The Liberated Learning Project:[1] Improving Access for Persons with Disabilities in Higher Education Using Automated Speech Recognition Technology

David S. Coco
The University of Texas at Austin

Jane Bagnall
Saint Mary's University, Halifax, Nova Scotia, Canada

Abstract

The Liberated Learning Project, an international research effort spearheaded by the Atlantic Centre at Saint Mary's University, Nova Scotia, Canada promotes the use of automated speech recognition (ASR or speech-to-text) in the university classroom as a tool to improve access to lecture material for students with disabilities. This paper discusses some of the different techniques that have been proposed for ASR in the classroom, reviews several recent investigations related to real-time ASR, and provides an overview of the Liberated Learning Project's role in developing this unique application of speech-to-text technology. ASR has positive implications not only for students and professors, but will also undoubtedly have a significant impact on a much broader scale.

Introduction

Accessibility to higher education for persons with disabilities has been a cornerstone of the philosophy of Saint Mary's University in Halifax, Nova Scotia, Canada for over 30 years. In 1980, the University's first deaf student was accommodated and in 1985 the Atlantic Centre of Research, Access and Support for Students with Disabilities began a formal program of support services to students with a wide range of disabilities. The Centre has since then developed an international reputation as a leader in providing support services for students with disabilities at the post-secondary level. Its programs and individualized support services include individual academic counseling, ASL/English interpreting, academic support, a volunteer note-taking program, alternative examination accommodations, adaptive technology and training in its use.

The introduction of true continuous speech recognition (speech-to-text) products with large, expandable vocabularies engendered a commitment from Saint Mary's University to explore the concept further in connection with increasing accessibility to lecture material for students with disabilities. It was clear that problems existed with both immediate intake of the material and with notetaking for later study purposes. For example, students who were deaf or hard-of-hearing usually required interpreters or assistive listening devices, and relied upon notetakers. As well, students with certain learning disabilities found it difficult to process information presented orally, and other students were physically unable to take their own notes. Finally, the notetaking skills of non-disabled students *were* often far from satisfactory. These shortcomings with both the teaching and the learning processes became the impetus for investigating the use of ASR in the classroom.

Techniques and Recent Investigations

In this section we discuss some of the different techniques that have been proposed for ASR in the classroom and also review several recent investigations related to real-time ASR. We should note

[1]A Project of the Atlantic Centre of Support for Students with Disabilities, Saint Mary's University, Halifax, Nova Scotia, Canada.

that these techniques and investigations are very closely tied to the introduction of several consumer oriented ASR software packages in 1997. These software packages (e.g. IBM Via Voice and Dragon Systems NaturallySpeaking) were the first products featuring large vocabularies and continuous speech recognition made available to the general public. For the first time researchers had easy access to the tools needed to investigate the use of ASR for classroom lectures and other real-time applications.

Three different ASR real-time techniques have been proposed by various groups. In the direct input technique the instructor speaks directly into the ASR system microphone, and the text is displayed in real-time and also saved for post-lecture notes. This technique requires that the instructor be trained in the use of the ASR system and feel comfortable using a wireless microphone during the lecture. The verbatim shadowing technique uses an intermediary (referred to as a 'shadower') to repeat the lecture verbatim into the ASR system. The shadower uses a mask to dampen the speech sounds so that the other people in the room are not disturbed. Because current ASR systems are not always capable of performing adequately when the speech rate is over about 150 wpm, it has been suggested by Ross Stuckless of NTID that an abbreviated shadowing technique may be more useful. In this technique, the shadower would repeat only the essential information in the lecture, producing readable and accurate text. This technique resembles the NTID C-Print technique in many ways.

Several groups have begun to investigate these techniques but overall the field of ASR for deaf applications is still in the very early stages of development. There has only been one conference which has focused on this topic in depth: the Frank W. Lovejoy Symposium on Applications of Automated Speech Recognition for Deaf and Hard of Hearing People held in Rochester, New York in April 1997. This symposium explored the potential of using ASR for automatically converting speech to text for deaf people in a variety of settings, including the classroom environment. This symposium was held before researchers had an opportunity to investigate the ASR software products mentioned above so it does not include results from the more recent software. An electronic copy of the proceedings of this conference is available at www.isc.rit.edu/~ewcncp/Lovejoy.html.

