

Dyslexia and Voice Recognition Software - Really the Perfect Match?

Anne Nelson and Caroline Parker

*Intuitive Internet, 13 Gillespie Drive, Helensburgh, G84 9BL
anne@intuitiveinternet.com*

*Centre for Research into Systems and People, Computing Division, Glasgow Caledonian
University, Glasgow G4 0BA*

Voice Recognition (VR) Software is commonly recommended as an assistive software tool for people with dyslexia. This paper investigates the current use of these software tools, both in the home and in education. A study in 1999 found that there were severe usability problems with the technology (O'Hare and McTear, 1999). Since then however there have been considerable improvements. This study replicates the earlier one to see whether it has improved and what the main usability issues are. The results suggest that despite voice recognition software being a more attractive and viable option than before, it still has many serious usability issues to overcome before it can be recommended.

Introduction

Voice Recognition Software (VRS) would appear to be an obvious choice for supporting users who have difficulty with spelling and writing, or using a standard keyboard. Using a microphone (usually mounted on a headset) it will translate spoken words into written text on the screen, and can read it back to you, as well as performing menu commands and various other instructions. It is often cited as a potentially powerful tool for children with Dyslexia (Miles, 1998; iANSYST Ltd)

VRS has been around for over 30 years, but has until recently been held back by its need for greater than average processing power. However the tremendous reduction in price of the technology has now made its introduction into schools feasible. How useful is it in practice? The most recently documented studies testing the suitability and accuracy of VRS were conducted in 1997 and 1999, respectively, using IBM's **VoiceType** V3.0 dictation package: (O'Hare and McTear, 1999), and **InCube** software: (Zhang, *et al*, Oct 1999). The study by Zang *et al*. however is not relevant to schoolchildren as the age range of the subjects was 20-30 years and the speed of development of the software now means that the O'Hare results are out of date.

This study examines the usability of the more recent VR technologies by replicating the O'Hare study with the most recent ViaVoice software. It also examines in more detail problems encountered by children with dyslexia when using computers, for everyday tasks, at school and at home.

Method

Replication of O'Hare study

Eleven pupils of comparable chronological (12-14yrs) and reading ages (9-10) with those who had undertaken the original study were selected. Testing was carried out over a period of three days, over two consecutive 55 minute sessions, with additional time for set-up and explanation of the tests and software. A short pilot study was undertaken.

The original format of the O'Hare and McTear study conducted in 1997, i.e. samples of text and pictures, was followed as closely as possible. All tasks were monitored for time taken and printed for analysis, Errors were counted and all the data was compared with the original study. **The two main tasks were:**

Structured Text Input (broken down as follows) -

- Hand-written dictation of a piece of text
- Typed dictation of the same piece of text into the computer using a standard keyboard
- Voice-dictation of the same piece of text into the computer using the VRS (following enrolment into the system and training each student to create their own personal voice model)

Spontaneous Text Input

- The children are shown a title for a story and a collection of pictures
- They are then asked to invent their own version of the story and spontaneously dictate it using the VRS

Web-based survey

A Web-based survey was also conducted, the purpose of which was to obtain information about VRS from dyslexic people of all ages and walks of life. The web address of the survey was posted on a large number of online forums and message boards.

Results

Replication study

Results from the practical testing showed that the latest version of the software was faster for voice input than the version used in the previous study. It also showed that voice input was, on average, up to eight times faster than typing, and five times faster than handwriting. However, the original study showed higher accuracy levels for voice input. This was mainly due to time factors relating to the level of training for each student's voice model within the software.

Closer examination of the voice dictation times shows just how much faster dictation times are now with the new version of the software. Average times recorded across all students show that dictation with the latest version was almost twice as fast as the older version (79 rather than 47 words per minute). Though slightly faster overall at all three tasks than their predecessors, the new students still only produced an average of 15 words per minute in the handwritten task, whilst achieving a slow 9 words per minute in the typing task. As far as accuracy is concerned, though, the new software performed worse than the original, averaging 55% accuracy, against 79% in the original study. Handwritten text was the most accurate, averaging over 85% accuracy in both studies.

One of the hardest things for the children to do in relation to the software was train it, training requires reading aloud from text on the screen. Given that many of these children struggle with reading (Orton, 1937/1989) and that reading from the screen tends to be harder than reading from the page (Elkind, *et al.* 1996) this is not a surprising result.

Web Survey Results

The survey generated 220 responses, 59% female, and 41% male, made up of a variety of age ranges. The majority (46%) were from the 22-40 range, while only 12% were under 16. 68% of those who responded had been formally diagnosed with dyslexia, and the other 32% seemed in no doubt that they were. Several comments from this group highlighted the fact that they were only made aware of their own symptoms after a child or sibling had been formally diagnosed.

Computer use

88% of those surveyed use computers every day. When asked what type of user they considered themselves to be, 75% were either experienced or advanced users, with only 7% regarding themselves as beginners. 73% also use computers at home *and* in the workplace.

It therefore seems that even though *all* of the respondents were either formally diagnosed (68%) or at least known to have symptoms of dyslexia, this has not led to an avoidance of computer use, rather, judging by the additional comments received; in most cases computers have made work (and life) much easier for them.

More than half of those surveyed (56%) had never tried any specialist tools, such as text editors, readers or voice recognition software. When asked *why* no specialist tools had ever been used, 58% said it was because they were unaware of their existence. This could also partly be caused by 52% - including those formally diagnosed - never receiving any kind of formal support for their dyslexia. Of the 44% who had tried specialist tools to assist them, the majority seem to have found them fairly easy to use *and* helpful. This would suggest that, were these tools more readily available, they would provide significant help. This is further highlighted by answers given to the question, '*What would encourage you to use/try any software support tools in the future?*' 40% indicated that more publicity and greater availability would encourage them, while 33% would require support and training on how to install and use the software.

