

# Speech recognition

- 1 Speech recognition systems**
- 2 Speech recognition for learners with writing difficulties**
- 3 Discrete and continuous speech recognition**
- 4 Continuous speech recognition programs**
- 5 Digital recorders and speech recognition**
- 6 Assessment for a speech recognition system**
- 7 Introduction, training and use in schools**
- 8 Speech recognition summary**
- 9 Suppliers**
- 10 References**

---

# SPEECH RECOGNITION

---

*“For so long they [dyslexic pupils] have had to rely on others to scribe for them that they have begun to despise themselves and ‘switch off’. VoiceText [DragonDictate] allows these pupils to respond independently, gives them control, and gives them the freedom to express themselves as others do.”*

(Donald, 1998)

Speech recognition, as the quotation above suggests, has great potential for helping pupils overcome the barriers presented by their spelling and writing difficulties. The technology and its application are both relatively new, however, and many issues have still to be investigated and resolved. Which groups of students are liable to gain from speech recognition? Which system is most effective? What is the best way to introduce and train students to use the programs? How can students and schools get adequate technical support? Will the use of speech recognition give students with specific learning difficulties an advantage over their peers?

We do not attempt to answer these questions in this book because the necessary research and experience does not yet exist and because this book was originally conceived to address the older, more established supportive writing tools such as word processors, spellcheckers and word predictors. This section therefore attempts to draw together current research and experiences of speech recognition and suggest partial answers to some of the issues raised above.

## 1 Speech recognition systems

A few years ago, speech recognition programs were very expensive and required a powerful (and expensive) computer. Now, a speech recognition program can be bought for around £50 and will run on almost all Windows PC computers on sale today (but may not run on machines a few years old). A speech recognition program takes spoken commands and dictation from a headset microphone worn by the user, and tries to match the sound to the correct word or command. Because voices vary so much from person to person (male/female, different ages, accents, dialects etc.) most speech recognition programs need ‘training’ before they can be used effectively. Even after a program is trained to the user’s voice, it will still mis-recognise words, so the writer must be able to spot any errors and correct them. Speech recognition programs are therefore not at present suitable for people with very limited literacy skills.

There are several speech recognition programs available: *DragonDictate*, *VoicePower Pro* and *Dragon NaturallySpeaking*; *IBM ViaVoice* (the older *IBM SimplySpeaking* has been discontinued, although it may still be available in some shops); Lernaut and Hauspie *VoiceExpress*; and Philips *Freespeech*. Some of the programs are available in several different versions. The most popular programs used in schools at the present time are: *DragonDictate Classic* and *Dragon NaturallySpeaking*, and *IBM SimplySpeaking* and *IBM ViaVoice*. We will not go into great detail about the different programs here: see CALL’s information sheet on speech recognition for more information (Wilson, 1998); iANSYST’s web site at <http://www.dyslexic.com>; and the other web sites listed at the end of the book. By the time you read this, products and prices listed here will probably have changed, so you should contact the suppliers to get up to date details.

Almost all of the speech recognition programs work on Windows PC computers, apart from *VoicePower Pro* for Macintoshes. Speech recognition is a complex task which requires a fairly powerful machine: the ‘discrete speech’ *DragonDictate* program

can run on older 486/66 Windows 3.1 systems, but all the other PC programs need Windows 95 and a fast Pentium computer. *VoicePower* requires a PowerPC Macintosh with MacOS 7.5.1 or later, and at least 32 MB RAM. The current version is fairly limited – it comes with a US rather than UK dictionary so many words are spelled incorrectly, and you need an extra adapter getting it to work with some Mac Performas. Good news for Mac users is that Dragon has started developing a version of *Naturally Speaking*, for iMacs and G3 Macs, which they say will be available at the end of 1999.

There are two types of speech recognition programs: ‘discrete’ and ‘continuous’ speech. With discrete speech systems the user must leave a short gap (around 0.25 seconds) between each word; with continuous speech, the user can speak more or less normally. Advantages and disadvantages of discrete and continuous systems are discussed below.

## 2 Speech recognition for learners with writing difficulties

Speech recognition systems have been used by people with disabilities for several years. They have proved effective for writers who have problems with a standard keyboard because of arthritis, repetitive strain injury, or spinal injury. They have also been used by writers with reading or spelling difficulties, and visual impairment.

Experiences in schools have been mixed however: some pupils have found the systems extremely beneficial, while others have found them extremely frustrating and of very little use. We will explore the reasons for these mixed results by looking at two recent successful studies.