In the Fall of 1998 Saint Mary's University in Nova Scotia, Canada, initiated the first trial using ASR in the classroom. Three professors used ASR systems in the classroom to display the text of the lecture in real-time and also provide post-lecture notes for students. In this trial the professor spoke directly into the ASR system using a wireless microphone and the text was displayed overhead for all students to view. This study concluded that the technology was promising and warranted further investigation and development. This trial established the groundwork for the Liberated Learning Project described elsewhere in this paper.

In the Spring of 1999 Sprint, IBM, and the University of Texas collaborated in a trial at the University of Texas at Austin using the verbatim shadowing technique. The objective of this trial was to investigate the feasibility of this technique for classroom applications and identify the critical elements of the system that needed improvement. This study found that the rate of speech for normal classroom lectures (180–220 wpm or higher) was too fast for the current ASR systems and that extensive training of the shadower was critical to the success of the system. In both of Saint Mary's and UT trials the ASR system was used to augment rather than replace the usual support systems for deaf students. Ross Stuckless at NTID also performed a series of tests in 1999 using the verbatim shadowing technique in simulated classroom conditions. He found that the error rate was unacceptably high even when the shadower was a court reporter who had been extensively trained in verbatim re-voicing of speech in the courtroom. Stuckless concluded that abbreviated shadowing might be a more feasible approach for the classroom.

In the Winter of 1999 Sprint and Ultratec initiated a six month trial of ASR in the Maryland relay service. Although this trial was not held in a classroom environment the relay ASR technique is very similar to the verbatim shadowing technique, and the results have some important implications for classroom applications. In this trial the relay agent re-voiced the spoken telephone conversation in real-time rather than typing it as is normally done. This trial used commercially available ASR software that had been modified so the agent could edit the text before it was sent out to the deaf caller. This editing feature allows the agent to correct the ASR errors so that the agent has more control over the error rate. Although an official report on this trial has not been released to the public, several beta testers (including the author D. Coco) found that ASR did not provide a significant improvement in the speed of transmission compared to regular relay. In addition, the error rate for the ASR system varied widely from excellent to unacceptable depending on type of conversation and the skills of the agents.

Although ASR has an enormous potential for automatically converting speech to text for deaf people, all of these initial trials have demonstrated that simply asking an untrained speaker to use an off-the-shelf commercial ASR product does not produce acceptable results. These trials have indicated that speaker training is probably the most critical issue in the implementation of this technology. This is not unexpected because the ASR software used in these trials was not developed specifically for real-time applications but rather for dictation applications. This is an important point because there are significant differences in these two applications.

An ASR system designed for dictation can set limits on the rate of speech and on the type of speech that is allowed to provide optimal system performance. The dictation user can usually adjust the rate and content of his speech to meet these requirements. In a real-time application, however, it is usually rather difficult for the user to limit the rate of speech or specify the content of the speech. In addition, the effect of the ASR errors on dictation and real-time users is quite different. The dictation user knows exactly what the output from the ASR program should be (after all, he is doing the talking!), so errors are easily detected and corrected. However, the deaf real-time user has only a vague idea of what the output should be (in a math class the text is expected to be math related) so errors will be more difficult to detect and correct, especially in real-time.

Even if the error rates were the same for the two users, which is highly unlikely, the *effect* of the errors on the real-time user would be much more significant. The dictation user can simply correct the error and move on, whereas the real-time user may completely misunderstand the meaning of the sentence or paragraph. Simply looking at word count error rates may not be the best approach for evaluating real-time ASR systems. An evaluation of the comprehension of the user may be required.

Can speakers be trained to use ASR dictation products effectively for real-time deaf related applications? This is the key question to be addressed. One proposed approach is to train the instructors themselves (the direct input technique), whereas another approach is to train intermediary speakers who might modify or summarize the original speech (verbatim or abbreviated shadowing). No conclusive answers have been provided as to whether any of these approaches will actually work. However, the potential for using ASR to improve access for deaf students in the classroom and in other settings is tremendous, and further exploration in this area is certainly warranted.

The Liberated Learning Project

As mentioned earlier, the Atlantic Centre's pilot project in 1998 found the initial testing of this application for ASR to be enlightening. Brief exposure to the concept suggested it could indeed provide an alternative to conventional note taking for students with disabilities. Serendipitously, it was also noticed that non-disabled students were using the instantaneous display of the lecture as a reference check for their own notes: ASR technology gave students access to both auditory and visual learning channels, helping them better integrate the lecture content. They could also use the software-generated notes to augment their own notes. Therefore, the successful application of ASR technology was seen to have valuable implications for every student in the classroom.

