

Development and Use of a Speech Recognition System for Physically Handicapped Users

James Monaghan & Christine Cheepen

University of Hertfordshire, England, UK

Abstract This paper reports on the impact of the implementation of a totally hands-free work station based on an automatic speech recognition (ASR) user interface management system in a school for disabled children with learning difficulties. It focusses on two major areas of impact - firstly the general effects of the introduction of a speech input computer system into the organisational context of a school, and secondly the educational benefit to the student users which arose from the introduction of the system.

1 Introduction

The CALE project¹ had as its aim the provision and implementation of a totally hands-free computer system based on a standard workstation for disabled school children. The research team worked with a particular group of target users at Lonsdale School in Stevenage, Herts., UK, using the school's existing hardware platform - a Nimbus 386, and IBM's VoiceTypeTM as the speech recognition software base². The school caters for disabled children from 3 to 18 years old, all of whom have some degree of learning difficulty, and the particular group of users involved in the research and development project were taken from the 16-18 age bracket. Some work was done with sufferers from cerebral palsy, but in the main the target user group were disabled due to muscular dystrophy - a fatal disease which causes sufferers to become progressively weaker, so that they are unable to walk, and ultimately lose any manual ability.

¹ The CALE project was partly funded by a substantial donation from Save and Prosper Educational Trust.

² VoiceType provides a pre-set dictionary of 5000 words and allows the user to add a further 2000 words, which can be repeatedly updated.

2 Organisational Matters

Clearly in any school situation the smooth running of the day is dependent on observing the timetable and having an overview of where teachers, individual pupils and classes are at any given time. In a school for pupils with special needs this is exacerbated by the pupils having to engage in therapeutic bed rest periods, swimming and the like at specific times of the day. Any technology has to be made available in this context. An ASR based system also has special aspects such as the sound of the user talking possibly disturbing parallel speech-based activity in the same room, and, conversely, the possibility of noisy classroom activities interfering with the speech recognition. As a consequence, the system was installed in a separate computer room with three other computer systems.

All computer systems create tensions in that they represent a significant deskilling of all but specialist computing staff, and in the present environment the primary agenda of most staff is the care for the pupils as a whole. Any difficulties with such a system, due to inexperience, present a very strong temptation to abandon the experiment and substitute it by very hands-on supportive work such as taking dictation. While this approach can be beneficial in terms of creating and maintaining social bonds between staff and students, it inevitably creates an ever increasing dependence of the disabled child on the able-bodied adult. This can - and only too often does - result in deteriorating motivation in the student, and, as we will illustrate in section 4 below, a consequent drop in educational performance, particularly in terms of written language skills.

In order to ensure that this kind of avoidance of the system does not occur, it is necessary, particularly in the initial stage of development, to have support personnel from the research project on site to pick up on obvious teething problems until staff and at least some pupils find it worth while to solve problems using the ASR system rather than to revert to less efficient but more familiar solutions. It is also important to aim at a spread of competence among staff because uneven skill development causes frictions and tends to raise a perception that the technology is 'owned' by specific departments rather than by the school as a whole.

3 Physical Access to a Speech Input System

For pupils with conditions such as muscular dystrophy, numerous difficulties must be overcome. As well as the physical problems which characterise the disease, there is often poor morale and lack of motivation. This, coupled with a life-long physical dependence on able-bodied help, often results in a general lack of concentration and application, so that pupils are unwilling to cooperate in activities requiring any sustained effort. This problem of user attitude became particularly apparent in the matter of providing the users with physical access to the system.

The VoiceType speech recognition system, which is the core of the CALE development, provides a headset for use with the microphone, and this clearly militates against user independence, as it seriously constrains the severely disabled user, who needs help getting in and out of it. To cope with this, we provided a gooseneck microphone stand which they could simply approach in their wheelchairs, thus allowing them total independence when using the system.