### Perth High School, Scotland

Elaine Donald, Principal Teacher of Learning Support at Perth High School has had considerable success with *DragonDictate*. This was one of the first programs to be used in schools, and is a ‘discrete’ speech recognition program. In 1995 Donald started a pilot study looking at the use of *DragonDictate* by twelve pupils from S1 to S5. All of the pupils, apart from one, had specific writing and spelling difficulties. It is important to note that those who took part in the study had ‘reasonably good word recognition skills’ (Donald, 1998). Donald concluded that *DragonDictate* was appropriate for writers with the pattern of skills and difficulties shown in Table 1.

Skills	Difficulties
wide general knowledge	inability to spell consistently
ability to read silently/good word recognition skills	poor organisational skills
good oral ability (indeed some are highly articulate)	fluctuating concentration
good comprehension of orally delivered material	sometimes immature handwriting
some may have good comprehension of written material	reluctance to put pen to paper
good social skills, interacting well within the peer group	difficulty with sequencing including following instructions
	poor use of language in written form
	poor punctuation skills
	low self-esteem

By August 1998 more than 30 pupils were using the system in Perth High School, sharing six laptops between them. Pupils are able to book a laptop computer to take to a specific class for an assignment, or to complete work in the Learning Support Base. Improvements observed included:

- an increase in volume and quality of work;
- great improvements in self-esteem;
- improvement in reading ability;
- greater motivation to write;
- more sophisticated language and vocabulary;
- better punctuation.

Pupils have also used speech recognition successfully in Scottish Standard and Higher Grade examinations.

Factors which appear to have contributed to this success (and which have been lacking in schools where speech recognition has been unsuccessful) include:

*Technical support* – The computers used for speech recognition in Perth High School were supplied by a local firm. They were supplied with the program already installed, tested and working, and later technical difficulties were overcome quickly. Some other schools have purchased a computer from one source and a program from another and found that the two do not work well together. Unfortunately this seems to be a fact of life with PC's – because of the wide range of different computers and sound cards available, some PC's seem to work better than others. Therefore, it is essential to either buy the computer and program together from the same supplier, or check that the computer is able to run the program before you order it.

*Microphones* – Initially at least, Donald used high quality Shure microphones which were more reliable and comfortable than the relatively basic headsets supplied with the programs. If you are intending to use speech recognition seriously, it is worth buying a better microphone such as the TalkMic (£39 – £49, iANSYST).

*Staff resources* – Perth High School is a *Resourced Location for Dyslexic Pupils* and therefore has a relatively large learning support department. This means that time was made available to train the pupils to use *DragonDictate* properly.

*Training procedure* – Donald developed a systematic procedure for introducing speech recognition to the pupils, which involved first completing the standard *DragonDictate* training and then tackling set exercises based on Cassell's English Comprehension Cards (Cassell, 1979). This process ensured that the pupils were confident with the system before they started to use it for classwork.

## Devon LEA

Martin Miles, Di Martin and Jim Owen (Miles, Martin and Owen, 1998) investigated the effectiveness of *IBM Simply Speaking* for older school pupils with specific learning difficulties. *SimplySpeaking* is a discrete recognition system, like *DragonDictate*. It is generally regarded as easier to get working than *DragonDictate*, but has a different method of operation. With *Dragon*, the writer speaks, checks and if necessary corrects one word at a time. With *SimplySpeaking*, the writer usually dictates a sentence, or paragraph, and then goes back to correct any mis-recognised words.

These different methods of interaction may seem trivial, but in fact they do demand quite different approaches to dictation and writing and so it is important to try the program which best suits the writer's abilities.

Eleven pupils aged between 14 and 16 were trained to use the system and their progress monitored over a 10 week period. Ten of the pupils had Statements of Need. Training involved showing the pupils how to: set up the microphone and select their own voice file; read from set texts; and correct any mis-recognitions. Later in the term the pupils started to use the systems for curriculum based work and free writing.

The research team reported:

- an improvement in the programs' recognition accuracy;
- an increase in the amount of hand-written work, fewer spelling errors and a significant improvement in the quality and quantity of curriculum based work;
- the reading age of the pupils improved by 13.4 months, on average, over the 10 week period (according to the British Ability Scales Reading Test);
- the average spelling age improved by 6.1 months (measured with the Schonell Spelling Test);
- improved self-esteem and motivation.

Miles suggests that these improvements in literacy skills were due in part to the writers "focusing intently upon the words that they were dictating and the computer's attempt to match these" (Miles, Martin and Owen, 1998).

