How I Use It: Voice recognition software

Speech recognition software appeared in about 1952 but the first widely available example was the system developed by IBM and incorporated, in 1996, into its personal computer operating system OS/2 Warp. The following year, Dragon Systems released its first version of NaturallySpeaking for Microsoft Windows. At about the same time, IBM released a Windows version of its recognition software, ViaVoice. Competition in the commercial field disappeared in 2003 when ScanSoft, now Nuance, obtained the rights to produce and develop both systems. Since then several versions of NaturallySpeaking have been released, the most recent being version 11.5 in June 2011 and running on a range of Windows systems from Windows XP to Windows 7. There are programs available for Mac computers, with the market leader being a version of Dragon.

What does it do?

On personal computers, voice recognition software has two roles: to convert speech to text and to control the computer. For most doctors, the former will be its principal use but for many users, particularly those with physical disability, having hands-free control of a computer system is a considerable benefit. I first tried IBM’s offering in 1996, but even though OS/2 was greatly superior to Windows in its use of the computer hardware then available, the software was too slow and too prone to error to make it a useful tool.

For a few years, I worked for a national occupational health provider which used a central typing pool to transcribe dictation tapes from peripheral clinics. I was dismayed by the quality of transcription and the careless checking of reports before they went out to customers. Tightly constrained turnaround times prevented me from checking my letters. Some of my reports were badly distorted, so when the company agreed to a trial of dictation software I was very happy to be a guinea pig.

This experiment coincided with the introduction of NaturallySpeaking version 7. This proved to be an effective speech-to-text system that was of practical use in my everyday work. Subsequent upgrades have each introduced a substantial improvement in the speed and accuracy of dictation, although the improvement has to some degree depended on a parallel improvement in the processing capacity of personal computers.

Voice control

There are two aspects of computer control by voice. Within a program, it can be useful to issue commands so that activities such as opening and saving files, and moving through and selecting text, require no keyboard input. NaturallySpeaking is pre-configured to provide this facility within commonly used programs such as Microsoft Word. Using these commands quickly becomes second nature.

It is also possible to control the computer by voice commands that replace mouse and function key inputs to the operating system itself. This feature requires time and effort to learn and is only suggested for those who are unable to use the keyboard.

Dictation

I now use NaturallySpeaking as my normal input method when writing text. This includes letters, reports and this article. The system is not perfect and for me it is not a complete substitute for keyboard entry so I use whichever happens to be more convenient. It is, of course, essential to proofread reports as carefully as if they had been prepared by a typist. Because I use Dragon NaturallySpeaking, I can produce a text file that I know is an accurate transcription of my intended meaning, ready to go to the client and usually ready to be e-mailed or posted less than 24 h after I have seen the patient. A suitable working environment and a good-quality microphone are crucial to success. Background noise and reverberation from hard surfaces can degrade the speech signal and cause multiple recognition errors, especially with an omnidirectional microphone. The headset microphone may not be practical during clinics and hand-held microphones are a nuisance when leafing through a set of notes. I use a desk microphone, the Samson Go Mic, which meets essential criteria: highly directional discrimination, good signal quality and sufficiently robust to travel around in a laptop bag.

Summary

The current version of Dragon NaturallySpeaking certainly does the job. For less than £150, I have software and a high-quality microphone that quickly convert my thoughts to text while allowing me to avoid the expense of using a typing service.

Andrew Graham-Cumming

e-mail: andy@gableside.org.uk