Return to Menu

SPEECH RECOGNITION – THE FUTURE IS OUT THERE

LIBERATED LEARNING PROJECT

An innovation to Improve Access to Higher Education
Using Speech Recognition Technology

INTRODUCTION

The LIBERATED LEARNING PROJECT (LLP) is a large and complex research project studying two core questions:

1) Can speech recognition technology be successfully implemented for real-time transcription and display of text in university lecture theatres?
2) Can speech recognition technology be used successfully as an alternative to traditional classroom notetaking for persons with disabilities?

This paper is intended to inform discussion on the nature and extent of research required to develop and implement the Liberated Learning Project.

The Liberated Learning Project evolved from a pilot project conducted by the Atlantic Centre of Research, Access, and Support for Students with Disabilities, Saint Mary's University, Halifax Nova Scotia in 1998. After voice training using computers, three Saint Mary's University faculty members wearing cordless microphones utilised speech recognition in their lecture theatres. Their lectures were digitized and simultaneously displayed on a large screen at the front of the lecture theatre. This meant, for the first time in a mainstream university, students had access to the spoken lecture in its entirety. The students could see the lecture as it took place. They could also obtain a hard copy or disk copy of the lecture transcript for later use. Saint Mary's University learnt much from this brief exposure to speech recognition as an alternative to traditional lecture theatre note taking and could foresee great promise for this technology, not only for students with disabilities, but for faculty participants and the general student population. It was obvious that the implications of voice recognition technology for teaching and learning were great.

Liberated Learning concept at a glance:

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

- Lecturer uses a wireless microphone 'connected' to a robust computer system during lectures. An IBM Intellistation computer running specially designed speech recognition software, working in conjunction with IBM's ViaVoice technology, receives digitized transmission of lecturer's speech.
- Using lecturer'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** the lecture as it is delivered.
- After the lecture, text is edited for recognition errors and made available as lecture notes (electronic or hard copy format) for all students through an on-line notes system.
- Lecturer's individual voice profile is continuously updated and expanded through intensive system training.

Saint Mary's University received major funding in 1999 from the J.W. McConnell Family Foundation in Canada to further research and refine a unique application of speech recognition technology to assist students with disabilities in the university lecture theatre. Saint Mary's University through Dr David Leitch, Director of the Atlantic Centre, is now heading a consortium of Canadian and international university and corporate partners to develop the Liberated Learning concept. These strategic alliances will collaborate on the development and testing of speech recognition technology in the lecture theatre, and study its implications for pedagogy and learning.

Universities involved in developing and testing the concepts during the project initially include Saint Mary's University and Ryerson Polytechnic University in Canada and the University of the Sunshine Coast in Queensland Australia. Additionally, the project has a specific mandate to expand the consortium over its three-year life and numerous international universities are poised to join the project team.

Saint Mary's University will begin testing the software in lecture theatres during October 2000, followed by the University of the Sunshine Coast and Ryerson Polytechnic Universities in February 2001.

PROJECT OBJECTIVES

The main objectives of the Liberated Learning Project are to develop and evaluate a model for using speech recognition in the university lecture theatre. Further the project intends to focus global attention on this concept as a method of improving access to learning for people with disabilities. During this three-year project, researchers will thoroughly develop and test multiple applications of speech recognition as a tool to enhance teaching and learning. The project culminates with an international conference on the importance of speech recognition in the university lecture theatre.

To illustrate the potential impact of this teaching and learning tool on students with disabilities, a demographic study of students with a disability, undertaken in Canada in 1995, revealed approximately 7,000 students with a disability

were attending one of the 47 universities surveyed by McLean's Magazine. In Australia, according to the 1999 statistics produced by the Department of Education, Training and Youth Affairs (DETYA), there were 18,084 students with a disability enrolled at the 39 public universities. Therefore, the immediate implications for speech recognition technology in tertiary education in Australia and elsewhere will be great.

SUPPORT FROM THE CORPORATE SECTOR

As well as university partners, the Liberated Learning Project has also corporate partners, namely IBM-Research and MTT (formerly Maritime Telephone and Telegraph Company), Canada.

