

Transcribe Your Class: Using Speech Recognition to Improve Access for At-Risk Students

Keith Bain
Saint Mary's University

Eunice Lund-Lucas
Trent University

Janice Stevens
Saint Mary's University

Through a project supported by Canada's Social Development Partnerships Program, a team of leading National Disability Organizations, universities, and industry partners are piloting a prototype Hosted Transcription Service that uses speech recognition to automatically create multimedia transcripts that can be used by students for study purposes. This paper focuses on the accessibility and educational impacts of receiving complete and accurate lecture transcripts for students with disabilities.

Accessibility Challenges Facing Students

The post-secondary educational environment can be extremely challenging for students with disabilities, international students, and other at risk populations. A key challenge facing these stakeholders is access to lecture content and the requirement for adequate notetaking skills.

In post-secondary education, research demonstrates that studying lecture notes results

in a greater learning experience and higher grades (Peverly et al., 2007; Titsworth & Kiewra, 2004). While notetaking is clearly connected to academic achievement (Kiewra, 1985a; Leitch & MacMillan, 2002), studies show that students with disabilities record up to 70% less lecture information than students without disabilities (Hughes & Suritsky, 1993).

Manual notetaking can be extraordinarily difficult for students with disabilities. Learners with physical disabilities may lack the fine motor skills necessary to take copious notes throughout a class

period. Students with learning disabilities often experience difficulties capturing lecture notes and may benefit from audio recordings or instructor-generated notes (Hughes & Suritsky, 1993). Hearing impaired students utilizing visual language interpreting typically require notetaking support, as simultaneously attending to an interpreter, the instructor, and the notetaking task is nearly impossible.

Notetaking support services

Notetaking is one of the most requested support services for students with disabilities (Elliot, Foster, & Stinson, 2002). Services can be categorized by the notetaking methodology used and the extent and form of resulting notes.

Various intermediary notetaking models are used in Canadian post-secondary institutions including peer notetaking, computerized notetaking, and remote stenography. Volunteer, peer notetakers can be helpful, but students with disabilities have little control over the content of lecture notes being recorded and are dependent on the skills of the notetaker (Elliot et al., 2002). Professional services, such as stenography, Communication Access Realtime Translation (CART), and remote captioning are usually more effective at capturing content, but are expensive and sometimes in short supply.

The format of notes is also varied and ranges on a continuum from traditional shorthand outlines, summarizations, to verbatim transcripts. Because less than 40% of information presented during lectures is typically captured by traditional notetakers (Titsworth, 2004) and summarization techniques using computerized notetakers are subject to interpretation errors, verbatim transcription provides the most complete lecture resource. Transcripts have been used to support diverse learners in college courses through increased content access and engagement (Leitch, 2002; Rose, 2006; Wald, 2006). However, creating transcripts is traditionally expensive, relies on professional intermediaries, and limits benefits to those with disabilities.

Speech recognition: A new support

Speech recognition technology can be used to convert

a recorded lecture into a searchable transcript for review and study purposes (Bain, Basson, & Wald, 2002). Case studies have demonstrated that post-secondary students with disabilities found speech recognition generated transcripts of academic lectures useful and enhanced access to lecture material (Leitch & MacMillan, 2002). Students without disabilities also noted that transcripts, “helped fill the gaps” in their notes, facilitated revision of notes in preparation for exams, and improved understanding of the lecture material (Hede, 2002; Leitch & MacMillan, 2002). Creating accessible, digital archives of key information allows all users to retrieve accessible information on demand in the format that best suits their individual learning needs.

Longstanding challenges facing post-secondary students with disabilities regarding notetaking and evidence that speech recognition based transcription can improve accessibility for at risk students prompted a group of researchers to launch a new initiative targeting Canadian youth.

Project Description

Through a project supported by Canada’s Social Development Partnerships Program, a team of leading National Disability Organizations, universities, and industry partners ran an applied research project to explore the perceptions of students with disabilities towards a Hosted Transcription Service and SR generated multimedia transcripts.

Hosted Transcription Service (HTS)

Hosted Transcription Service (HTS) is a prototype, speech recognition powered platform that performs speaker-independent, offline transcription of audio and video files. To use HTS, authenticated users visit an online portal, log into secure user accounts, and upload a media file for automatic transcription.

Multimedia transcripts

While traditional transcripts are typically generated by listening and manually typing what is heard, multimedia transcripts refer to machine-generated

text that is synchronized with the original spoken language source. Users access multimedia transcripts through a conventional web browser in various accessible formats. These learning resources facilitate multimodal access to content and allow students to review, search for key words/phrases, and interact with these new learning resources according to personal preferences.