Difficulties observed included:

- frustration in the early stages;
- some pupils with reading difficulties needed a lot of one-to-one support to help them read the set sentences to train the system;
- some pupils with reading difficulties needed help to identify words that were mis-recognised by the system;
- disruption from background noise.

### **Other research and reports**

The Speak to Write project (<http://www.edc.org/spk2wrt/>) in the United States has been investigating discrete speech recognition systems for people with disabilities and suggest that the systems can help writers:

- use longer and richer words;
- write more creatively;
- organise their work better;
- complete more work.

In the UK, BECTA are currently monitoring speech recognition system in 12 centres in England and a report is due in the early part of 2000. Information about the project can be found on the BECTA web site at <http://www.becta.org.uk/projects/voicerecog/>. The Scottish Office Education and Industry Department are also considering a proposal from the CALL Centre, to investigate and report on the current

use of speech recognition systems in schools in Scotland.

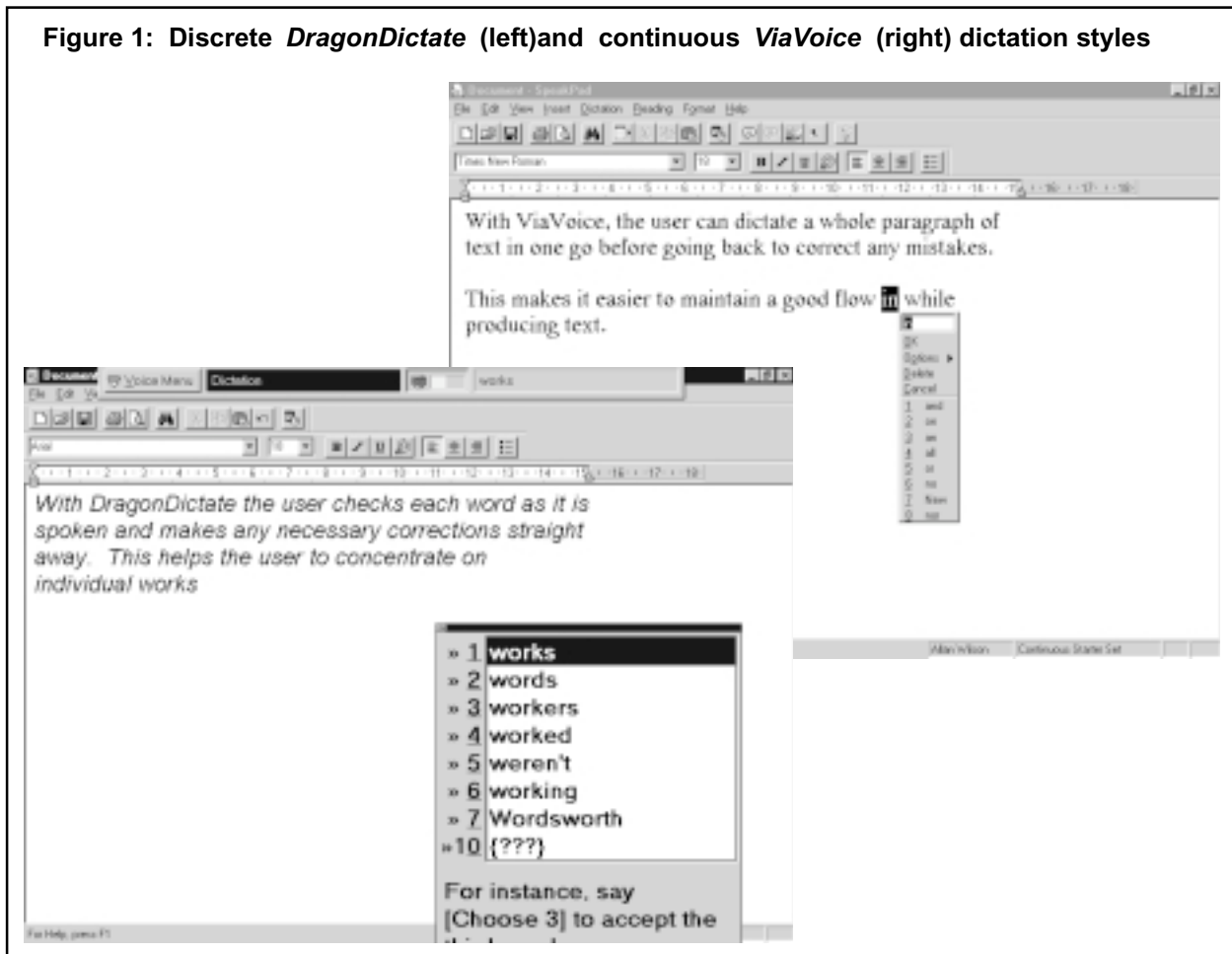
These studies, together with experiences in schools, have shown that speech recognition systems can have benefits for students with specific learning difficulties. Success appears more likely where the school has adequate technical support and staff resources, and where the students are trained to use the programs systematically. Success is also more likely with older primary or secondary school pupils who have reasonable general and word recognition skills, rather than young writers or students with significant reading and literacy difficulties.

