Chapter 6

Low Cost Voice Recognition Systems for the PC

By Allan Wilson

6.1 What is Voice Recognition?

Voice recognition (VR) systems allow people to control a computer by speaking to it through a microphone, either entering text, or issuing commands to the computer, e.g. to load a particular program, or to print a document.

Voice recognition systems have been around for about twenty years, but until recently they have been very expensive (upwards of £500) and required uncommonly powerful computers to run. However, the three main manufacturers, Dragon, IBM and Kurzweil, have recently released 'cut-down' versions of their programs selling for less than £100 and have also reduced the prices of the 'full' versions. At the same time, the Windows PCs that are now being bought for schools are powerful enough to run the voice recognition software.

The majority of VR systems run on Windows PCs. There is a system (*Power Secretary*) for MacOS computers but it still costs around £300. There is no voice recognition system for Acorn RiscOS machines.

All the low-cost programs reviewed here recognise 'discrete' speech – the speaker must leave a short pause (about 0.25 seconds) between each word. There are also now programs (*Naturally Speaking* and *ViaVoice*) which recognise 'continuous' speech but they are more expensive, more complex to use and require more training and so may be less appropriate for use in schools (at the present time). Of course, technology is moving so fast that it is likely that the continuous speech systems will become much simpler and less expensive by the end of 1998. The programs are all supplied on CD-ROM, with a small instruction card, and head-set microphone. They all claim to be accurate enough to use after a short setting up procedure but effectiveness improves considerably with use or by using special training schemes. Since none are 100% accurate, the user must be able to recognise when the program has made a mistake, and correct it.

To use the low-cost VR systems you need:

- a fairly new and fast Windows computer (a 486/66 is the minimum, although some systems require a Pentium)
- at least 16 MB RAM
- SoundBlaster or compatible card
- a CD-ROM drive

Table 6-1 compares the minimum requirements for each VR program.

Table 6-1 System requirements for Voice Recognition Software

	Price Guide	Operating System	Processor	RAM	Disk Space
DragonDictate Solo	£ 49.99	Win95 / Win 3.1	486 / 66MHz	16 MB	37 MB (+ 13MB)
DragonDictate Solo Pro	£ 99.99	Win95 / Win 3.1	486 / 66MHz	16 MB	30 MB (+ 5MB)
IBM SimplySpeaking	£49.95	Win95	Pentium 100	16 MB	38 MB (+ 2 MB)
IBM SimplySpeaking Gold	£98.70	Win95	Pentium 100	16 MB	75 MB (+ 2 MB)
Kurzweil VoicePad	£69.33	Win95 / Win 3.1	486 / 75MHz	16 MB	20 MB (+ 8 MB)
Kurzweil VoicePlus	£99.90	Win95 / Win 3.1	486 / 75MHz	20 MB	30MB (+ 6 MB)

Notes: These are the minimum recommended requirements for processor and computer memory. Some of the programs will run, but not very well, on a machine with a slightly lower specification, while the programs would all work better on a machine with a higher specification.

The figures for the amount of disk space required represent a full set-up for a single user (+ additional space required for an additional user of the same system).

6.2 Who might use Voice Recognition Systems?

The commercial VR systems were originally designed for lawyers, medical staff and other professionals who want to be able to enter text into a computer at speed without having to learn to type. As the systems have become cheaper and more reliable, they have become useful for all computer users, including people with disabilities.

People with Physical Access Difficulties

People who have difficulty typing due to RSI (Repetitive Strain Injury), arthritis, or muscle weakness are likely to be successful users of VR systems, particularly if they have experience of using computers. The manufacturers claim that you can dictate at 70 to 100 words per minute so VR is potentially extremely fast compared with most other access methods, as well as requiring less physical effort. Writers with more severe physical involvement (perhaps due to cerebral palsy) may also be able to use them, provided their speech is reasonably clear and consistent. The choice of system will depend on the level of disability and the nature of the task. Where a high level of text entry is required and a person can still make limited use of keyboard and mouse for navigation and editing, *SimplySpeaking* would be very suitable. *DragonDictate* or *VoicePad* would be a better choice if a person cannot use a keyboard.

People with Specific Learning Difficulties (Dyslexia)

Pupils with spelling or language difficulties may benefit from using VR, but in practice making effective use of a system takes time and commitment. One advantage of a VR system for pupils with spelling difficulties is that the words are generally correctly spelled because a VR system uses words from its own dictionary. Another is that some people find they can concentrate more on the content and style of language because they no longer have to concentrate on spelling or on using the access device. However, even though the manufacturers claim word recognition accuracy of up to 98% (after practice), this means the program will still get around 12 words wrong per page of text. Therefore, the user must be able to spot any errors and correct them and as a result voice recognition demands considerable technical and language skills. They are not very suitable for people with severe language difficulties.