Although effective in providing user independence, this created problems in the case of some users. Because of their dependence on teachers and helpers they accepted the constraints of the headset without complaint, but, given the freedom of the gooseneck stand, and having no experience of self-discipline, they simply stopped working whenever they were inclined to do so, and wheeled away, leaving their work unfinished and unsaved. When this was observed, the research team considered removing the gooseneck stand and returning to the headset, but in time, as the users became more proficient, and their satisfaction with the system increased, they became willing to work for longer and to enjoy completing word processing tasks. Our recommendation for future installations in schools is to provide the headset in the early stages, and, when the users have begun to experience considerable success with the system, to introduce the gooseneck stand in order to allow them full independence.

4 Users' Educational Problems and some Solutions

As we have indicated above, severe permanent disability frequently leads to poor morale and lack of motivation in those afflicted. The pupils at Lonsdale School all have some degree of learning difficulty, this being primarily evident in terms of their reading age (which was in most cases some five years lower than their chronological age) and associated written language skills. The reason for this poor level of literacy is, however, often difficult to establish, as experts in special needs education have no way of determining whether the problem is due to a cognitive disability in addition to the physical disability, or simply a consequence of the lack of practice and impoverished feedback which are due to the constraints of the physical disability.

The speech input system which we introduced to the pupils at Lonsdale School proved to be an important diagnostic and remedial tool in this area of literacy skills. In some cases, it became clear that poor literacy levels were not merely apparent, but a fundamental educational problem. This was most certainly the case with a 17 year old girl who suffered from a form of deteriorating cerebral palsy. She appeared to have great difficulty using the system, and was reluctant to come to grips with the effort required to learn how to use it, employing a variety of techniques to distract the researcher who worked with her, such as incessantly chattering, displaying new jewellery etc. After some weeks of attempting (unsuccessfully) to teach this pupil to use the system, it became evident that her level of literacy was extremely low, so that, when faced with the system's 'offers' of candidate words on the screen in response to her voice input item, she was unable to read them, and was therefore unable to choose the right one. For her, and other such users, a speech input system such as VoiceType, used to access a standard word processing package, such as Word, was totally inappropriate. Any attempts to proceed with this kind of user would be likely to result in demotivation, due to the experience of repeated failure rather than success in using the system. Our recommendation for these users - and a problem currently under consideration - is to use speech input for special software packages which provide for high age, low ability education in literacy.

However, in other cases the introduction of a speech input system dramatically improved users' satisfaction, and, as a consequence, their observable literacy skills. Three boys who suffered from muscular dystrophy showed substantial progress after a few months of using the system, showing that their previously observable level of literacy had been artificially depressed by poor morale due to deteriorating ability in using conventional input methods.

One boy, a 17 year old called Glen, had particularly outstanding success with the system. At the beginning of our work with Glen he was causing the staff some concern, as his deteriorating condition had resulted in him becoming severely demotivated, not only regarding his school work, but in the social aspects of his life as well, so that he would not enter into conversation with team members, with the staff, or even with his fellow pupils. He had an apparent reading age of about eleven, and his last written work with a pen was very poor indeed, with badly spelt, and often incomplete words.

Glen's progress over approximately 6 months was remarkable, both in terms of his attitude, which improved enormously, so that he began to communicate with people again, and, very dramatically, in terms of his literacy skills. It became clear that in Glen's case most of his learning difficulties (certainly in terms of literacy skills) were due to lack of feedback and reinforcement. Because, in his later stage, he was unable to produce much written work (in a typical half-hour session of writing he would produce an average of 18-25 words), he was missing out on the practice of writing lengthy pieces of text, and the reinforcement of seeing lengthy texts which were his own work. After six months of using the CALE system, he was able to produce around 150 words of text in a half-hour session, and because using speech as an input medium was less tiring, was able to continue working enthusiastically for much longer. The improvement was not only in the amount of work done, but also in the range of vocabulary used and (as our detailed analyses of a wide range of his work revealed) in his sentence length. When using a pen, Glen's average sentence length was 7-9 words, but with the CALE system his sentence length rose to around 16 words, which is the average sentence length for normal, literate adults.