### 3 Discrete and continuous speech recognition

The studies described in Section 2 used discrete speech recognition programs, where the user has to pause briefly between each word. *SimplySpeaking* has now been discontinued so for the purpose of this discussion 'discrete program' now means '*DragonDictate*', or for Mac users, '*VoicePower Pro*'.

While the newer continuous speech programs can recognise more natural speech and should therefore be faster and more intuitive to use, experience suggests that some writers with specific learning difficulties may actually be better off using the older *DragonDictate* technology. This is mainly due to the differences in dictation technique between *DragonDictate* and the continuous systems. With *Dragon*, one word at a time is spoken, checked, and if necessary corrected by the writer. With the continuous systems, the writer dictates a sentence, or paragraph of text, and then proofreads to correct any mis-recognised words.

Figure 1: Discrete *DragonDictate* (left) and continuous *ViaVoice* (right) dictation styles



Discrete systems suit writers who:	Continuous systems suit writers who have:
<ul style="list-style-type: none"> <li>• have difficulty reading and expressing thoughts clearly</li> <li>• race ahead or have erratic speech patterns or a tendency to language blocks</li> <li>• need direct contact with the written word so that he or she can bridge the gap between speech and writing</li> <li>• want to improve and develop spelling, reading and writing</li> </ul>	<ul style="list-style-type: none"> <li>• clear, controlled thoughts</li> <li>• no difficulties finding words to express themselves</li> <li>• poor spelling and prefer to get down thoughts quickly</li> <li>• strong reading skills</li> </ul>

Table 2 presents advantages and disadvantages of discrete and continuous systems from Emma Elliott and Judy Thomas (Elliott, 1998) who have used *DragonDictate* extensively.

The Speak to Write project (<http://www.edc.org/spk2wrt/>) suggest the following about continuous speech recognition programmes:

Advantages:	Disadvantages:
<p>¥ they are faster and easier for the average user who does not have significant linguistic or writing difficulties;</p> <p>¥ they may be more accurate to start with, and require less training because they are newer and have better algorithms.</p>	<p>¥ the writer must dictate chunks of text so must be able to formulate sentences before starting to dictate; good pronunciation is needed — it needs to be clearer than when using a discrete system</p> <p>¥ with older, slower computers, there may be a delay before the text appears on screen after the words have been spoken, which can be disconcerting (although this is not a problem with newer machines)</p> <p>¥ sometimes a word is changed by the program after it has been typed on screen, because the program takes context into account when working out the correct word — this can be confusing for some writers</p> <p>¥ the writer must learn to speak written rather than spoken English - a skill which emergent writers may find difficult</p> <p>¥ the voice models used in discrete systems are based on single words so the systems may adapt more quickly to the voices of young children.</p>

The Speak to Write web site suggests that discrete systems are suitable for writers who:

- need direct contact with the written word so that he or she can bridge the gap between speech and writing
- want to improve and develop spelling, reading and writing

Elliott and Thomas are unequivocal about which program is best for writers with severe literacy difficulties: “*the most supportive system is DragonDictate and Keystone*” (Elliott, 1998). The key reasons for their preference are the word-by-word dictation style which forces the writer to correct each word immediately after it has been

spoken, plus the use of *Keystone*, which echoes back the word as it is typed. There seems little doubt that continuous systems demand different, if not greater literacy skills from the user: the writer must be able to compose a sentence before starting to write, and must be able to read back over the text to identify mis-recognised words, and then correct them.

However, continuous systems do have advantages: they are faster and more accurate (provided the writer can train and operate them); the dictation style suits some writers who prefer to get the text down then go back and correct any mis-recognitions; some users find the necessity with *DragonDictate* to check every word as it is spoken tiring and stressful; and some people find that *ViaVoice*, for example, is easier and less frustrating to train. The authors' experience with disabled students at the University of Edinburgh has also favoured *SimplySpeaking* (and now *ViaVoice*) because the programs have proved so easy to set up and use quickly and reliably compared with *DragonDictate*.

