REACH Journal of Special Needs Education in Ireland, Vol. 22.2 (2009), 83–92

Hands Free Typing: How do Students with Dyslexia Benefit from Using the Speech Recognition Programme in Microsoft Word?

This article describes a study carried out to research the effectiveness of the Speech Recognition tool in Microsoft Word 2003 used by third level students with dyslexia. Eight students each studying a variety of courses took part in the experiment. At the same time they were registered on an Assistive Technology training programme during which they learned to operate other Assistive Technology software over a number of weeks. The results show the students succeeded in using the Speech Recognition tool when it came to dictating single words, numbers and many commands in a word processor, but struggled with the technology when it came to participating in the initial training setup process and dictating text and other information even after participating in additional training sessions.

JOHN PHAYER is an Assistive Technology tutor in a third level institution and a post-graduate researcher at Mary Immaculate College of Education, Limerick.

INTRODUCTION

The aspect of writing is a highly complex and difficult form of language production (Connelly, Gee and Walsh, 2007). Hatcher, Snowling and Griffiths (2002, cited in Connelly, Campbell, MacLean and Barnes, 2006) state that although little research has been carried out to investigate the writing skills of students with dyslexia, it is believed that writing and not reading is their greatest problem at third level. Consequently, these students must look into the possibility of composing information using different methods – by typing on a computer or even through the use of specialised computerised software such as Voice Recognition Technology – all of which can be decided upon depending on the writer's ease and speed of composing (De La Paz, 1999). Lerner (2000) makes the point that word processing has become a boon for students with learning disabilities who have difficulties in handwriting, spelling and written composition. Therefore, the potential benefits of using Voice Recognition Technology in a word processor could address many of the writing obstacles experienced by students with dyslexia in their studies. De La Paz (1999) points out that as Voice Recognition Technology improves, there is great potential in the way it can revolutionise the composition process and as Gardner (1980, cited in De La Paz, 1999) states 'it permits thinking to unfold in a natural and unimpeded way' (p.173).

Studies carried out by Conn and McTear (2000) show that Voice Recognition software can eliminate the need for using a keyboard. Higgins and Raskind (1997, cited in Lerner, 2000) emphasise that this type of technology is particularly helpful for those who have problems with typing, writing and spelling. Other studies carried out by Curran, Crawford and O' Hara (2005) maintain that using Voice Recognition software allows some people to perform daily tasks without assistance. This paper

discusses an experiment carried out to evaluate the effectiveness of dictating (a) words and numbers; (b) commands and (c) sentences using the Microsoft Word speech training tool by third level students with dyslexia.

RESEARCH QUESTION

The main research question addressed is "Can the Microsoft Word speech engine help to overcome many of the writing difficulties experienced by third level students with dyslexia?"
Within this main research question, a number of sub-questions arise such as:

- How best can the Speech Recognition facility in Microsoft Word 2003 be used to help students with dyslexia with writing in their studies?
- What are the students' experiences of using the Speech Recognition facility in Microsoft Word?
- What are the weaknesses and strengths of using various dictation tools in Microsoft Word by these students?
- How can the speech recognition tool be further enhanced to address other problems experienced by students with Dyslexia?

METHODOLOGY

This experiment was a site specific case study which took place at a third level institution and was based on a 'closed and structured' approach which enabled 'patterns to be observed and comparisons to be made' (Cohen, Manion and Morrison, 2000, p. 248). A small in-depth case study approach was deemed appropriate because of the limited number of participants available. The benefits of a small number allowed for a concentrated study yielding what Cohen et al. (2000) call 'rich and personal data'. The researcher also chose a case study for the following reasons:

- The researcher was limited to only those third level students who had a diagnosis of dyslexia and who were offered Assistive Technology support in the college. The sample size was therefore determined by the limited number of students available and who were willing to participate in this project. The relatively small number of participants was also determined by the uniqueness of the particular group of third level students with literacy difficulties.
- The researcher was limited to those students who attended the Assistive Technology lessons at a specific time on a regular basis and gave their consent to participate in the project.
- Many of these students with dyslexia were assigned other specific Assistive Technology programmes (Visual Managing Software and Voice to Speech Software and other Voice Recognition Software) to use in their studies. Depending on the nature of their difficulty, certain individuals were not assigned the Microsoft Word Speech Recognition tool and were therefore excluded from participation in this study, thereby further limiting the number of potential participants.