The primary research question during this project was how students with disabilities responded to and described their experiences using HTS and multimedia transcripts. A second key area of the project investigated pedagogical benefits for faculty participants. Given the project was designed to both iteratively improve the technology and simultaneously evaluate its impact on students, an applied research model with primarily qualitative methods was selected as the mode of inquiry. Research objectives included:

- determining whether students found HTS easy to use;
- determining how students used HTS and multimedia transcripts;
- identifying difficulties encountered when using HTS and accessing multimedia transcripts; and
- articulating faculty perceptions of HTS and/or multimedia transcripts.

Methodology

Students with disabilities were recruited from Canadian post-secondary institutions to participate in the study. Applicants completed an online form that required prospective participants to self-disclose the presence of a disability. An initial cohort was selected for pilot testing in September 2010. They were required to complete a participation agreement that included their instructor's signature, which granted permission to record lectures and submit them for transcription.

Before using HTS, participants completed a brief, online pre-technology survey that attempted to capture current perceptions of SR technology, current

services being used to access lecture content (i.e., note taking services) and to identify any logistical barriers that may impede participation.

Participants completing the pre-technology survey were given user accounts for HTS. Students subsequently recorded lectures and uploaded the media for transcription. Once the system processed uploaded jobs, an automated email notification was sent to the student indicating that the job was complete and a raw, unedited version of the multimedia transcript was available. While speech recognition systems use advanced statistical models to improve recognition, they still contain recognition errors. Project organizers therefore passed the raw transcript to a human editor who corrected errors. The edited version subsequently replaced the raw multimedia transcript available through the student's account.

At the end of the academic semester, students who submitted recordings for transcription were given a post-technology survey to assess their experiences with HTS and multimedia transcripts.

Preliminary Results

During this first phase, 69 students applied to participate, 33 received their instructor's permission, and 17 ultimately submitted recordings for transcription. A total of 65 lectures were made available to students as multimedia transcripts.

The results of this initial phase were encouraging as the majority of students that completed post technology surveys indicated the technology improved their understanding of lecture content and increased independence compared to traditional support services. One student expressed her feelings about the system and how it provided a new level of information:

"I have reviewed the transcript and am thrilled! I am just beginning to realize the magnitude of how much I am actually missing! She spoke very clearly and I 'thought' I was understanding her primarily because

I was very familiar with the content. However, reviewing the transcript, I began to understand her presentation at an entirely different level. Thank you!”

Some students voiced concerns with the system: “I was very frustrated and discouraged when I received the unedited transcripts...too much lecture not enough notes...and way too long.” These comments spoke to the perceived difference between traditional hand written notes and the availability of transcripts. The student in this case was one of the few that utilized unedited versions of the speech recognition generated text. Most participants based their experiences on the final outputs that did not include recognition errors.

Although it was not originally part of the evaluation, researchers were curious why so many students signed up to participate, yet did not follow through.

Through discussions with many service providers, a number of factors emerged. Finding a method to properly record lectures was an unforeseen obstacle. Many students tried to use digital recorders, simply placing them on or near the instructor’s desk. Recordings captured through this method were very low quality and contained too much ambient noise. Audio quality is a key variable in using speech recognition. Initially, technical specifications mandated that students record audio in .wav format, which was not feasible with some devices students tried to use. Other reports suggested that some of the participation requirements, such as a five hour transcription limit per applicant, were cited as disincentives.

Some instructors did not grant permission, voicing concerns about having their lectures recorded and transcribed. Despite the presence of participation agreements that expressly prohibited unauthorized reproduction and dissemination of recordings and resulting transcripts, a small number of instructors felt uncomfortable with having their lectures recorded. Others worried that peers or superiors could potentially use the resulting transcripts to monitor or evaluate teaching methods and in class performance.

Two instructors noted that given HTS was initially hosted on servers residing in the US, it was subject to the US Patriot Act that allowed authorities access to the records of internet service providers. In one case, the instructor felt the lapel microphone was too inconvenient and uncomfortable to wear.

The above objections were the exception rather than the norm, as most instructors welcomed the opportunity to assist students with supports that could improve learning opportunities.

Future Activities

A second phase of testing began in September 2011. Based on preliminary findings, a number of key changes were instituted. Researchers developed a series of videos to explain key aspects of the project and increase awareness of this opportunity. The application process was streamlined and organizers removed a five hour transcription limit and opened the timeline to cover the entire course period. Developers also changed HTS to allow a wider variety of multimedia formats, including high-quality mp3s. As an added incentive, a \$50 stipend was offered to those students that successfully participated in the project.

A complete evaluation and publication of final research results occurred in 2012. During the second phase, researchers also collected feedback from participating academics/instructors, with the goal focusing on understanding pedagogical implications associated with these technologies.