The method of correcting mis-recognised words varies between programs. With *DragonDictate* and *VoicePad* the likely words are listed after every word spoken so the user has to recognise and choose the desired word from the list or enter it him/herself, leaving opportunities for mistakes to occur. If the user has difficulty identifying the correct word, a text-to-speech program such as *Keystone* can be used to read out the words in the list. With *SimplySpeaking* a few sentences or a paragraph are dictated, then any errors are corrected, allowing the use of a text output program such as *TextHelp!* to read back the words on screen. This speech can then be compared with an actual ‘recording’ of the dictated text taken by *SimplySpeaking* so that incorrect words can be identified and changed. *SimplySpeaking Gold* allows the spoken text ‘recording’ to be saved so that the editing can be done later, possibly by another person.

6.3 What factors can influence success?

Some people can use a VR system and get good results more or less straight away, others need to complete a training procedure and spend many hours using the system, painstakingly correcting mis-recognised words, before a satisfactory level of recognition can be achieved.

Speech Consistency

Consistency of speech is much more important than voice quality. Many people with quite dysarthric speech are able to use voice recognition systems provided that they pronounce words consistently. The training period will generally be longer and it may be necessary to use facilities in *DragonDictate* and *VoicePad* to train words individually to match the speaker's idiosyncratic pronunciation.

Voice recognition systems are generally able to cope with regional and even national variations in pronunciation and accent, although initial results may be disappointing and the training period will again be longer.

Literacy Skills

Given the frequent need to choose the desired word from a list of choices, voice recognition will be most useful for users with reasonably reliable word recognition skills. They are not suitable for non-readers. A program such as *VoicePad*, which allows individual words to be trained easily to the user's voice might be used in conjunction with a reading scheme as an aid to learning new words.

Cognitive Skills

The cognitive load involved in using *DragonDictate* or *VoicePad* can be quite high. The person using the system must not only think about what they want to say, but also how to say it; they must monitor whether the word they have used has been recognised, or has appeared in the list of choices; and must be able to learn the International Communications Alphabet (alpha, bravo, charlie, delta, etc.) – or use the keyboard – to enter a word that has not been recognised. In the case of *DragonDictate* a common problem is remembering to change from CommandMode, used for navigation and control, to DictateMode, before entering text. Dictating text is easier with *SimplySpeaking*, but this must be set against the need to correct mistakes using the keyboard and the limited voice control over the computer.

Visual Skills

VR systems present a large amount of information on the screen: the text that has been dictated; choices for a mis- or un-recognised word; information on how the program is running; even basic information as to whether or not the microphone has been switched on. Despite this, people with a visual impairment can successfully use voice recognition systems and the best results have been obtained with 'full' versions of programs combined with an appropriate screen magnifier or screen reader program, such as *DragonDictate* with *Keystone*. There is an inevitable increase in the cognitive load when using a screen reader because the user has to dictate to the computer and monitor the synthesised speech playback of what the system thinks he/she has just said.

Support

The level of support required will vary from user to user. An experienced computer user who has developed RSI, for example, will require much less support than a young child with multiple disabilities. Nevertheless, even the former would benefit from advice from an experienced user as to whether initially poor recognition levels might be due to a microphone problem, or some other factor.

A young child would need a lot of support, particularly in the training process and it is arguable whether any but the most able child under 11 could manage a VR system, even with good support.

Children with special educational needs will need lots of support to get to grips with the program, especially if they have not used a computer or word processor before.

Motivation

This is probably the most important factor for most people who try to use a voice recognition system. Initial results are often disappointing, particularly in comparison with the manufacturers' claims. There will also be occasions when levels of recognition will seem to drop for no apparent reason (usually a change in the microphone setup, although this will not be obvious). Unless the user is well motivated to use the system, working to overcome the difficulties that arise, it is very easy to give up before the system has had a chance to adapt to the user's voice.

Health Issues

A number of people have reported having voice strain and other problems arising from the use of VR systems. Speaking in a monotone with a short pause between each word for lengthy periods, usually sitting down, without moving, is unnatural and can lead to voice strain. People making extensive use of VR systems are advised to take frequent short breaks, to drink plenty of water and to vary their tone of speech. If a problem does arise, users should seek professional advice from a speech and language therapist.

6.4 Using Voice Recognition Systems

How to Speak

All of the low cost systems described in this chapter use 'discrete' speech, with a short pause between each word. This can encourage users of VR systems to speak in a flat monotone style with a 'dalek' tendency, which may achieve a high level of recognition, but puts an unnecessary strain on the voice. It is quite possible to vary the tone of speech in a natural manner, provided that there is still a pause between words.