In order to conduct a concentrated study of the Speech Recognition tool in terms of its impact on a student with dyslexia, the researcher decided to use interviews and observations to gain a 'fly on the wall' insight of the opinions and views that these

students had about this facility. The researcher tried to develop the study by 'methods of inquiry, an outcome and a resultant record of the inquiry' (LeCompte and Preissle, 1993, cited in Cohen et al., 2000), an approach which is defined as ethnographic research. This was achieved through creating the most realistic and practical environment possible in order to observe and record the students actions and opinions or as LeCompte and Preissle (cited in Cohen et al.) define it, to create 'as vivid a reconstruction as possible of the culture or groups being studied' (p. 138).

What is a Word Processor?

A word processor such as Microsoft Word is a:

> powerful authoring program that gives users the ability to create and share documents by using a comprehensive set of writing tools within an easy to use MS Office user interface (Microsoft Cooperation, 2008a).

The application allows individuals to create professional looking content and documents more quickly and simply. This program also contains a collection of tools and facilities to construct documents from a variety of predefined parts and styles in addition to composing information within Microsoft Word (Microsoft Cooperation, 2008a).

Microsoft Word Speech Recognition Toolbar

The speech recognition facility in Microsoft Word is available in Simplified Chinese, Traditional Chinese, English (US/Ireland) and Japanese language versions of Microsoft Office (Microsoft Cooperation 2008b). This facility allows one to dictate text into any Office program by using a selection of (a) menu (b) toolbar (c) dialog box and (d) task pane options. Speech recognition is not specifically tailored for completely hands-free operation, but individuals will obtain significant results if they use a combination of voice with mouse and/or keyboard interactions (Microsoft Cooperation, 2008b). The Microsoft Word speech recognition facility is located underneath the 'Tools' facility at the top of the main menu as outlined in Figure 1.

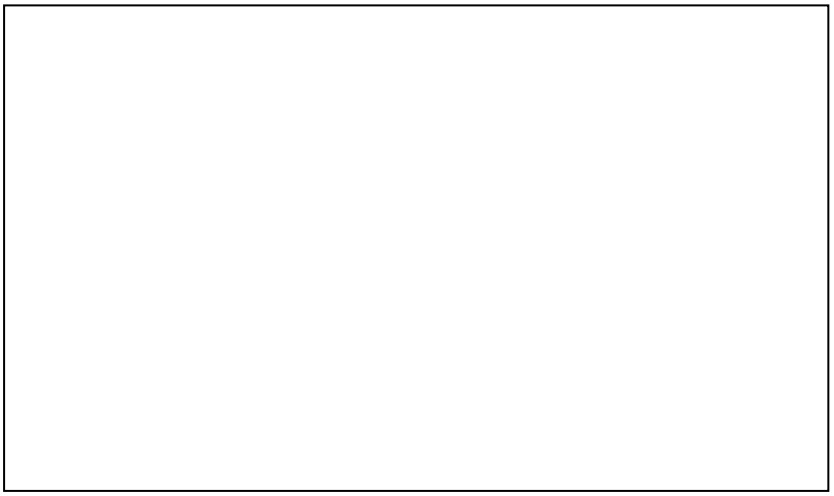


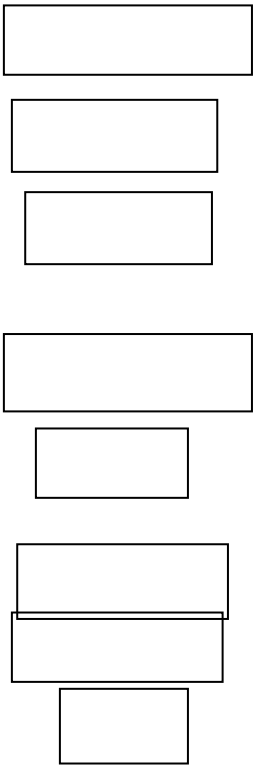
Figure 1: Location of the Microsoft Word speech recognition tool

Once this has been activated, the speech recognition toolbar (Figure 2) is positioned or 'floats' above Microsoft Word and appears as a 'language bar' facility with hidden text labels (Microsoft Cooperation, 2008b).