Behind the scenes, a technical team conducted statistical modeling using the transcribed lectures to improve recognition performance. A new stream of customized acoustic and language models designed for the lecture environment were developed. These scientific endeavors may ultimately lead to more robust technologies, which should improve the overall ease-of-use of the systems; encourage increased adoption; and ultimately lead to increased access to information for at risk students. Complete project information, including demonstrations of HTS technology/multimedia transcripts and details about research outcomes, are available at www.transcribeyourclass.ca.

Beyond this project lifecycle, future lines of inquiry could explore the benefits and experiences of other students, including international students and second language learners. Using quantitative methods to investigate whether the presence of HTS and multimedia transcripts improve academic performance could generate further evidence supporting the expanded use and adoption of these technologies in higher education.

References

- Bain, K., Basson, S., & Wald, M. (2002). *Speech recognition in university classrooms: Liberated learning project*. Proceedings of the Fifth International ACM SIGCAPH Conference on Assistive Technologies, ACM Press, 192-196.
- Elliot, L., Foster, S., & Stinson, M. (2002). Student study habits using notes from a speech-to-text support service. *Exceptional Children*, 69(1), 25-40.
- Elliot, L.B., Stinson, M.S., McKee, B.G., Everhart, V.S., & Francis, P.J. (2001). College students' perceptions of the c-print speech-to-text transcription system. *Journal of Deaf Studies and Deaf Education*, 6(4), 285-298
- Hede, A. (2002). Student reaction to speech recognition technology in lectures. In S. McNamara & E. Stacey (Eds.), *Untangling the web: Establishing learning links*. Proceedings ASET Conference, Melbourne.
- Hughes, C.A. & Suritsky, S.K. (1993). Notetaking skills and strategies for students with learning disabilities. *Preventing School Failure*, 38(1).
- Kiewra, K.A. (1985a). Students' note-taking behaviors and the efficacy of providing the instructor's notes for review. *Contemporary Educational Psychology*, 10.
- Kiewra, K.A. (1985b). Investigating notetaking and review: A depth of processing alternative. *Educational Psychologist*, 20, 23-32.
- Kiewra, K.A., DuBois, N.F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. *Journal of Educational Psychology*, 83, 240-245.
- Leitch, D. & MacMillan, T. (2002) How students with disabilities respond to speech recognition technology in the university classroom. *Final Research Report on the Liberated Learning Project*, 10-22.
- Peeverly, S.T., Ramaswamy, V., Brown, C., Sumowski, J., Alidoost, M., & Garner, J. (2007) What predicts skill in lecture note taking? *Journal of Educational Psychology*, 99(1), 167-180.
- Rose, D.H. (2006). Universal Design for learning in postsecondary education: Reflections on principles and their application. *Journal of Postsecondary Education and Disability*, 19(2).
- Titsworth, B.S. & Kiewra, K.A. (2004). Spoken organizational lecture cues and student note-taking as facilitators of student learning. *Contemporary Educational Psychology*, 29, 447-461.
- Wald, M. (2006) Creating accessible educational multimedia through editing automatic speech recognition captioning in real time. *International Journal of Interactive Technology and Smart Education: Smarter Use of Technology in Education*, 3(2), 131-142.

Biographies

Keith Bain is the lead of the international Liberated Learning Consortium and Adjunct Professor in the Department of Finance, Management Science, and Information Systems at the Sobey School of Business. Keith helped co-found an international

research group dedicated to improving accessibility through speech recognition based captioning and transcription systems. He has presented keynote addresses and special sessions on speech recognition and accessibility in over ten countries, including the renowned CSUN Technology and Persons with Disabilities Conference, Australian Pathways National Conference, and the China Information Accessibility Fo-rum. In addition to invited lecturers at prestigious universities such as MIT and the Indian Institute of Technology, his work was honoured by a National Education Award from the Learning Disabilities Association of Canada and an International Faculty Award for Innovation from IBM Research.

Eunice Lund-Lucas has been the Manager of the Disability Services Office at Trent University for 22 years. Eunice has been a member of several provincial committees looking at support for students with disabilities at the post-secondary level and served on the executive of the Inter-university Disability Issues Association and as President a national organization of service professionals in post-secondary education in Canada. Currently, Eunice represents Trent's involvement in the Liberated Learning Consortium, an international research network to introduce speech recognition in the learning environment.

Janice Stevens served as the Project Coordinator for the Liberate Learning Youth Initiative. Janice Stevens utilized here experience in the area of speech recognition and learning technologies to educate individuals and organizations on the benefits of accessible educational media. Janice has a Computer Programming diploma from CDI College and a diploma in Business Information Technology from NSCC.