The Project has been assisted by the world's top speech recognition scientists at the IBM - T.J Watson Research Centre in Yorktown Heights, New York. Assistance from IBM Research includes:

1. Providing access to the world's top speech recognition scientists;
2. Providing access to IBM's top technical/software experts;
3. Providing access to the software codes for ViaVoice to enable modifications for use in lecture theatres;
4. Copies of the latest IBM speech recognition software;
5. Serving in an advisory role to lecturers as they are trained in the use of ViaVoice technology;
6. Providing a forum for team members to get updates about advances in speech technology;
7. Working with partners in the Project to design experiments evaluating the efficacy of the technology;
8. Serve as a liaison with IBM Special Needs Division to determine whether the results of the Project can be disseminated more widely in other university settings.

MTT, Canada are also partners in the Project as they see unlimited opportunities in the advancement of this technology for the purpose of enhancing the global learning environment. MTT provides the sophisticated telecommunications support structure necessary for the collaboration of individuals and institutions in Canada, the United States, Britain and Australia.

PROJECT CHALLENGES

There are a host of questions generating from the initial conceptualisation phase, which will be addressed through the life of this Project. For example, an obvious question is the effectiveness of speech recognition for real-time transcription and display of university lectures. If speech recognition does work then information needs to be gathered on the impact on learning and pedagogy. Furthermore, by the end of the Project the outcome needs to be able to go beyond the subjective and declare with confidence the value and utility of speech recognition for the lecture theatre, firstly for students with disabilities and then for non-disabled students and teachers.

The first attempt with speech recognition in the classroom in 1998 was under financed and as a result Saint Mary's were not able to provide much in the way of documentation on the actual tests. Consequently, the project must start from scratch in terms of producing baseline data on which later comparison studies and evaluation can be undertaken. At the project's inaugural 1999 meeting, Dr. Ross Stuckless, a consultant and team member from Rochester Institute of Technology, reinforced the importance of documentation for the purposes of establishing baseline data. He recommended remedying this situation by documenting an examination of a variety of variables tried in mini-demonstrations prior to formal classroom testing. However, the baseline data which Professor Stuckless suggested for collection, could only occur when the actual configuration of speech recognition hardware/software to be taken into the lecture theatre was determined.

In other words, samples of voice and displayed text cannot be produced until such time as a computer, microphone, sound card and software modification of Via Voice Millennium (VVM) was settled on. This baseline system is now operational in Saint Mary's University classrooms.

During the conceptualisation or first phase of the project, a working group was constituted including a programmer and the scientists at IBM, to decide on the speech recognition technology that faculty plan to test in their lecture theatres. It is only after many hundreds of exchanges amongst the team and partners that current state of readiness has been reached. The intensity and thoroughness of analyzing equipment such as the microphone and sound card has been impressive to say the least. Equally impressive, the Project Manager, Keith Bain, with the support of MTT, has ensured that an electronic archival strategy was in place to chronicle the exchanges and developments to date with respect to the technology. This represents a significant change from the pilot study.

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 lecturer 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 non-obtrusive way of integrating these markers to enhance readability and thus comprehension.
- Developing a model capable of effecting better learning and teaching. Lecturers 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.
- Ensuring as many elements of the research, as is possible, are standardised to ensure universal comparison.
- Improving the technical model's cost efficiencies - an important task when considering more universal application of the Liberated Learning concept.

CONCEPTUAL FRAMEWORK

Thus far, much of the energy has been focussed on technology. However, in anticipation of studying the impact of speech recognition on students and lecturers we now turn the attention to the collection of baseline data specific to this domain.

The conceptual framework comprises six cells representing all of the components of this project that we are investigating.