Microphone Setup

If the level of recognition suddenly drops, it is most likely that there is a problem with the microphone setup. Other programs which make use of sound will often change volume settings on the PC's sound card which can be carried over into the VR system. If the computer is shared between different users or used for running other programs in the classroom, it is good practice to first check the microphone setup whenever you use your system.

Background Noise

Background noise can be a problem with voice recognition systems, causing words to be mis-recognised, or 'phantom words' to appear in your text. It is best to use the system in a quiet room – a typical classroom is not a good location! The problem can be reduced by setting up the microphone in an environment with a slightly higher than normal background level of noise, e.g. by having music playing gently in the background. The *Talk Mic* microphone (£57.58, from Talking Technologies) is claimed to operate more effectively in noisy environments.

Dictionary Size

When the user dictates a word the program first looks for the word which matches the sound in an *active* vocabulary. The active vocabulary is a set of common words to which extra words can be added to suit the individual user or topic. *DragonDictate* and *VoicePad* also have a *background* vocabulary which comes into play when corrections are made. The dynamic list of words that appear as the user starts to spell an unrecognised word come from the background vocabulary. A word selected from this background vocabulary will normally be added automatically to the active vocabulary during the course of a session. If a word is not in either the active or the background vocabulary – as would be the case with an unusual proper name, for example – it can be added to the vocabulary so that it will be recognised on future occasions. The active dictionary has a limited size: when it is full, new words can still be added and they replace less frequently used words. *SimplySpeaking* and Kurzweil *VoicePad* have *user* dictionaries where new words which are not in the active dictionary can be recorded.

It is possible to get additional dictionaries with vocabularies for legal, medical and other specialised subjects for *DragonDictate* and *SimplySpeaking*. Although there is no medical dictionary for *VoicePad*, Kurzweil have another product, *Clinical Reporting*, specifically designed for use in medicine.

Table 6-2 Comparison of dictionaries in Voice Recognition Programs

	Active	User	Background
DragonDictate Solo	10,000 words		120,000 words
DragonDictate Solo Pro	30,000 words		120,000 words
IBM SimplySpeaking	30,000 words	27,000 words	
IBM SimplySpeaking Gold	30,000 words	27,000 words	
Kurzweil VoicePad	17,000 words	3,000 words	200,000 words
Kurzweil VoicePlus	30,000 words		200,000 words

DragonDictate and VoicePlus do not have a formal user dictionary – if a word has to be added, an unused word will be dropped from the active dictionary to make room for it. SimplySpeaking does not have a background dictionary, making up for this by having a large amount of space available for the user to add words.

6.5 Training

There are two separate issues involved here: training the system to recognise a person's voice, and training a person to make best use of a system.

Training the System

Although it is possible to get reasonable levels of recognition 'straight from the box' with each of the systems, most children and many adults who have not used computers extensively will require a lot of support. To start with, the user chooses male or female voice 'template' – *VoicePad* also has the option of setting up for a child's voice, which would be a useful starting point for use in a school. On-screen tutorials guide the user through the initial process of setting up the microphone and establishing a basic level of recognition. Initially, accuracy of 80% may be achieved (particularly with *SimplySpeaking*) but often the recognition will be less accurate. The user may need help and support to persevere if initial recognition rates are poor. Even 80% accuracy means that one word in five is mis-recognised and must be corrected.

Once the initial setup and tutorial have been completed you can start dictating, correcting errors as you go. As mis-recognised words are corrected, the accuracy improves. Some users find this

satisfactory (particularly with *SimplySpeaking*, which is usually quite accurate after only a short time).

Alternatively, the user can go through a more comprehensive training procedure to train the system to recognise their voice. Training involves speaking several hundred words, phrases or sentences read from the screen into the computer. It is possible to do part of the training process on one occasion and then complete it another time. The vocabulary used in training is not always particularly easy and many words would be unfamiliar to a child. People with dyslexia or other reading difficulties could also have difficulty in identifying some of the requested words. In these cases another person would be needed to 'coach' each individual word, with the microphone switched off, before the user speaks the word into the computer. This will inevitably increase the amount of time required to complete the training process. When all of the words have been trained, the computer will take some time to prepare a voice profile for the user. Depending on the speed of the computer and how close the voice is to one of the standards, this can take up to an hour to complete. Given the time and complexity of the training process, children may be better to just dictate and correct so that the system adapts itself to the vocabulary used by the child.