The study by Miles, Martin and Owen (Miles, Martin and Owen, 1998) may also suggest that continuous speech systems can be effective for writers with literacy difficulties. Although Miles and his colleagues used the discrete *Simply Speaking* program, it operates like a continuous system because complete sentences are dictated and then corrected afterwards. Therefore, one might expect similar results with a continuous speech system. For this to happen, Miles (Miles, 1999) suggests it may be helpful to train pupils to speak *discretely* when using *ViaVoice*, so that they can focus on each word and therefore benefit from the multi-sensory learning that appears to be associated with discrete speech systems. For this to work, the "Attention word required" option in the *ViaVoice* Command options must be selected, otherwise the program can interpret a pause followed by a word as a command, rather than dictation. One advantage of using continuous systems 'discretely' is that the writer has the possibility of increasing speed and talking more continuously once the program has been trained and he or she has learned how to use it. The latest version of *textHELP! Read and Write* is claimed to be able to speak back each word recognised by *ViaVoice*, which will further help writers focus on individual words during dictation.

In summary: *DragonDictate*, or a continuous program used discretely, is probably more suitable for writers who have the sort of difficulties summarised by Elliott in Table 1. Writers with older computers may also choose *DragonDictate* because the newer continuous speech programs need Pentium 166 or faster machines. Continuous speech programs are worth looking at if the writer can compose and dictate complete sentences; and can proofread to identify and correct mis-recognitions.

#### 4 Continuous speech recognition programs

Of the continuous systems available, *IBM ViaVoice* and *Dragon Naturally Speaking* seem to be the most popular programs in schools at present. *Dragon* usually needs training before it can be used productively, and the training involves reading a relatively large amount of complex text, so it is only suitable for good, motivated readers. Table 4 gives a summary of the available programs. Refer to <http://www.dyslexic.com> for a more detailed comparison of the systems.

There are several versions of both programs. *Dragon NaturallySpeaking Preferred* is the most useful version for schools because it can be used to dictate into any application, it has text-to-speech feedback for proofreading, and records the writer's speech which makes editing easier. Of the *ViaVoice* products, *ViaVoice Executive* is better because it

can type into any application – *ViaVoice Home* can only be used to dictate into a simple *SpeakPad* word processor, or *Microsoft Word 97*. If you want to try one of the less expensive versions first, you can upgrade from *ViaVoice Home* to *Executive* and from one of the cheaper Dragon systems to *Preferred*.

Program	Machine	Description	Cost	Supplier
<i>Dragon Point &amp; Speak</i>	Win 95/98, NT	Requires Pentium 166 processor (minimum) and at least 32 MB of RAM. Allows dictation into a number of Windows packages, but with limited editing and correcting functions. No text to speech playback.	£45	iANSYST, PC World, Dixons, etc.
<i>Dragon NaturallySpeaking Standard</i>	Win 95/98, NT	Requires Pentium 166 processor (minimum) and at least 32 MB of RAM. Allows dictation into any Windows package, with full editing and correcting functions. Navigation by voice, but no text to speech playback. Training required to create a voice file.	£81	iANSYST, PC World, Dixons, etc.
<i>Dragon NaturallySpeaking Preferred</i>	Win 95/98, NT	As <i>NaturallySpeaking Standard</i> , but records user's speech and allows text to speech playback.	£175	iANSYST, PC World, Dixons, etc.
<i>Dragon NaturallySpeaking Professional</i>	Win 95/98, NT, Mac	As <i>NaturallySpeaking Preferred</i> , but with improved command and control functions and more dictionary options, including facility for creating different topic vocabularies.	£539	Various, including iANSYST
<i>IBM ViaVoice98 Executive</i>	Win 95/98, NT	Requires Pentium 166 MMX processor (minimum) and at least 32 MB of RAM. Allows dictation into any program and navigation between, and within, programs by voice. Records user's speech and allows text to speech playback. Many users can get reasonable levels of recognition without training the system.	£137	iANSYST, PC World, Dixons, etc.
<i>IBM ViaVoice98 Home</i>	Win 95/98, NT	Same as above, but without the voice navigation features and restricted to dictating into <i>Microsoft Word97</i> , or it's own basic <i>SpeakPad</i> word processor.	£49	iANSYST, PC World, Dixons, etc.
<i>Lernout &amp; Hauspie Voice Express Standard</i>	Win 95/98, NT	Requires Pentium 166 MMX processor (minimum) and at least 32 MB of RAM. Can only be used with Voice Xpress word processor, though text can be cut and pasted into other applications.	£40	iANSYST, PC World, Dixons, etc.
<i>Lernout &amp; Hauspie Voice Express Advanced</i>	Win 95/98, NT	As <i>Voice Express Standard</i> , but with the ability to dictate directly into <i>Microsoft Word97</i> . Includes speech files designed for use with children.	£80	iANSYST, PC World, Dixons, etc.
<i>Lernout &amp; Hauspie Voice Express Standard</i>	Win 95/98, NT	As <i>Voice Express Advanced</i> , but with the ability to dictate directly into any <i>Microsoft Office</i> application. Comes with improved microphone.	£120	iANSYST, PC World, Dixons, etc.
<i>Philips FreeSpeech98</i>	Win 95/98, NT	Requires Pentium 166 MMX processor (minimum) and at least 32 MB of RAM. Comes with SpeechMike combined microphone, speaker and trackball. Good quality recognition, but poor editing facilities and can only be used by one person.	£70	iANSYST, PC World, Dixons, etc.