Figure 2: Microsoft Word Speech Recognition toolbar

First time users must participate in a speech training task so that the computer can recognise how the individual speaks in addition to increasing speech recognition accuracy (Microsoft Cooperation, 2008c). Once this has been completed, the users are offered a variety of tools from which to use when dictating. Table 1 offers a summary and explanation of the main features of this Speech Recognition toolbar.

Table 1: Summary and explanation of the main features of the Speech Recognition toolbar



(a) English (Ireland): Speech recognition engines are language specific and available in Simplified Chinese, US/Ireland English and Japanese.
(b) Microphone: Allows individuals to turn the microphone on/off.

(c) Dictation: Users can dictate information almost anywhere on the screen. When information is being dictated, a blue bar appears on the screen which means the computer is processing the user's voice.
(d) Voice command: Users can select various menu, toolbar, dialog box and task pane commands by simply "saying" their names.
(e) Tools: Allows users to participate in further interactive speech training programs.

(f) Handwriting: Permits individuals to use the handwriting tool to write text.
(g) Drawing Pad: Allows users to draw a variety of shapes in a separate window.
(h) Help Menu/Extra Facilities Menu: Provides the user with a Help facility and an extra facilities menu from which to select other settings.

The students were observed and interviewed over a four week training course in their usual classroom environment. The researcher made every effort to ensure the observations and interviews took place during class in a non-intrusive and natural manner. Initially, all of the students received one-to-one instruction on how to operate the Speech Recognition tool. Each class lasted up to 50 minutes and the training element of the class lasted between 10-15 minutes depending on the individual needs of the students. Some students who experienced difficulties availed of additional training sessions during the four weeks. Anonymity and confidentiality was guaranteed.

RESULTS

From a group of twelve eligible students, eight students with dyslexia participated fully in the study. The following results provide a summary of the students' scores and opinions.

Speech Training

The study began by looking at how long it took each student to participate in one speech training passage in order to build up his/her own individual speech profile. This took each participant approximately ten to fifteen minutes to complete. Each student experienced a certain level of difficulty in training the system to recognise his/her own voice. For example, Andrew[1] who "speaks very slowly" but "is quite accurate and clear at pronouncing words/sentences" took "approximately eight minutes to complete the first training session". Susan "took eleven minutes to complete the training facility". Interestingly, the researcher had to dictate the training passage for this student while she repeated this information (approximately 40% of the training session) using a "Play and Pause" method. The student remarked that "if an easier piece of text was made available for dictation she would have completed this task in a shorter time".

A closer analysis of this question showed that even though many of the participants succeeded in completing the initial speech training setup, a number of them encountered difficulties pronouncing certain "long" words in various passages which they selected to train. Since this feature is one of the most important tasks to achieve before using the Voice Recognition technology itself, the results would suggest that using this facility is quite difficult to operate especially by students with dyslexia who were slow to speak or who were slower to understand what they were doing. A possible resolution would entail providing the participants with simpler reading passages containing easier content to dictate.

Dictation of Numbers and Words

The next part of this facility to be examined was the dictation of numbers. On certain occasions students would have to dictate numbers in Microsoft Word e.g. entering ages into a table, and would therefore require knowledge of being able to use this feature. The researcher asked each student to read aloud random sets of numbers from one to twenty and other numbers with more than three digits in the figure e.g. 12,000,

[1] Student names have been changed.