If there are still problems with recognising particular words after practice and/or training, *DragonDictate* and *VoicePad* both allow individual words to be trained. *VoicePad* handles this task particularly well, providing opportunities for a child to train new words as they are introduced in a reading scheme, for example. A similar technique could be used for training some commands. *DragonDictate* and *VoicePad* commands tend to consist of two or more words compounded together, e.g. *Button-double-click*, which can be difficult to say. It is possible to assign a different word, or sound, that is easier to say to the command.

Training the User

Like any other computer program, there are various commands and operating procedures that the user has to learn. Each of the systems has an on-line tutorial and help facility, that can be used as an aid to learning the system. *DragonDictate* and *SimplySpeaking* provide a help card with a summary of commands, while *VoicePad* provides an on-screen window with a summary of the commands available at any time. It is not necessary to learn all the commands in one go – it is possible to make effective use of any of the systems with knowledge of only a handful of commands.

As well as learning how to operate the program, the writer must learn how to make best use of the system, by monitoring how the system is responding to his/her voice. This can only be done on an individual basis. The user should be prepared to experiment with the volume, tone and rate of speaking until the best level of recognition is achieved. For example, some people have found that the level of recognition with *SimplySpeaking* drops dramatically if they speak too slowly. Short words, such as *a*, *the* and *of* are always more difficult for VR systems to cope with. The person using the system should monitor how well these words are recognised and may have to vary how they are pronounced until the system consistently recognises them.

Use of Language

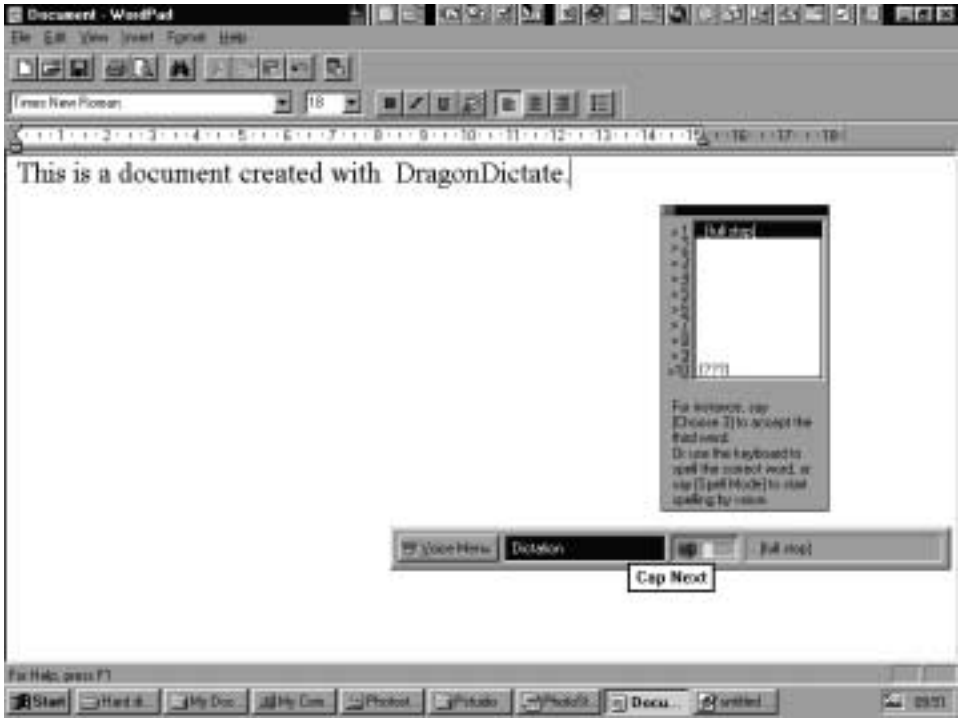
Using dictation requires reasonable sentence building skills which might be a difficulty for users with a language disorder. In any case, learning to dictate *written*, rather than *spoken* language is a skill that has to be learned, particularly by children. Although the user is *speaking* to the system, the outcome is *written* text. Many people using VR systems tend to use longer words than they would if they were typing. Two factors are involved in this: firstly, longer words are easier for the system to recognise (so a person may say "demonstrate", rather than "show"); secondly, people who have difficulties with spelling or typing can experience a considerable sense of liberation because they can produce a long word just as easily as a short word (which they might previously have used as a substitute typing manually).

6.6 Reviews of Voice Recognition Systems

DragonDictate Solo

DragonDictate Solo can be used to dictate into all of the basic Windows accessories, such as *WordPad* and the calculator, and any one from a choice of more advanced programs: *MS Word* (not *Word97*), *Excel, Powerpoint, Access* and *Internet Explorer*. The decision on which program to set up for is taken when the software is installed. Another more expensive program, *DragonDictate Solo Pro Classic* (£99), can be used with a range of programs, although the level of recognition is not quite as good as in *DragonDictate Solo*.

