Voice - to - Text

A Reality

By Colin Mitty

(Colin Mitty is corporate counsel for Osborne Computers Australia.)

As a lawyer who has dictated to a secretary or tape machine for the whole of my career, my keyboard skills are far from adequate but the text which you read in this article has been "typed" by me without the assistance of a word processing operator and very little use of the keyboard.

New Software

Since late last year I have had the privilege of using a new software product developed by IBM Speech Recognition Laboratories in Boca Raton, Florida. I am informed that the software has been in the course of development for more than 20 years. It is only now that personal computers which have the power to successfully run the software have become readily available.

The software is truly exciting. I am able to speak to my PC almost normally, and quite quickly (more than 70 words per minute) and the system can determine what it is that I am saying and convert those sounds to appropriate text, almost as quickly as I say the words. Perhaps the most exciting aspect of this program is that the system is able to assess each word in context of the three preceding and the three following words and to make a "logical decision" as to the appropriate word and, in the case of like-sounding words, the appropriate spelling. To demonstrate this latter ability of the system I will dictate the following sentence:

"Please *write to* Mr. James *Wright right* away and tell him that *our two* o'clock appointment is *one hour too* early."

Some of the basic aspects of the program are also impressive. For example, when I dictate the words "full stop" as one word, the system inserts a full stop, 2 spaces and starts the following word with a capital letter.

Corrections

Occasionally, the system will make an error in recognising a word, which must be corrected. In such a case, I am able to highlight the word, using either my voice or the mouse, following which the system will present a list of alternative words which it believes might be the correct word. If the word I require is not shown in the list, I can either say the word again or type it using the keyboard. Therefore, the editing process is very simple and correcting a document is really no harder than checking a draft document which has been typed by another person.

Easy Shortcuts

I am also able to create voice activated "macros" to facilitate frequently used paragraphs, addresses, or phrases. For example, the words "sign off", when said together, will produce the following:

Yours faithfully,

Colin Mitty
Corporate Counsel

One further interesting aspect of the program is that it incorporates voice-activated "navigation" of the computer's applications. For example, using only voice commands, I am able to open the word processing program, open a particular file and paste the text which I have dictated into a specific part of that file. One way in which I use this facility is to paste dictated text into a file which is effectively a letterhead. After pasting the appropriate text, I can activate the spell checker and print the document, again without touching the mouse or the keyboard.

Use with other Programs

I am also able to train the system to perform tasks which would normally achieved with key strokes in programs other than word processing. I use an electronic diary and address book program on my PC and I have been able to create voice operated commands to move to different parts of that diary.

Large Vocabulary

No less remarkable is the way in which the system is able to recognise more than 20,000 words, whereas I read aloud only about 250 sentences to enable the system to recognise my voice. After reading those sentences, the program takes many hours to process all of the sounds which it has "heard" and become able to recognise the large vocabulary of words which reside within the program. In addition to the resident vocabulary, I am able to create an expanded vocabulary of particular words which I use frequently. These particular words are most frequently the names of places or people which cannot be in the standard program, for obvious reasons. One of very interesting aspects of the program is that it will use its best endeavours to interpret a sound and produce what the program determines is the most likely word. With names of people, it will produce some interesting interpretations. The program's first attempt at my name as "column 80".

Reduce Overheads and Save Time

The benefits of the system, I believe, will be substantial. As the cost of a PC, software, and microphone from Osborne will approximately $10,000, which includes a very high level computer, it will be easy to justify the purchase of a personal dictation system which could be used by a number of authors. The capital cost of the system will be far less than the annual cost of a word processing operator and the long term economy of such a system is obvious. At the same

Voice - to - Text

A Reality

By Colin Mitty

time, one of the most significant benefits which I achieve by using the system is that a document can be created by me and become immediately available to me without my relying on a favourable queue in the word processing department. I can create confidential reports and have only the recipients see them. I can attend to short personal correspondence and office memos almost instantly. If I attend the office outside usual hours, I can readily create and print any urgently-needed documents. This remarkable system will be a necessity for any organisation which fears claims for "RSI" and for authors who do not wish to employ a word processing operator for economic reasons.

Hardware

The software will not run properly on just any personal computer. The minimum specifications are a 16 megabytes of RAM, a minimum of 62 megabytes of available disk space, a 25 MHz 486 SX or faster processor, an external speaker and microphone. Osborne will be supplying the software, adaptor card, and microphone as integral parts of the package. The software and the manual will be supplied on compact disk which will mean that a CD drive will be an important component of an appropriate PC. Any existing personal computers will be capable of upgrade to the required specificiations. In addition, the software capable of upgrade to the required specifications. In addition, the software presently requires OS/2 operating system and word processing package that is compatible with the voice recognition software. Packages which are presently compatible include OS/2 System Editor, Lotus Notes for OS/2, Lotus Ami Pro, Microsoft Word for OS/2 and Word Perfect 5.2 for either OS/2 or Windows. I have been using Wordperfect for OS/2 and

have found the system generally very easy to use.

New Approach to Personal Computers

The voice recognition software has the capability of radically changing the way in which people deal with personal computers. The ability to enter text without having to develop adequate keyboard skills is a remarkable development in the user-friendliness of the personal computer. The ability to navigate around a program or between programs, using voice controls, is equally impressive. I should add that it is also possible to create custom commands to enable computer games to be played without using the mouse or the keyboard.

Additional Benefits

The program also provides a startlingly effective tool for people who for whatever reason are unable to use a computer keyboard. I am presently studying for an MBA part time at Macquarie University and I am required to submit typed assignments at reasonably frequent intervals. I have in the past imposed on the services of my wife to type these assignments or I have struggled away in the small hours of the night "hunting and pecking". I am now able to talk an essay into my PC, make any necessary changes and print a final version, all without difficutly. People with physical disabilities will find the system enormously helpful. With the use of a head-set microphone, dictation can proceed "hands free" and any documents which I want to refer to can be handled easily.

Specialised Versions

The version I use is designed for general officer correspondence. In some industries, this version will not have a large enough vocabulary and

special versions are presently available for radiology, medical and journalism. For these, the vocabulary is substantially expanded with specialised words. It is anticipated that a legal version will be available in due course, although I am sure that the general office verison has sufficient vocabulary for most legal dictation.

Mutliple Users

More than one person can use the same PC, as the system will allow numerous authors to train the system to recognise their voice. To use the PC, it is simply a matter of initialising your personal voice recognition file by clicking an icon and you are ready to start dictation.

It is easy to see that we are close to the point where we will be able to operate our front door lock, lights, television and stereo system by voice commands. In the meantime I will be more than happy to talk to my PC and have it happily produce the text I desire.