

Plugged In

Chatting . . . with . . . a . . . PC

Software that lets you dictate to your computer needs to boost accuracy / **Simson L. Garfinkel**

CAPT. JAMES T. KIRK could speak to the computer of the Starship Enterprise. David Bowman could talk to the HAL 9000 in Stanley Kubrick's movie "2001: A Space Odyssey." So how come I can't talk to my personal computer and have it understand what I'm saying?

Actually, I can.

At least three companies, two in the Boston area, sell software packages that let you control a conventional Windows-based PC using merely the sound of your voice. Unfortunately, none of the systems works as well as the computers in the movies.

Although all of the systems claim to understand tens of thousands of words, giving you the ability to start programs, choose menu commands and even dictate a memo into your word processor, using the systems in practice is far from ideal.

For the computer to understand what is being said, the speaker must pause for a split-second between each word, turning his or her voice into a machine-like staccato, a parody of human speech. It . . . also . . . helps . . . to . . . speak . . . in . . . a . . . monotone.

Despite these limitations, all of the systems seem to have a grow-

ing following. Kurzweil Applied Intelligence Inc. is strongest in the medical sciences: The Waltham company claims 10 percent of hospital emergency rooms in the country use its Kurzweil Clinical Reporter system for the dictation of radiology reports. This professional-grade system is priced between \$5,000 and \$8,000 per user. Kurzweil also makes a general-purpose system called VOICE 2.0, priced at \$695. A medical vocabulary and language module is available for \$295.

For people curious about voice-recognition technology, there's a free demo on the company's Web page — <http://www.kurz-ai.com> — that you can download and try for yourself.

Dragon Systems, meanwhile, has gained a good reputation among students and writers for its flexible general-purpose system that can handle a variety of tasks. The Newton company sells a Personal Edition with a 10,000-word vocabulary for \$395, a 30,000-word Classic Edition for \$695 and a 60,000-word Power Edition for \$1,695. Dragon also sells language modules containing 60,000 words specialized for medicine, law, business, the computer industry and even journalists. These modules cost \$495 each. It's Web site can be found at <http://www.dragonsys.com>.

International Business Machines Corp. (<http://www.ibm.com>) is also making a push into the field, basing its own voice dictation product on research out of its Yorktown Heights laboratory in New York. The company's Voice-Type Dictation 3.0 has a retail cost of \$695.

These systems are actually more similar than different. They can run on standard Windows-based PCs equipped with a 90 MHz (or better) Pentium microprocessor, a 16-bit sound input system and at least 16 megabytes of RAM. This means that you can run them on a laptop — a big improvement over a few years ago, when both Dragon and Kurzweil required a special-purpose board to be plugged into your PC.

All of the systems allow a speaker to say words and have them entered into a program such as a word processor or a spreadsheet. And the programs have a rudimentary understanding of grammar, so that words at the beginning of sentences are capitalized, and so that the program can make an educated guess among words like "two," "to" and "too."

Of course, none of these programs is perfect, so they all display a small window in the corner of your screen showing their top 10 guesses — if they make a mistake, which they often do, you can correct them.

Ironically, one of the hardest parts of developing these systems isn't getting the voice recognition to work, but developing a good user interface and links to the me-

SURF SITES FOR THE INTERNET EXPLORER

HARD NEWS

<http://wsj.com>

Wall Street Journal Interactive Edition

<http://www.cnet.com/Content/News/>

c/net

<http://www.cnnfn.com/>

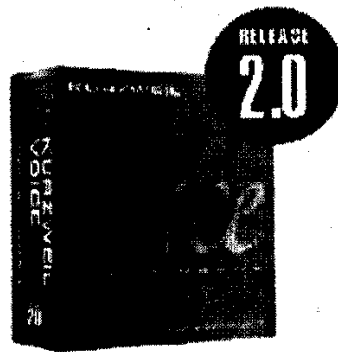
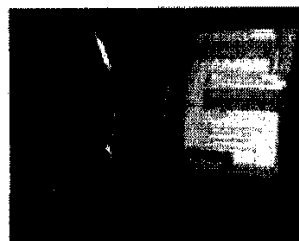
CNNfn

<http://www.usa.ft.com/>

Financial Times of London

<http://www.zdnet.com>

ZDNet



chanics of the Windows operating system. That might soon change, thanks to new features Microsoft is adding to Windows.

A system's accuracy depends, to a large extent, on the speaker.

"You can find a very small number of people who can benefit right away," says David Nahamoo, senior manager of the human language technologies department at IBM's Yorktown Heights laboratory. It also helps, says Nahamoo, "if you are a very controlled speaker, and you use your system in a very quiet environment."

The biggest benefit, says Nahamoo, is for people who are slow typists or who have a disability that affects their ability to type. With practice, speakers can achieve typing rates of 100 words per minute and 95 percent accuracy or better.

Practice is key, as I discovered during a week of using both the Kurzweil and the Dragon systems. Although all of today's speech-recognition systems are speaker-independent, which means that they work out of the box with an accent-free English speaker, the systems learn how you talk as you use them, so their accuracy gener-

ally increases. And, of course, the more you use one of these systems, the more you learn how to enunciate your words so that they will be easier for the system to understand.

Nevertheless, all of today's systems fall short of the science-fiction promise of computers that can understand what we say. The reason, says Janet Baker, chairwoman and chief executive of Dragon Systems, is that trying to recognize a continuous stream of spoken words is dramatically harder than trying to recognize words spoken one at a time.

Back in 1989, both Baker and Kurzweil chief executive Ray Kurzweil told me that all that was necessary for a commercial continuous speech recognition system was a tenfold improvement in the speed of computers. Well, computers are 10 times faster today than they were in 1989, but we still don't have such a system.

Free-lance technology writer Simson Garfinkel can be reached at simsong@acm.org. Plugged In columnist Hiawatha Bray has the day off.