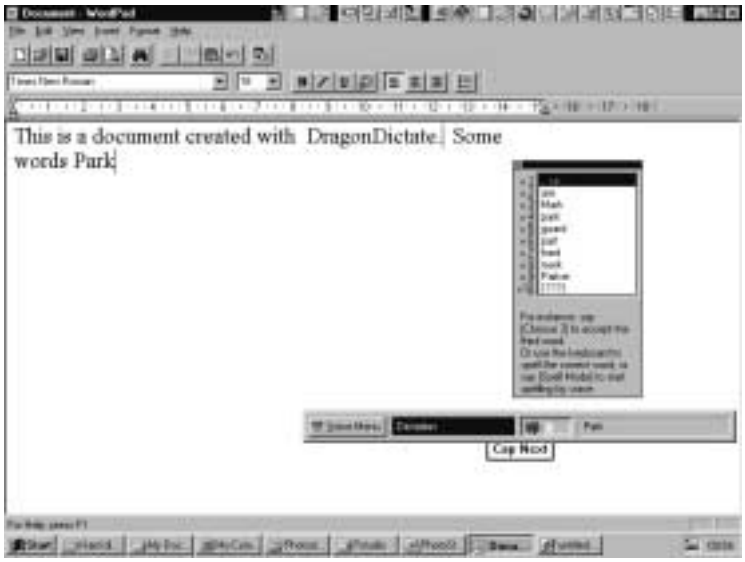
Figure 6-1 *DragonDictate Solo* used in conjunction with the *WordPad* wordprocessor



DragonDictate Solo is the only one of the cut-down programs described here that can provide complete hands-free control over the computer. This has to be balanced against a greater overall complexity and longer initial training than there is in *SimplySpeaking*. The program has two basic operating modes: *Command Mode* and *Dictate Mode*. *Command Mode* provides a comprehensive set of commands to control the computer. It is possible to load and switch between programs, print and save documents. The mouse pointer can be controlled by giving directional commands, e.g. *Mouse-left*, or by telling the pointer to jump to a particular location within a numbered grid which appears on screen. *DictateMode* is used purely for entering text in a document. The user has to be careful to distinguish between these different modes when using the system.

Mis-recognised words are corrected as they are entered. A box containing the likely words to match the sound appears on screen as a word is spoken, with the 'best guess' at the top of the list. If this word is correct, there is no need to do anything – simply carry on and say the next word. If the required word is, say, second in the list, say *Choose 2* or press Alt+2 on the keyboard to select it. If the word does not appear in the list of choices it must be typed either by using the keyboard, or by spelling it aloud using the International Communications Alphabet. The list of choices is dynamic, changing as the first couple of letters are entered, so it is quite possible that the word will appear in the list and can be selected without the need to complete it.

Figure 6-2 Correcting a mis-recognised word in *DragonDictate Solo*



Editing a document by voice is relatively straight forward: there are commands to move the cursor to a particular word or phrase, to delete it, select it, cut it and paste it elsewhere.

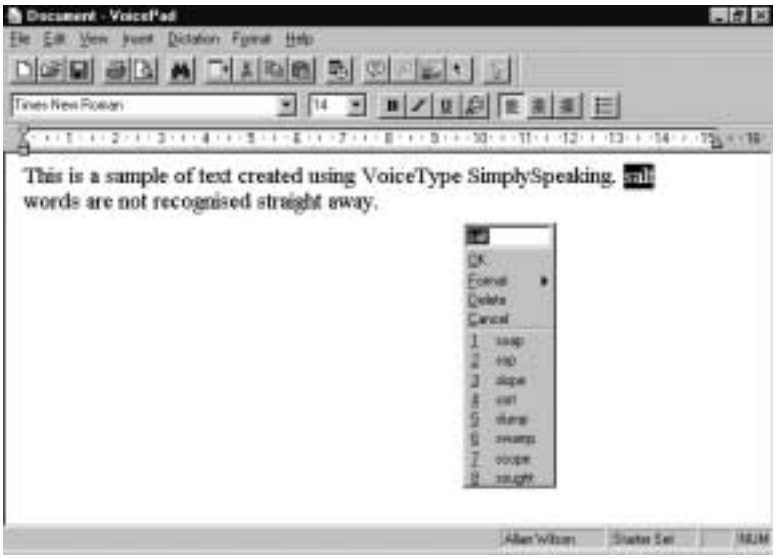
Use of *DragonDictate*, particularly at an early stage, can impose heavy cognitive demands. The user has to think about what he or she is wanting to say, watch the screen to see that the word is in the choices box, remember to choose the word if it is not the first choice and be prepared to spell it if it is not on screen. Given time, this will become less of a problem as the level of recognition will improve to an extent that the user will not need to check that words have been correctly identified, unless it is an unusual word that may not be in *DragonDictate*'s dictionary.