Cost was also a priority, with 26% indicating they would try it if it were cheaper, or greater availability of trial versions.

Voice recognition use

Opinions of Voice Recognition software were less positive, however. Of the 20% of respondents who *had* tried VRS, 62% found them to be *very unhelpful* and *very difficult* to use, in comparison to specialist tools in general which 64% of users found either *very* or *quite* helpful or easy to use.

General discussion

The results are mixed. While the new software appears to have improved in terms of the speed with which it can recognise text the accuracy still leaves a lot to be desired. Both the increase in speed and degradation of accuracy appears to be related to the new mechanism that IBM has adopted for speech input. In the old system 'discrete' speech had to be used, slowing the reader down, whereas the new system allows for 'natural' or continuous speech. Practicing isolated speech gave the original students a chance to focus on the pronunciation of each word, encouraging them to speak more clearly. This would undoubtedly produce more accurate results. The new students, being asked to speak in their usual voice produced less accurate results as the software substitutes words it does not recognise with others, which in most cases have completely different meanings. This did however lead to much hilarity and was a great source of entertainment for most of the students.

Background noise is still difficult to filter, and has an effect on accuracy. If the noise is there consistently, the software *will* adapt to it, but in a classroom situation, where noise levels are changing constantly, this could be a problem.

Although voice dictation appears to be considerably less accurate from this short study, it does however become much more accurate over time. Taking that into consideration along with the extremely slow input times for typed and handwritten text, and even allowing additional time for correction of voice dictated text, it would still be possible for some dyslexic users to produce an accurate, word processed document in less than half the time it takes to type. The productivity levels achieved by each student for both typed and voice dictated tasks shows just how significantly faster voice dictation is. However, by analysing each student's results individually and observation during the tests, it was clear that voice dictation would be extremely productive for some students; but not necessarily for others.

Summary

To summarise, the voice recognition software testing shows that:

- The method of 'discreet speech' used in the older version of the software produced much greater accuracy than the newest version, using 'natural speech'
- The interface design for the training procedure in the newest version is still too inflexible to be used by the majority of those with dyslexia – if they cannot read the training dialogue from the screen, they cannot record their voice model, therefore they cannot use the program
- On the positive side, the newer version was much faster than typing or handwriting for all of the students tested
- The children who produced the slowest times and least accurate documents in the handwritten and typed tests, produced the greatest accuracy and fastest times in the voice recognition tests
- Given more time to train the voice model accurately, and more practice reading aloud, the new version would produce much higher accuracy than in the initial short tests
- The children *enjoyed* using the software more than typing or handwriting

The survey results show that:

- Over half of those surveyed were unaware of the existence of any specialist software available to them
- Those who had tried some form of specialist software, generally found it helpful
- Voice recognition software was, on the other hand, found to be unhelpful and difficult to use

In conclusion then these findings show that VRS could be more viable now as an aid to children with dyslexia than it was before, however, time, patience and persistence of both student and teacher are needed. A greater level of publicity and support for these tools is also needed if those who could benefit from it are to be given the chance to do so. Apart from accuracy the main issues which affect VRS usability for dyslexic school age children are:

- The inability to change the background colour on the screen
- The inability to enlarge or change the font used on the screen
- No option to have the training dialogue read aloud – a feature available within the system to read portions of dictated text back to you – but only *after* you have trained it!
- A sensitivity to background noise

Doctors, scientists, and even the Government are now being seen to be actively supporting children and adults with dyslexia in their quest for normality. It seems it may be up to developers to turn these findings into realistic goals.

References

- BBC News Online, July '02, Figures from Gartner Dataquest
- Elkind, J. Black, M.S. Murray, C. Nov 1996 'Computer-based Compensation of Adult Reading Disabilities' *Annals of Dyslexia*, (The Lexia Institute) **46**
- Johnston, D.J. & Myklebust, H.R. 1967, '*Learning Disabilities: Educational Principles and Practices*' (Grune and Stratton, New York)
- iANSYST Ltd. <http://www.iansyst.co.uk/> *Computers and technology for people with disabilities*, Cambridge
- IBM Research Projects/Publications 'Think Research' www.research.ibm.com [online]
- Miles, M. Martin, D. Owen, J. Feb 1998, 'A Pilot Study into the Effects of Using Voice Dictation Software with Secondary Dyslexic Pupils' (Devon LEA)
- Orton S.T. 1937/1989 '*Reading Writing and Speech Problems in Children and selected papers*' (Austin, TX: Pro-Ed)
- O'Hare, A. McTear M.F. Aug 1999 'Speech Recognition in the Secondary School Classroom: an exploratory study' *Computers and Education*, **33**, 1,
- Stephens, C.A. (1997) 'The Application of Voice Recognition Software by Tertiary Students Who Have Specific Learning Difficulties' *Speech and Image Technologies of Computing and Telecommunications*, IEEE Tencon Conference '97
- Seymour, Prof. P.H.K. 1986, '*A Cognitive Analysis of Dyslexia*' (Routledge & Kegan Paul)
- Zhang, W. Duffy, V.G. Linn, R. Luximon, A. Oct 1999, 'Voice Recognition based on Human-Computer Interface Design' *Computers and Industrial Engineering*, **37**, 1-2