13,121. From the observations and interviews which evolved in the study, none of the participants demonstrated any major difficulties performing this task or mentioned any negative commentary worth citing. The findings indicate that the participants who use this feature consider it to be a useful tool which can be easily tailored for their own needs and would therefore not require any further improvements.

The third task explored in this study was the dictation of words. This facility allowed students train in the 'add/delete' words facility so that the speech engine tool would better identify various words spoken by them. Many students used this tool as much as possible so that they would achieve better accuracy when it came to dictating information on the screen. None of them experienced any problems using this tool. As a result, there was no interesting commentary about this feature. The findings suggest that this feature is a well designed tool, can be successfully tailored to suit the student's own requirements and therefore would not require any further improvements.

Dictating Commands

The next feature examined by the researcher was to explore how these students succeeded in dictating commands. The students were asked to dictate a selection of commands in Microsoft Word e.g. "File → Save As → Save", "Edit → Select All → Copy → Paste". The findings which emerged in the study showed that all participants achieved a high rate of success with this task and consequently, they showed great enthusiasm using this facility. For example, Amy dictated a "variety of table and insertion commands for about five minutes" and felt "she achieved 99.99% accuracy". Interestingly, Martin succeeded quite well in "dictating certain commands which did not involve the pronunciation of 'T's' such as dictating "Tools → Track Changes" or "Table → Table Autoformat" but overall, the student "felt he achieved about 90% accuracy using this command". The student pointed out to the researcher that "he always had difficulty pronouncing words beginning with the letter 'T' or 'Th'". Andrew felt "he was quite good at dictating commands" and another student, Susan told the researcher that "she dictated commands with an eight out of ten times accuracy". As a result of this commentary, it would appear that this feature is a relatively easy and manageable tool to use by these participants which would not really require further modification.

Dictating Sentences

The final part of the study examined the student's accuracy at dictating sentences in Microsoft Word which is often regarded as one of the most important procedures to achieve when using any type of dictation technology as opposed to computer typing. Amy began to "dictate a variety of words/sentences". Initially, the system did not recognise the spoken words but after the student had participated in two additional training programmes approximately seven/eight words out of ten appeared correctly on the screen. Amy also "dictated an essay from her college studies but it had a lot of incorrect and inaccurate words". The student dictated a further five lines of text from

memory and felt she "achieved 75% accuracy, with the system actually recognising the words she spoke". Even though Martin "experienced many problems pronouncing words which began with 'T's, or 'Th's ", at times "he mixed up letters, e.g. 'a' for 'g' and 'g' for 'a'. The student believed this was due to his "English accent" and resulted in him speaking "at a very slow pace in order for the system to recognise the words/sentences". The participant also pointed out to the researcher that "his greatest problem was misspelling large words in Microsoft Word". This became quite evident when Martin spelt large words such as 'altogether', 'technology' or 'assistance' on the screen. On occasions, he would repeat specific words that he had spoken. Certain words and sentences dictated by the student caused him a great deal of stress and resulted in him becoming quite anxious about using the technology.

Andrew found that "the speech training programme would not play back the information which he spoke". James was "quite sharp at spotting his own mistakes". Quite Interestingly, James "believes the voices in TextHelp" (a Text to Speech programme) "are a lot better than those in Microsoft Word but at the same time would prefer to type rather than using Speech Recognition software".

Summary

It is apparent from the above study that no single feature of the Microsoft Word speech engine emerges as being the most useful feature for these students. The initial training setup of the speech engine and the dictation of sentences facility did cause a lot of 'teething problems' for those students who participated in the study but over time, these difficulties reduced with additional training. The participants who used the 'dictating commands' and the 'dictating numbers' facilities considered them to be very useful in meeting their needs. Overall, some students indicated that they found the Voice Recognition technology to be a useful study support and would consider using this application for dictating (a) essays, (b) reports and (c) numbers into tables as well as using it to (a) read back information and (b) listen to the spelling of words in their studies.