5 Customisation of Software

VoiceType is a functionally very powerful system, which allows full voice access to any standard, pc based software package which can normally be accessed by the keyboard. This functionality can, however, be rather overwhelming for new users of low educational ability. For the pupils at Lonsdale School we were concerned to ensure that the CALE system would be sufficiently simple so that they could quickly achieve success, in order to encourage their further use of it. This consideration was particularly important in Glen's case, as it was a matter of some urgency to grant him some independence with the system in order to improve his confidence and general morale.

Many of the spoken commands provided by the VoiceType system were unsuitable for Glen, because they were too long for him to remember, for example the sequence which is the default way of loading personalised voice files:

Example 1	voice console
	wake up
	new user
	golf
	lima
	echo
	november
	enter key
	yes please

With this kind of sequence, when the voice recognition is less than perfect (a fact of life with any speech input system in the early stages) the number of errors necessitating correction (without access to the keyboard for a disabled user) can be very discouraging to a novice user of the system. We therefore developed customised commands so that he and his teacher, Sonia, could easily access their own voice files. We set up the system to boot up in a general voice file, and provided Glen and Sonia with their own, trained voice files. They were then able to simply say:

Example 2

change to glen or change to sonia

to carry out the lengthy procedures of Example 1 above, but without the need to make each step explicit by a command word. To unload their voice files and return to the general file, they said:

Example 3

get general

This was a very successful improvement for Glen, as it avoided most of the problems of misrecognition while trying to begin work, and it allowed him to experience early success with the system rather than early failure. The resulting boost to his motivation meant that he was able to tackle the business of using a word processing package by speech with confidence and the expectation of success. We built in several similarly simplified commands for him to use for basic word processing, and provided him with a set of instructions to use instead of the VoiceType manual, so that he would not be discouraged by having to use commands which were lengthy or difficult to say.

6 Summary and Recommendations

With any computer system and any group of users there is inevitably a tension between the provision of full functionality and the human factors requirements of the user. A balance must be achieved between these often competing needs in order to make the system properly usable on a day-to-day basis. The problems are most acute in the case of novice users, as they can be discouraged only too easily from using the system, and in the case of disabled users - particularly in a school context, where there are many able-bodied helpers on hand - the possibility of simply ceasing to use the system is very real indeed.

As preceding sections have shown, our solution to this potential problem was to provide access to the full functionality of the system by gradual stages. In terms of the users' direct interaction with underlying speech software, we initially catered only for their most basic requirements, hiding facilities and functionality, so as not to create confusion. This was done largely by producing customised user documentation and requiring the users to operate according to that, without any reference to more sophisticated instructions and information, such as the VoiceType manual. As the expertise and confidence of the user increases, more facilities can be made available to them, permitting more comprehensive functionality of the system.

With regard to users' physical access to the system, our work at Lonsdale School showed very clearly that special problems tend to occur when the users suffer from life-long, severe disabilities, as they have little self-discipline, and are disinclined to commit themselves to the lengthy periods of sustained effort which are necessary if they are to become proficient with a speech input system. Again, our approach is to allow them fully independent access only in gradual stages, either by restricting novices to using a microphone headset, or, where appropriate, by ensuring that there is close supervision of the user during the early stages of learning how to use the system.

This process of gradually introducing the users to the full power of the system inevitably has a beneficial result in the area of the impact of the new system on the organisation as a whole. Because the progress of disabled school children from full dependence on the teaching staff and helpers to independent use of the system is completed in stages over a period of months, the staff also become increasingly familiar with the system, willing to promote its use, and confident of their new role in relation to pupils who are newly independent of able-bodied help.

Bibliography

Cheepen, C. & J. Monaghan (1993) *Speech interfaces and their effects on written language skills*, Proceedings of Eurospeech '93, Berlin

Gilchrist, P. (1994) *An Evaluation of the CALE System at Lonsdale School*, unpublished BA dissertation, University of Hertfordshire