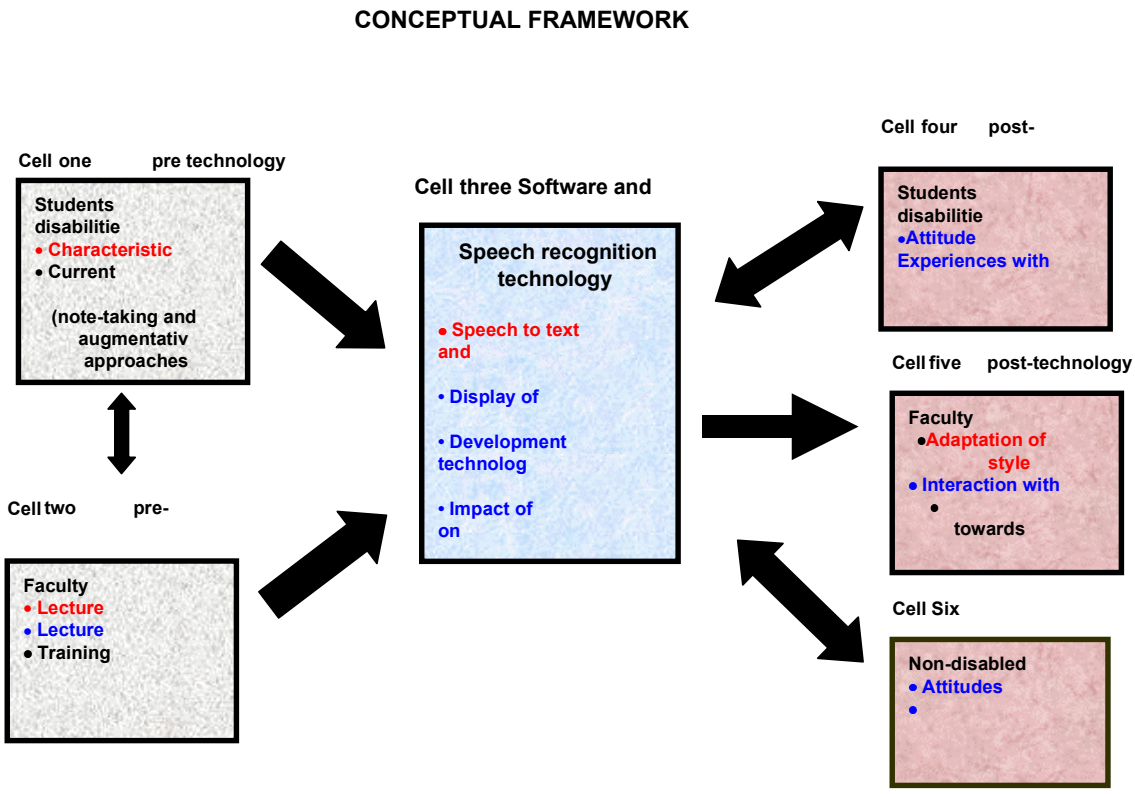

Figure 1. Conceptual Framework.

Each cell is accompanied by a list of relevant research questions. Figure 1. Conceptual Framework details the inter-relationship between the various aspects of the Project.

CELL ONE (Students with disabilities) is elaborated upon with a list of possible research questions. It is important to note that Cell One will provide the baseline data we require on students' experiences with note-taking and other augmentative processes in the lecture theatre. Joby Fleming, President of the National Educational Association for Disabled Students (NEADS), working out of the Faculty of Education, Memorial University, Newfoundland, Canada has commenced a literature search and descriptive study on the current barriers students with learning disabilities have experienced with notetaking and with paid/unpaid notetakers in university. In light of the apparent absence of literature on the effectiveness or otherwise of note-taking and other augmentative approaches in the lecture theatre, this baseline data will be effectual in allowing for comparative measurements later in the study.

An important feature in the gathering of such data on experiences of students is to allay any anxieties about the replacement of other augmentative approaches by speech recognition.

CELL TWO provides an understanding of the impact of speech recognition on teaching. This necessitates a close examination and documentation of current lecture preparation and styles, attitudes, concerns and expectations of the lecturer in the pre-technology phase as well as the training priorities of faculty.

This data will provide the foundation against comparisons of the degree of change that occurs with respect to lecture preparation and style once speech recognition has been implemented in the classroom.

Though the relationship between faculty and students with disabilities is not a central focus of this research, it nevertheless provides further insight into the experiences of students with disabilities in post secondary institutions. For example, do students with disabilities feel lecturers are aware of their specific needs and/or limitations? How do lecturers describe their own understanding of and experiences with teaching students with disabilities?

CELL THREE of the Conceptual Framework concerns itself with the hardware and software itself. This cell will examine the specific technological issues associated with the development of the software and hardware configurations and its impact on the project and test sites.

Secondly this cell examines speech to text accuracy and readability. Word accuracy is a major component in the readability of the text of a lecture, and in turn its value to the student, whether it is displayed in the classroom to be read in real time, used as lecture notes, or both.

It is common to read or to be told by a user of speech recognition, that 98% accuracy is readily achievable. And indeed this is so, under favorable

conditions such as dictating or reading selected materials aloud. However, in introducing speech recognition into the classroom, and asking it to recognize a lecturer's spoken lecture, we are asking both the technology and the instructor to undertake a much more challenging application.

What can we expect by way of word accuracy? The Liberated Learning Project has set for itself the goal of 90–95% accuracy by completion of the Project. Is this attainable? Only time and the best efforts, both technically and pedagogically, will tell.