IBM VoiceType *SimplySpeaking*

SimplySpeaking is probably the simplest VR system to use, and seems to achieve the best level of recognition 'straight from the box'. Unlike the other systems described here, which work purely on the sound of an individual word, *SimplySpeaking* will try to recognise individual words on the basis of context. This leads to a high level of recognition, but can be disconcerting on screen as the 'first guess' at a word may change rapidly on screen as further words are added.

Mis-recognised words are not be corrected by voice as they are entered. Instead, a couple of paragraphs of text are dictated and then any mis-recognised words are corrected by clicking on them with the mouse and selecting an alternative word from a list that appears on-screen, or by typing the correct word. Unfortunately, the list of words is static, i.e. it does not change as the user starts to type, so that, rather than being able to choose the desired word from the list after typing only a couple of letters, the user must be able to spell the word correctly and in full. It is vital to correct errors otherwise *SimplySpeaking* will not learn the user's voice. A useful feature of *SimplySpeaking* is the ability to play back a recording of the actual word and sentence that were dictated as a reminder of what was said. People who have difficulty in reading text from the screen may wish to use *textHelp!, TextAssist*, or some other speech output utility to read back the text from the screen as a check that it matches what was intended. When the user has finished dictating and editing the text it can be saved as a *Word 6* file, or copied into another program.

Figure 6-3 Text correction within VoiceType *SimplySpeaking*



Text correction is usually carried out after a passage of text as been entered. When an incorrect word is selected, a box with possible alternatives appears on screen. If the desired word is in the list, then it can be selected with the mouse. If it is not in the list, then it has to be typed in from the keyboard.

It is possible to dictate at speeds of between 70 and 90 words a minute – faster than most people can compose text – but the slow speed for correcting mis-recognised words must be set against this.

SimplySpeaking can be used only for dictating text, making it useful for creating new documents, but unsuitable for editing existing documents. There are no facilities for speaking commands, e.g. to save a document, or for moving the mouse on-screen. This makes it difficult for people with a severe physical disability to use unless they have an alternative method of accessing the computer.

IBM VoiceType *SimplySpeaking Gold*

This is an expanded version of *SimplySpeaking* providing options to enter text directly into any application (including *Word 97*), correct and edit text by voice and use commands to navigate through programs. Two extra programs provide these facilities – *VoiceType Direct* and *VoiceType Navigator* – together with new Correct and Spell modes (the latter using the International Communications Alphabet) within the basic *VoicePad* word processor used in *SimplySpeaking*. The addition of these features can make the program more complex to use than *SimplySpeaking*, and it also needs a slightly more powerful computer to run well. For some users it may be sensible to concentrate on the basic dictation features of *SimplySpeaking* to build up a good voice profile, before moving on to the other facilities.

There is a facility to save the recorded voice file, as well as the text, making it easier to edit text later to ensure that it matches the original dictation. This facility would be particularly useful if the editing was being done by somebody else.

Figure 6-4 Choosing a program to run from the *VoiceType Navigator*

Provided that the program has been put in the right place and its name appears in the list of choices, it is possible to run most programs by using the VoiceType Navigator in *SimplySpeaking Gold.*

