioning produces inaccurate captions with silly errors. Paying a captioning company to manually correct errors from speech recognition automatic captioning is too expensive for universities as when the author gave the captioning companies a speech recognition generated transcript and asked them to just correct the errors they said it would cost just as much as captioning a recording because they still had to listen to the whole recording.

The argument about speech recognition errors is overplayed because actually speech recognition accuracy continues to improve and Xiong et al (2017) have demonstrated that speech recognition can now even be more accurate than professional human transcribers.

The Equality Act 2010 requires universities to make anticipatory reasonable adjustments (Disability Rights UK Factsheet F56, 2017) and so universities should caption all their lecture recordings rather than only caption a lecture recording if requested by a deaf student.

While universities might claim that paying for commercial manual captioning is not reasonable as it is too high a cost, universities cannot justifiably claim that paying a few pounds an hour for unedited automatic speech recognition captioning is not a reasonable cost.

# **Recording Quality**

It is important for teachers to make a good quality recording and this can be achieved by wearing a wireless microphone and adjusting the recording level to provide a good signal to noise ration. If a teacher uses a fixed lectern microphone and turns or moves away from the microphone to write on the board or walk round the room then the recorded speech level and signal to noise ratio will decrease. If the lecturer repeats any questions or comments or answers from the students then the speech of the students does not needs to be transcribed. It is possible to also record and transcribe the speech of the students using a wireless microphone, either handheld and passed around or throwable (Catchbox 2017) or using an app on a mobile phone (Crowdmics 2017). It is also possible to use mobile phone speech recognition app to transcribe the students’ speech live in the classroom and display it on the main screen and include it in the transcript and enable students to correct any speech recognition errors live in the class (Wald 2012).

# **Students Collaboratively Correcting Speech Recognition Errors**

As the quality of the recording degenerates then speech recognition may still struggle more than human transcribers and a solution to this problem of improving the accuracy of any speech recognition transcription is that it is possible to use students to collaboratively error correct errors and verify the transcript by automatically comparing their corrections (Wald, 2013). Scoring the corrections can increase student motivation, whether through self interest in getting a better transcript, altruism in wanting to help others less fortunate than themselves, or rewards such as micropayments or print credits, badges, high score tables or academic credit. Students correcting errors engage more strongly with the lecture content to improve their learning and so justifying the awarding of academic credit for correcting errors. Universities and students could choose appropriate rewards.

Students correcting errors in their own lectures involves little extra effort while they listen, watch, and read the recording and captions as it is difficult not to notice errors and so it is not like a real job. They also generally know the subject better than a professional captioner as captioning companies do not guarantee to provide a specialist in that subject.

A questionnaire given to 30 students in a class by the author found that approximately one third of the students in the class said they would like micropayments or academic credit but two thirds would not because of various reasons including: “For my own personal revision of lectures”, “You shouldn’t need rewarding for using a tool like this”; “Wouldn’t really need motivation, if I saw a mistake I would correct it”; “It just being there would be enough motivation to use it”; “More accurate transcript gives better search”

# **Features Required To Enhance Learning from Lecture Recordings**

A system like Synote (Synote 2017) works as shown in Figure 1 by the speech of the lecture recordings being transcribed by speech recognition to automatically produce the captions. The images and slides are automatically synchronised with the transcript to enable printing out all of the information. Any errors in the captions can be collaboratively corrected by the students resulting in accurate captions for the recordings and the scoring of corrections can be used as a basis for the student rewards.



Figure 1 Schematic of Synote

Figures 2, 3 and 4 are screen captures of Synote screens that show some of the features.

Figure 2 shows the caption edit button with the caption shown underneath, the button to show the shortcut key list to speed up correction, the searchable transcript, the button to add a clip to the playlist, the synchronised notes and bookmarks that can be created and searched and filtered. Any section of a recording can be bookmarked to create a replayable clip and a playlist can replay selected clips in any order. This for example allows a student to create a revision playlist for all their lectures in a course.

Figure 3 shows the print friendly selection button option, the next or previous caption selection button to help speed up editing and the button to add bookmarks with notes and tags.

Figure 4 shows the print friendly low bandwidth mobile friendly option which replays only the audio with time synchronised video images, transcript and bookmarks with notes and tags which can be selected to copy to the clipboard for printing or pasting into a word processor. A QR code is shown under each image and when you print everything out you can look at all the notes anywhere and if you want to listen back to something or watch the video you can use your mobile phone QR code reader to scan the QR code and Synote will go to that precise point in the recording and play that video and the audio back on your phone.



Figure 2 Synote screen capture showing some features of video replay and caption editing



 Figure 3 Synote screen capture showing some more features of video replay and caption editing



Figure 4 Synote screen capture showing Print Friendly option

While speech recognition, caption editing and annotation may be available in some other systems, they do not offer all the above benefits and features specifically designed for disabled students. While small scale trials have been undertaken using collaborative editing conclusive evidence awaits future larger scale research trials.

Learning from a lecture recording without annotations and captions is rather like trying to learn from a text book that has not got any contents, index, page numbers, chapter or section headings, and does not allow you to add annotation, notes or bookmarks: which is not like a useful textbook but more like a story book. Similarly a lecture recording with no captions, transcript, chapter or section headings or annotation, notes or bookmarks doesn’t allows you to search and interact with the recording and so would appear to encourage students go into ‘movie mode’ wanting to be entertained along with coca cola and popcorn!