When deciding on a speech recognition program, bear in mind this comment from Ian Litterick at iANSYST:

*'A final factor is that different people and different computers sometimes get inexplicably different results from what it would be reasonable to expect.'*

(Litterick, 1999)

## 9.5 Digital recorders and speech recognition

A new development is the use of digital recorders designed to work with speech recognition programs. The idea here is that the writer can dictate into the recorder, then plug the recorder into the computer, load up the speech recognition program, press 'Play' on the recorder, and sit back while the computer converts the recording into text on screen. This approach has obvious advantages for schools: although digital recorders are more expensive than Walkman-style standard types, they are a lot cheaper than Windows laptop computers, so several students could use their own recorders and then download, correct, edit and print out their work on a desktop PC. However, at the present time these systems are relatively unproven and should be approached with some caution. First, the additional recording-playback process is likely to reduce accuracy. Secondly, and more importantly, the tape recorder does not provide the same quick feedback on screen of the text dictated as is provided when dictating straight into the computer. So while a practised writer with good literacy skills may be able to produce good work with a recorder, it is not a suitable approach for writers who have difficulty composing sentences and structuring their work or who have poor short term memory skills. The use of digital recorders in this way is also unlikely to produce the same improvements in basic literacy that have been reported by Donald, Miles et al, and others.

Program	Machine	Description	Cost	Supplier
<i>Dragon NaturallyMobile V3.5</i>	Win 95/98, NT	This is <i>Naturally Speaking Preferred</i> with a <i>Voice-It</i> digital recorder which gives 40 mins recording time.	£169	iANSYST, PC World, Dixons, etc.
<i>Olympus D1000 Recorder with Via Voice Exec V4</i>	Win 95/98, NT	This recorder gives only 15 minutes recording time and comes with a slightly older version of <i>ViaVoice Executive</i> .	£255	iANSYST, PC World, Dixons, etc.

## 9.6 Assessment for a speech recognition system

The previous discussion presented factors to be considered when investigating whether a speech recognition program is appropriate for a student with writing difficulties, and if so, which system should be evaluated. These assessment issues are summarised below:

*Age* – Speech recognition programs at the present time are designed for adult voices, and therefore do not usually work well with children younger than about 9 or 10.

*Cognitive skills* – Writers must have reasonable general cognitive abilities, word recognition skills, and good articulation.

*Motivation* – Speech recognition systems can be extremely frustrating to train and

learn to use, so the writer must be very motivated and supported during the difficult initial stages.

*Dictation skills* – The essential skill needed for speech recognition is to be able to compose and dictate clear, well-structured sentences. If the student cannot put a meaningful sentence together orally, speech recognition is unlikely to help. Spoken English is also different from written English, and students must learn how to speak ‘written’ English: ‘*Anecdotal evidence suggests that students who have started to use continuous speech systems to alleviate RSI, for example, have often developed a much “chattier”, less formal, style of writing.*’ (Wilson, 1998).

Assuming the writer has the necessary dictation skills for speech recognition, the first thing to check is his or her preferred dictation style. This can be done without a computer – ask the student to read or listen to a short text then ask questions about it. Can the writer compose a complete sentence in his or her head, before speaking it? If so, then continuous speech recognition is a possibility. If the writer has difficulty thinking of the next word, stumbles over words, speaks incorrect words or words in the wrong order, then *DragonDictate* is more likely to be suitable. Review Elliott’s suggestions in Table 9.2: does the writer fit the profile of a discrete or a continuous program user?