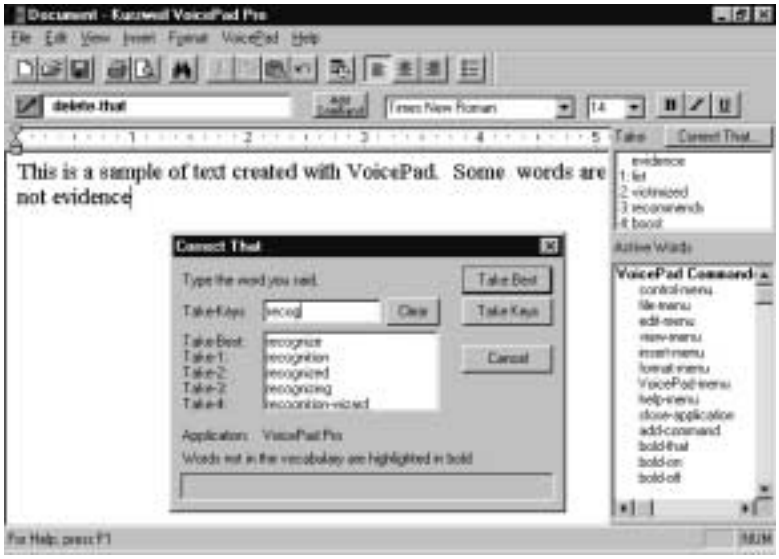
Kurzweil *VoicePad*

VoicePad uses a similar simple word processor to that of *SimplySpeaking* for entering text, but has the advantage of allowing voice control over menu items and over the cursor. Text entry is very similar to *DragonDictate*, with a box of choices appearing as words are spoken, but the commands are different, e.g. *Take 2*, instead of *Choose 2*. It uses the International Communications Alphabet to spell words in the same way as *DragonDictate*, but words are more likely to appear in the list of choices as *VoicePad* uses a much larger dictionary. The dictionary uses American spellings and pronunciations, which can cause problems for people in the UK. Provided that the system is used carefully by a person who is aware of these differences, the problem will disappear over time as UK spellings and pronunciations replace the American equivalents when a word is used and corrected. One of the main UK suppliers, Talking Technologies have produced a package, *TalkFast*, which provides a UK English command list, an improved microphone and other helpful features to accompany *VoicePad*.

Figure 6-5 Kurzweil *VoicePad*

With *VoicePad*, the list of current choices and 'Active Words' (commands and words available from the dictionary) remain fixed on screen.

Figure 6-6 Correcting a mis-recognised word in *VoicePad*



Mis-recognised words are corrected in a similar manner to *DragonDictate*, with the desired word being typed letter by letter until it appears in the list of choices. Access to a bigger dictionary means that words are more likely to appear in the list of choices offered by *VoicePad*, but the use of American spellings can be distracting.

Kurzweil *VoicePlus*

This is derived from the original *Kurzweil Voice* program on which *VoicePad* is based. In addition to the *VoicePad* features, it also includes navigation between and within programs and the ability to type into any application. There is immediate access to a dictionary of 30,000 words, increasing the likelihood of any particular word being recognised. Another useful feature is the ability to use 'macros' – the ability to use a single command to enter a piece of commonly used text such as an address.

6.7 Comparison of features

Table 6-3 Comparison of features of Voice Recognition Systems

	Text Entry	Menu Control	Navigation	Macros	Playback
DragonDictate Solo	Yes	Yes	Yes	Yes	No
DragonDictate Solo Pro	Yes	Yes	Yes	Yes	No
IBM SimplySpeaking	Yes	No	No	No	Yes
IBM SimplySpeaking Gold	Yes	Yes	Limited	Yes	Yes
Kurzweil VoicePad	Yes	Yes	Limited	Yes	No
Kurzweil VoicePlus	Yes	Yes	Limited	Yes	No

Menu control refers to the use of voice commands to select the commands that appear in the menu bar of the program, e.g. 'Print', 'Save'.

Navigation refers to the ability to run different programs on the desk top, copy, delete files, etc. Kurzweil VoicePlus allows the user to 'speak' anything that appears on screen, which would be enough for most purposes, but only DragonDictate allows full mouse control by voice.

Macros are used to enter a block of text or a series of commands with one voice command.

Playback is a feature of SimplySpeaking to record the words spoken by the user and play them back when corrections are being made.

Table 6-4 Advantages and disadvantages of different Voice Recognition Systems

	Advantages	Disadvantages
DragonDictate Solo	• Complete hands-free control over the computer, including the mouse	• Relatively small dictionary • Can appear complex in the early stages • Can be used with only 1 main application
DragonDictate Solo Pro	• Can be used with any application	• Uses 'old' voice recognition system
SimplySpeaking	• Easy to use • Recognition very good	• No facility for navigation / commands • Can be used only with its own word processor • Poor for editing text
SimplySpeaking Gold	• Can enter text in any application	• No mouse control.
VoicePad	• Big dictionary • Good for training individual words and commands • Can be set up to use a 'child's voice'	• American dictionary • No mouse control • Can be used only with its own word processor
VoicePlus	• Can be used with any application	• No mouse control

Continuous Speech

IBM and *DragonDictate* have now released voice recognition systems which offer continuous speech recognition, without the need to pause between words. IBM's *ViaVoice* (£99) lets the user dictate directly into its own word processor in a similar way to *SimplySpeaking*, and can also be used to dictate directly into Word 7 and 97. Text can be copied from these programs into another application. The minimum hardware specification is a Pentium 150 MMX Processor, with 32MB of RAM and 120 MB of hard disk space. More than one user can share the same system.

Naturally Speaking (£199) from *DragonDictate* also uses its own word processor, relying on 'cut and paste' to move text into other applications. It requires a Pentium 133 processor and 32 MB of RAM to run, but works better with a more powerful processor, and 60 MB of hard disk space. The basic version of the program does not allow the system to be used by more than one person. Although the program requires the use of hands for editing, mouse control, etc., it can be used in conjunction with the standard *DragonDictate* for 'hands-free' use.