Flexible ways and benefits of taking notes with a speech recognition captioning/transcription system such as Synote that allows collaborative editing and annotation include:

# No need to write down what is said during the live lecture because you know that all the information will be available.

# Search transcript and pause and rewind recording when replaying the recording.

# Make brief personal digital notes on a mobile device during the live lecture and copy into Synote.

# Make personal digital notes on Synote when replaying the recording

# Copy the digital transcript, slides, notes into a word processor

# Print and paste/staple digital transcript, slides, notes into Synote print out

# Flexible paper notes which supports diagrams can be pasted or stapled into the Synote print out which can be edited on paper and/or scanned and pasted into Synote

# The recording can be replayed using the Synote print friendly QR time stamped codes and listened to or watched on a mobile device

# **Conclusion**

Speech recognition captioning with collaborative editing could provide affordable transcription and captioning of lecture recordings and so support inclusive learning and help universities comply with equality legislation while also having the potential to improve retention & recruitment. *References*

Boyle, J. R. (2012) Note-Taking and Secondary Students with Learning Disabilities: Challenges and Solutions. Learning Disabilities Research & Practice, 27(2), 90–101.

Burgstahler, S. (2015). Opening Doors or Slamming Them Shut? Online Learning Practices and Students with Disabilities. Social Inclusion, 3(6).

*Catchbox* (2017) [online] Available at: http://getcatchbox.com/ [Accessed 18 Oct. 2017]

*Crowdmics* (2017) [online] Available at: http://crowdmics.com/ [Accessed 18 Oct. 2017]

*Disability Rights UK Factsheet F56* (2017) [online] Available at: https://www.disabilityrightsuk.org/understanding-equality-act-information-disabled-students [Accessed 18 Oct. 2017]

Hanafin, J., Shevlin, M., Kenny. M., & McNeela, E. (2017) Including young people with disabilities: Assessment challenges in higher education. Higher Education, 54, 435– 48. (2007)

James, A., Draffan, E.A., Wald, M., (2016) Learning through videos: are disabled students using good note-taking strategies? Miesenberger, Klaus, Buhler, Christian and Penaz, Petr (eds.) In Computers Helping People with Special Needs; 15th International Conference, ICCHP 2016 Linz, Austria, July 13–15, 2016 Proceedings, Part I. Springer. 580 pp, pp. 461-467.

Johnson, J. (2015). *Disabled Students' Allowances:Written statement - HCWS347* [online] Available at: http://www.parliament.uk/business/publications/written-questions-answers-statements/written-statement/Commons/2015-12-02/HCWS347 [Accessed 18 Oct. 2017]

Layer, G. (2017) Inclusive Teaching and Learning in Higher Education as a route to Excellence [online] Available at: https://www.gov.uk/government/uploads/system/uploads/attachment\_data/file/587221/Inclusive\_Teaching\_and\_Learning\_in\_Higher\_Education\_as\_a\_route\_to-excellence.pdf [Accessed 18 Oct. 2017]

Lewin, T. (2015) Harvard and M.I.T. Are Sued Over Lack of Closed Captions [online] Available at: https://www.nytimes.com/2015/02/13/education/harvard-and-mit-sued-over-failing-to-caption-online-courses.html [Accessed 18 Oct. 2017]

*National Association of the Deaf v. Netflix* (2012) [online] Available at: https://dredf.org/captioning/ [Accessed 18 Oct. 2017]

Piolat, A., Olive, T., & Kellogg, R. T. (2005) Cognitive effort during note taking. Applied Cognitive Psychology, 19(3), 291-312.

*Supporting disabled students* (2017) [online] Available at: http://www.hefce.ac.uk/sas/disabled/inclusive/ [Accessed 18 Oct. 2017]

*Synote* (2017) [online] Available at: <http://synote.com/> [Accessed 18 Oct. 2017]

*TED's Open Translation Project* (2009) [online] Available at: http://blog.ted.com/ted\_open\_trans/ [Accessed 18 Oct. 2017]

Wald, M (2017) Business Models for Captioning University Lecture Recordings. Journal of Management Science Surrathani Rajabhat University

Wald, M (2013) Concurrent Collaborative Captioning. Proceedings of SERP'13 - The 2013 International Conference on Software Engineering Research and Practice

Wald, M (2012) Important new enhancements to inclusive learning using recorded lectures At 13th International Conference on Computers Helping People with Special Needs (ICCHP 2012), Austria. 11 - 13 Jul 2012. 8 pp.

Whitney, L. (2011) Netflix sued by deaf group over lack of subtitles [online] Available at: https://www.cnet.com/uk/news/netflix-sued-by-deaf-group-over-lack-of-subtitles/ [Accessed 18 Oct. 2017]

Xiong, W., Droppo, J., Huang, X., Seide, F., Seltzer, M., Stolcke, A., Yu, D., Zweig, G. (2017) Achieving Human Parity in Conversational Speech Recognition: Microsoft Research Technical Report MSR-TR-2016-71 February 2017