Saint Mary's University received major funding in 1999 from a Canadian foundation, The J. W. McConnell Family Foundation, to further research and refine the unique application of ASR technology to assist students with disabilities in the university classroom. Saint Mary's University is now heading a consortium of Canadian and international partners, both universities and industries. These strategic alliances will help develop, test and evaluate multiple applications of ASR technology in the classroom and its implications for pedagogy and learning.

Project Concept

- Professor develops a personalized voice profile by "teaching" speech recognition software to understand his/her speaking style.

- Professor uses a wireless microphone 'connected' to a robust computer system during lectures. A computer running speech recognition software (project is using IBM's ViaVoice products) receives digitized transmission of professor's speech.
- Using professor's voice profile and acoustic information, the software converts spoken lecture into electronic text.
- Text is displayed via projector for class in real time: students can simultaneously see and hear lecture as it is delivered.
- After lecture, text is edited for recognition errors and made available as lecture notes (electronic or hard copy format) for all students.
- Professor's individual voice profile is continuously updated and expanded through intensive system training.

Project Objectives

The main objectives of the project are to develop and evaluate a model for using automated speech recognition in the university classroom and to focus global attention on this concept as a method of improving access to learning for persons with disabilities. During the three-year period, the project will thoroughly develop and test multiple applications of speech recognition in university classrooms. Global discussions of speech recognition as a tool to enhance teaching and learning will be stimulated. An effective model for using speech recognition in the university classroom will be developed and refined. Finally, an international conference on the importance of speech recognition in the university classroom will be sponsored.

Project Partnerships

Saint Mary's University has recruited several implementation and research partners that are essential to our success, including IBM, Maritime Tel and Tel, and individuals from universities in Canada and around the world, including England, Australia, and the U.S. This team will collaboratively forge the project's development, from the initial planning stages through to the in-class trials and beyond. The project has a mandate to pursue further partnerships and interested parties are invited to make contact regarding potential involvement in this international research consortium.

Partner universities will share a philosophical commitment to addressing issues of accessibility, inclusion of qualified persons with disabilities in academic programming, and providing support services to students with disabilities. They will designate an individual to lead the initiative on site and represent the university at the project level. Partners will attempt to provide resources to enhance the project's overall mission and assist in achieving objectives. Partner universities will be dedicated to implementing the Liberated Learning concept in university classrooms and will share a commitment at a research, technical, or consultative level.

Project Challenges

The Liberated Learning Project involves an intricate interaction of technological and human resources. As with any technological application in its infancy, there are obstacles to overcome before the Liberated Learning concept is more universally applicable. A few of the more pressing project challenges are:

- Improving recognition accuracy. As a professor delivers a lecture, the displayed text must be accurate and convey the intended message.
- Reducing the occurrence of errors. Errors affect the overall conceptual understanding of the lecture and thus remain our primary focus.
- Integrating non-obtrusive punctuation markers. Currently, speech recognition software requires the speaker to actually say the marker in order to have it appear (i.e., speaker says "period" or "new paragraph"). One challenge is to find a nonobtrusive way of integrating these markers to enhance readability and thus comprehension.
- Developing a model capable of effecting better learning and teaching. Professors must be able to learn the software quickly and use it easily. The project will be looking specifically at

the efficiencies of editing a lecture transcript produced via speech recognition software.

- Determining the right mix of associative technologies: sound card, operating system, microphone technology, memory, storage, etc.
- Customizing IBM's ViaVoice speech engine for lecture use.

Conclusion

It was 120 years ago that Alexander Graham Bell, who had strong ties to Nova Scotia, began experimenting with voice recognition to help deaf persons. It seems fitting that a Nova Scotia university is assuming a major role in advancing speech recognition to help persons with disabilities in the classroom. Speech recognition technology may potentially revolutionize the way students and professors interact in a university environment. It has the potential to spark exciting and unprecedented outcomes for both students and faculty. It is the hope of project participants that this innovative concept will be a stepping stone in developing and nurturing an educational environment free of boundaries, where all students have equal opportunity to pursue their educational aspirations.