*Identification of mis-recognised words* – Next, you need to consider if and how the writer can identify mis-recognised words. Type a paragraph of appropriate text into a talking word processor and then replace around 20% of the words with other words – or use an uncorrected sample from an earlier dictation session. Then ask the writer to read the text, with and without text-to-speech feedback to establish whether they can identify the mis-recognised words. Writers with good reading skills who can manage this may be more suited to a continuous system, and both *ViaVoice* and *Naturally Speaking Preferred* have text-to-speech facilities built-in. Note that the built-in text to speech systems are fairly basic, and iANSYST recommend *textHELP!* (£95, iANSYST, Inclusive Technology, and others) which gives more flexible control over the speech.

If the writer has difficulty identifying mis-recognised words in the paragraph then *DragonDictate* is probably more suitable.

*Text-to-speech feedback as you dictate* – Another consideration is whether the writer needs text-to-speech feedback of each word or sentence *immediately after it is dictated* (as opposed to selecting a complete paragraph and then speaking it out, as described above). At the present time, most suppliers and users recommend either *DragonDictate* plus *Keystone Lite*, or *DragonDictate* plus *textHELP!*, for word by word speech feedback. *textHELP!* is reviewed in Chapter 13. The authors have tried various combinations of talking word processors and speech output tools with programs like *SimplySpeaking* and *ViaVoice*, but we have not been able to make them work properly. However, iANSYST report that the latest version (3.2) of *textHELP! Read and Write* can speak back each word as it is dictated using *ViaVoice*.

*Correction of mis-recognised words* – Once the writer has identified a mis-recognised word, it must be changed for the correct word. The correction serves two purposes: to correct the text itself, and also to teach the speech recognition program the writer’s pronunciation of the word.

With the continuous programs, the writer clicks on the word to be changed, or selects it by voice, and is offered a list of likely alternatives. With *DragonDictate*, the list is usually set to appear as each word is spoken. The writer looks down

the list and if the correct word is offered on the list, selects it using voice, the keyboard or the mouse (different programs have different selection techniques). If the word is not offered in the correction list, the writer must either type it in, or speak it in letter by letter. *DragonDictate*, *NaturallySpeaking* and *IBM ViaVoice* all have a predictive facility which helps poor spellers type the word in correctly.

So, if the writer intends to correct *independently*, he or she must have good word recognition abilities (to spot the word in the correction list), and some spelling skills to type it in if it is not offered in the correction list.

## 7 Introduction, training and use in schools

*Hardware* – You must make sure that your computer, soundcard, microphone and speech recognition program are compatible. Some computers (particularly laptops) are very ‘noisy’ and the signal from the microphone gets degraded and therefore effects recognition accuracy. A quick way to test the quality of the sound input is to record ‘silence’ using the built in Sound Recorder in the Windows Control Panel. If the recording shows noise when it should be silent, your sound card may be poor quality and you may have difficulties getting the speech recognition program to work. IBM and Dragon both have lists of computers which are known to be compatible with their systems, on their web sites, or you can ask the computer supplier to demonstrate the machine working with speech recognition, before you buy it. Get a computer with a recommended sound card, such as the *Turtle Beach Malibu*, or a *Soundblaster Live!* (the cheaper *SoundBlaster PCI 64* is not recommended by iANSYST).

The microphone is an essential part of the speech recognition process. The microphones supplied with the programmes are adequate but often do not stay in place on the user’s head, so it is worth investing in a better device such as the *TalkMic* (£39 – £49, iANSYST). It is also a good idea to provide a microphone for each student using the system, rather than sharing a microphone between them.

*Management* – It usually takes many hours of practice to train a speech recognition program and create a ‘voice file’ for an individual writer. It is essential to avoid losing or damaging this voice file, so every user of the system must be trained to load in their voice file before they start, to avoid inadvertently using and modifying another writer’s. The best approach is to store voice files on ZIP discs, which are carried by the writer, along with his or her personal microphone.

*Training the system* – John Lubert and Scott Campbell (Lubert & Campbell, 1998) in Canada have developed a useful 10 step training program for *DragonDictate*. Even if you are using a different program, the insights and principles that are described in their manual are well worth following. The manual can be downloaded from <http://snow.utoronto.ca/best/manual2.html> A brief summary of Lubert & Campbell’s training is given below.