Dr Ross Stuckless has developed an instrument for a detailed scoring procedure for inter-scorer readability (Word Accuracy sub-test of the National Technical Institute for the Deaf Test of Automated Speech Recognition Readability). Dr Stuckless's instrument is designed to test three components of text readability, i.e. word accuracy, sentence markers and speaker changes. Dr Stuckless's experience in scoring using these previously documented procedures will undertake a random sample of 2,000 consecutive spoken words from one hour lectures videotape recorded by each of the participating professors to derive a word accuracy score. To enable accurate scoring, a lecture is videotaped is then dubbed to an audiocassette. A copy of the original, unedited text generated automated speech recognition in tandem with the lecturer's voice is then analyzed and compared to derive the accuracy score.

CELL FOUR (students with disabilities, post-technology) of the Conceptual Framework is intended to focus upon understanding the ways in which students with disabilities have responded to and interacted with speech recognition technology. Ultimately, the researchers need to understand how students with disabilities are responding to this innovation, whether or not it has a negative or positive impact on the learning experience. This particular component of the overall inquiry can be considered both extremely subjective and multi-faceted in nature, as such the data will reflect experiences and attitudes.

CELL FIVE directs our focus upon the effects of speech recognition on teaching. Ultimately, a close examination and documentation of how lecturers use and report on speech recognition can illuminate understanding of the efficacy of this innovation. Similarly, the data gathered during and after the implementation of speech recognition will be compared to data gathered in the pre-technology phase of the project. Such comparisons will lead the researchers to an understanding of how speech recognition has impacted the way in which lecturers prepare and deliver lectures, and whether this change has been positive one. Capturing this change is a critical step in demonstrating the implications of speech recognition technology for post-secondary education. As is noted in Steven Estey's evaluation document "…only if the technology makes a person's job easier are they likely to adopt it for use outside the parameters of this project."

CELL SIX of the Conceptual Framework directs our inquiry towards the attitudes and interactions of non-disabled students with respect to speech

recognition technology. As was noted in the proposal entitled “Liberated Learning Phase II: An Innovation to Improve Access in Higher Education Using Speech Recognition Technology”, initial lecture theatre testing of speech recognition in 1998 produced an incidental outcome: non-disabled students were using the display of lecture to enhance their own notes. In light of this, it is beneficial to examine the responses of non-disabled students in this research project. It is hoped that this would enlighten our research in two ways. First, the examination has the potential to strengthen the “generalizability” of the results. Secondly, a documentation and analysis of non-disabled students’ responses to speech recognition may help us to validate the assumption that the technology has broad implications not just for students with disabilities but for the entire sphere of education

CONCLUSION

The Liberated Learning Project will result in dramatic increases in the knowledge and experience base with respect to potential educational applications for speech recognition. The success of the efforts of the Liberated Learning Project team will encourage the continued support from the corporate sector as well as help in expanding the consortium of universities engaged in the Project. Members of the team are confident that the Liberated Learning Project will receive widespread acceptance as a model for universities to better accommodate students in lecture theatres.

Authors

Dr David Leitch
Director,
LIBERATED LEARNING PROJECT
Atlantic Centre of Support for Students with Disabilities
Saint Mary's University
Halifax, Nova Scotia, Canada, B3H 3C3
Phone: 1 902-496-8741 Fax: 1 902-496-8122
david.leitch@STMARYS.CA

Mr Keith Bain
Manager,
LIBERATED LEARNING PROJECT
Atlantic Centre of Support for Students with Disabilities
Saint Mary's University
Halifax, Nova Scotia, Canada, B3H 3C3
Phone: 1 902-496-8741 Fax: 1 902-496-8122
Keith.bain@STMARYS.CA

Ms Di Paez
Equity Officer
University of the Sunshine Coast
Maroochydore DC 4558
Queensland, Australia
Phone: 61 7 54 30 1258 Fax: 61 7 54 30 1210
dpaez@usc.edu.au

Ms Patricia McMillan
Research Assistant

Return to Menu

LIBERATED LEARNING PROJECT
Atlantic Centre of Support for Students with Disabilities
Saint Mary's University
Halifax, Nova Scotia, Canada, B3H 3C3
Work: 1 902-496-8741 Fax: 1 902-496-8122
patricia.macmillan@STMARYS.CA

PAPER IS TO BE PRESENTED BY

Mr Keith Bain and Ms Di Paez