The training process for each system is arduous, requiring the user to speak long passages of 'standard' text. The effort to achieve continuous speech recognition has to some extent been at the expense of tolerance for 'non-standard speech'. When the lack of navigation commands in the two basic systems is also taken into consideration, the systems have many disadvantages for people with disabilities, compared to the 'discrete' systems.

6.8 Further Information

Web Sites

Because of the speed of change with regard to the pricing and capabilities of voice recognition systems, the best sources for up to date information are on the world wide web. Here are some of the sites that are worth visiting:

General Sites

Comp.Speech FAQ (http://squid.eng.cam.ac.uk/comp.speech/index.html) This is a site with answers to frequently asked questions regarding all aspects of computers and speech. *Chapter 6* is devoted to voice recognition. Some of the information is quite technical, but the site contains a comprehensive guide to the many different packages available, with links to manufacturers' web sites, and a lot of useful background material, including material on health issues.

Typing Injury FAQ: Speech Recognition (http://www.cs.princeton.edu/~dwallach/tifaq/speech.html) Detailed reviews of most of the systems available and lots of links to other sites. People should be able to track down any information they want from this, or the Comp.Speech FAQ site.

National Centre to Improve Practice in Special Education (http://www.edc.org/FSC/NCIP/). This is an American educational site with resources on the use of technology to support children with disabilities. The site includes an extensive section on the use of voice recognition systems in schools.

Suppliers

Iansyst Ltd. (http://www.dyslexic.com/) UK supplier for the leading packages. The site is regularly updated with comparative reviews and the latest pricing information.

Endeavour Technologies (http://www.endeavour.co.uk/). Leading UK supplier for *DragonDictate*.

Talking Technologies International Ltd. (http://www.talk-systems.com/). UK supplier for *Kurzveil VoicePad*. The site has information about various 'add-ons' to make the software easier to use in the UK.

IBM Software (http://www.software.ibm.com/is/voicetype/uk_menu.html). Information on *SimplySpeaking, ViaVoice* and other IBM products.

VoiceUsers Mailing List (http://voicerecognition.com/voice-users/). Details of a discussion group for people using VR systems can be obtained from this site. The archive of previous discussions can be found at http://www-uk.hpl.hp.com/people/ange/archives/archives-97/voice-users-archive/index.html

Articles

Personal Computer World and other computer magazines regularly feature articles on voice recognition systems, for example:

Cole, George (1997) *Speak and Spell* **Personal Computing World** November 1997, pp 153 – 156.

Gann, Roger (1997) *DragonDictate Solo vs VoiceType SimplySpeaking* **Personal Computing World** July 1997, p 93.

Kelway, Peter (1996) *Speaking to Text* **Interactive** June 1996 pp 34 – 37.

Kotler, Ava-Lee & Thomas-Stonell, Nancy (1997) Effects of Speech Training on the Accuracy of Speech Recognition for an Individual with a Speech Impairment **AAC** Vol. 13 (2) pp 71 – 80.

Rahamim, Lesley (1997) *Talking with Computers* **Special!** Autumn 1997 pp 57 – 59.

Tam, Cynthia & Kotler, Ava-Lee (1995) *Assessment Criteria for Using Voice Recognition Systems as Writing Aids* **Closing the Gap** December 1995 pp 14 – 17.

Coleman, Colette & Meyers, Lawrence S. (1991) Computer Recognition of the Speech of Adults with Cerebral Palsy and Dysarthria **AAC** Vol. 7 (1) pp 34 – 42.

(1997) *Voice Recognition Systems for Windows - Frequently Asked Questions* CENMAC Information Paper 1997. (from CENMAC, Eltham Green Complex, 1a Middle Park Avenue, London SE9 5HL. Tel: 0181 850 9229).

(1997) Using Low Cost Speech Recognition Systems with Children CENMAC Information Paper 1997.

Suppliers

Systems are readily obtainable in most computer shops, e.g. Byte, Dixons, PC World. Dealers with particular expertise who will be more likely to be able to answer questions and provide basic support and training, if required, include:

Aptech Limited, Aptech House, Meadowfield, Ponteland, Newcastle-upon-Tyne NE20 9SD Tel: 01661 860999 Fax: 01661 822777

Endeavour Technologies, Colette House, 234 Station Road, Surrey KT 15 2PH Tel: 01932 827 324

Fax: 01932 858 200

IANSYST Ltd, The White House, 72 Fen Road, Cambridge, CB4 1UN, UK. Tel: 01223 420101 Fax: 01223 426644.

Talking Technologies Ltd., 34A Glazbury Road, London W14 9AS Tel: 0171 602 4107 Fax: 0171 603 2109.