### *Preparation*

- a) The trainer (i.e. teacher or staff member training the writer) must have a thorough knowledge of the system and must have trained and used it themselves before it is introduced to the writer.
- b) Before starting, the trainer should demonstrate the program and explain how it works.

- c) During each lesson the trainer and student keep a log book recording progress, difficulties, accuracy etc.

### *Training*

- Lesson 1. Connect and set up the microphone; start the program; create a new user; try 2 or three simple sentences, correcting the errors afterwards until the sentences are recognised accurately. Use *WordPad* with *Dragon* because it is simpler than writing into another word processor. Lubert & Campbell recommend that the trainer corrects errors *after* each sentence for the first five lessons – not as each word is dictated.
- Lesson 2. Do initial training (65 words); practice loading and saving voice files and work.
- Lesson 3. Do stage 2 training (110 Common commands); dictate a small paragraph for a ‘real’ application (e.g. a poster).
- Lesson 4. Do stage 3 training (230 Dictation words); complete the poster/ exercise from the previous lesson.
- Lesson 5. Do stage 4 training (365 Additional words); introduce text to speech facilities.
- Lesson 6. Show the student how to correct errors themselves using the ‘Oops’ and ‘Choose’ commands.
- Lesson 7. The writer completes a simple question and answer activity using the system, making their own corrections.
- Lesson 8. Teach the use of ‘Add Phrase’ and macros for typing phrases with short commands.
- Lesson 9. Teach correction with Word History. Complete a short reading comprehension exercise.
- Lesson 10. Teach the use of macros.

*Learning to use dictation to write* – Training the writer to speak clearly, and the computer to recognise speech accurately, is only the first stage in training. It is not usually sufficient to leave the student to get on with it once he or she has mastered the technical aspects of dictation. The writer will probably have struggled with writing previously and therefore may not have the necessary skills for planning, organising, editing and formatting work. Therefore, a training programme must be developed to teach these skills. This programme can be similar to the curriculum that is followed when teaching any student to write – whether with pen and paper, word processor or speech recognition system.

## **8 Speech recognition summary**

*Speech recognition systems are useful for pupils with writing difficulties because they:*

- ✓ let the student write without having to type or worry about spelling each word.

*Speech recognition systems are worth investigating if:*

- ✓ the writer has reasonable oral and general literacy skills, but poor spelling;
- ✓ the writer has reasonably consistent speech
- ✓ the writer is highly motivated to learn to use the system;

- ✓ there is adequate technical support and staff time to support the writer during the initial training.

*Things to look for in a good speech recognition system:*

- ✓ a program that works on your machine (some systems need a fairly new and powerful computer);
- ✓ ease of installation, training and use;
- ✓ accuracy and reliability;
- ✓ multiple voice files so that several writers can use it;
- ✓ text-to-speech output of written text;
- ✓ facility to play back (and store) a recording of the dictated speech;
- ✓ correction of errors and mouse and computer control by voice (for writers with physical access difficulties).

## 9 Suppliers

Speech recognition systems can be purchased from most “High Street” suppliers of computer software, e.g. Dixons, Electronics Boutique, and PC World. The suppliers below have specialist expertise in speech recognition systems.

Company Name	Address	Phone	Web address
iANSYST	The White House, 72 Fen Road, Cambridge CB4 1UN.	01223 420101	www.dyslexic.com
Words WorldWide	Ash House, Bell Villas, Ponteland, Newcastle upon Tyne NE20 9BE.	01661 860999	www.keyspell.com

## 10 References

- Donald, E. (1998) *Voicetext: Evaluation of System when used by Pupils with Special Educational Needs*. Perth High School, Perth.
- Litterick, L. (1999) *A comparison of widely available dictation systems*. iANSYST web site, <http://www.dyslexic.com>
- Lubert, J., Campbell, S. (1998) *Speech Recognition for Students with Severe Learning Disabilities from the Trilliom Speech to Text Project: A manual for Implementing Dragon Dictate. Special Needs Opportunities Windows (SNOW)*: <http://snow.utoronto.ca/best/manual2.html>.
- Miles, M., Martin, D. & Owen, J. (1998) *A Pilot Study into the Effects of Using Voice Dictation Software with Secondary Dyslexic Pupils*. Devon LEA.
- Wilson, A. (1998) *Voice recognition Systems*. SCET Sharing Practice Series 17, October 1998. SCET
- Wilson, A. (1998) *Voice recognition Systems*. CALL Information Sheet 15. CALL Centre, Edinburgh.