One way of using this technology in an efficient manner is to use a combination of voice commands with keyboard/mouse instruction (Microsoft Cooperation, 2008b). Further analysis of this facility would suggest the dictation facility does require further improvement in terms of displaying sentences in a more accurate manner for the reader which will determine its success or failure by those who use it. One factor which contributes to dictating accurate information on a screen is determined by the person's clarity and rate of speaking into the speech system. As participants train other dictation passages, the speech engine will better recognise the user's own voice and hence display more accurate information on the screen all of which require greater effort by the student. Additionally, providing a greater selection of playback voices to read back text for these students would be of immense help to them. This dictation facility in Microsoft Word does hold great promise for those students with dyslexia who use it, but how it is used depends on the way they can customise it to suit their own needs.

CONCLUSION

This paper examined how third level students with dyslexia use the speech training engine in Microsoft Word in their studies as well as finding out what they found to be the most beneficial/frustrating features of this facility. From the findings above, dictating commands, numbers and training words seem to cause less difficulty as opposed to the initial start up training facility and the dictation of sentences which are areas that require further investigation and improvement. The current study has its own limitations that might have been somewhat addressed with the use of a measurement tool to analyse the accuracy of this dictation facility and compare it with the accuracy of computer typing. Further research could address these issues.

REFERENCES

Cohen, L., Manion, L. and K. Morrison (2000) *Research Methods in Education* (5th ed), London: Routledge.

Conn, N. and M. McTear, (2000) Speech Technology: A Solution for People with Disabilities, Paper presented at IEE Conference, April 2000, http://ieeexplore.ieee.org/search/searchresult.jsp?query1=McTear&scope1=au&queryText=+%28%28mctear%29%3Cin%3Eau+%29+&history=yes&menu1=&query2=&menu2=&scope2=&query3=&menu3=&scope3=&op1=&op2=&queryblock=&reqloc=basic&py1=1950&py2=2009 (accessed 21st February 2008).

Connelly, V., Campbell, S., MacLean, M. and Barnes, J. (2006) 'Contribution of Lower Order Skills to the Written Composition of College Students with and without Dyslexia', *Journal of Developmental Neuropsychology,* Vol. 29, (1) pp. 175 – 196, http://ssl.brookes.ac.uk/sslDataForms/up1oads/MMacLean/Developmental%20Neuropsychology%202006.pdf (accessed 19th July 2008).

Connelly, V., Gee, D. and E. Walsh (2007) 'A Comparison of Keyboarded and Handwritten Compositions and the Relationship with Transcription Speed', *British Journal of Educational Psychology*, Vol. 77 (2), pp. 479 – 492.

Curran, K., I. Crawford and L. O' Hara (2005) 'Catering for the Disabled Surfer: A Case Study in Web site Navigation for Disabled Students', *EASI (Equal Access to Software and Information) for Information Technology and Disabilities E-Journal*, Vol. 10 (1), pp. 1-12, http://www.rit.edu/~easi/itd/itdv10n1/curran.htm (accessed 3rd November 2008).

De La Paz, 1999 'Composing via Dictation and Speech Recognition Systems: Compensatory Technology for Students with Learning Disabilities', *Learning Disability Quarterly,* Volume 22 (3), pp. 173-182, http://www.jstor.org/stable/pdfplus/1511284.pdf (accessed 30th July 2008).

Lerner, J. (2000) *Learning Disabilities: Theories, Diagnosis and Teaching Strategies* (8th ed), Boston: Houghton Mifflin Publishers.

Microsoft Cooperation (2008a) 'Microsoft Word – Microsoft Office Word 2007 Product Overview', *Microsoft Office Online,* http://office.microsoft.com/en-gb/word/HA101656411033.aspx (accessed 30th July 2008).

Microsoft Cooperation (2008b) 'Microsoft Word – About Speech Recognition', *Microsoft Office Online,* http://office.microsoft.com/en-gb/word/HA101656411033.aspx?mode=print (accessed 30th July 2008).

Microsoft Cooperation (2008c) 'Microsoft Word – Getting started with Speech Recognition', *Microsoft Office Online,* http://office.microsoft.com/en-us/word/HP030850221033.aspx?mode=print (accessed 30th July 2008).