

## Tuesday Worklife

# 'YO, computer!

More and more machines can understand English. It's a technology people are talking about.



The Philadelphia Inquirer / MICHAEL S. WIRTZ

**Dr. Charles Reese of the Medical Center of Delaware can quickly update patient records by speaking into a computer.**

By Gayle Sims  
INQUIRER STAFF WRITER

**T**his ... story ... was ... written ... by ... talking ... into ... a ... computer.

The task was slow and cumbersome, and sometimes the computer didn't hear right and had to be told how to fix the mistakes.

But it worked. And soon, it will work a lot better.

A new generation of high-powered personal computers, expected out by the end of this year, will come equipped with sophisticated voice-activated technology, giving you the choice of typing on your computer or doing it all by voice.

By speaking into a microphone hooked up to a computer, you'll be able to "voice-type" 70 words per minute, the manufacturers predict, with 95 percent accuracy. The computer will recognize about 32,000 different spoken words and has a

100,000-word dictionary. The price initially will be about \$3,000 — about what a new Pentium chip computer costs today — but is expected to drop quickly.

The development will mean a quantum leap for the computer age. It will turn computers that respond only to a complex series of typed commands into assistants that you can, literally, boss around. "We tell our assistants, 'Fax to John,' and the task will be completed," said Kai-Fu Lee, technical director for voice-activated systems for Apple, one of several companies developing voice products.

"With speech recognition and sophisticated new software," Lee said, "it will no longer be necessary to type or use a mouse to open the document, find the phone number, type the number and then run the fax program. Instead, the computer will carry out a simple voice command."

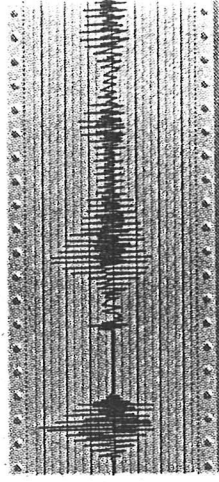
While a primitive version of this technology, involving only limited commands,

## Understanding Us

Advances in voice-recognition technology are bringing us closer to the day when we will be able to communicate with our computers by talking to them as naturally as we talk to each other. Much of the available technology requires a person to pause awkwardly between each word. This helps the computer to differentiate between the end of one word and the beginning of another. However, breakthroughs have been made in the area of continuous-speech recognition, which promises to let people speak naturally when using their computers. Here's how it works:



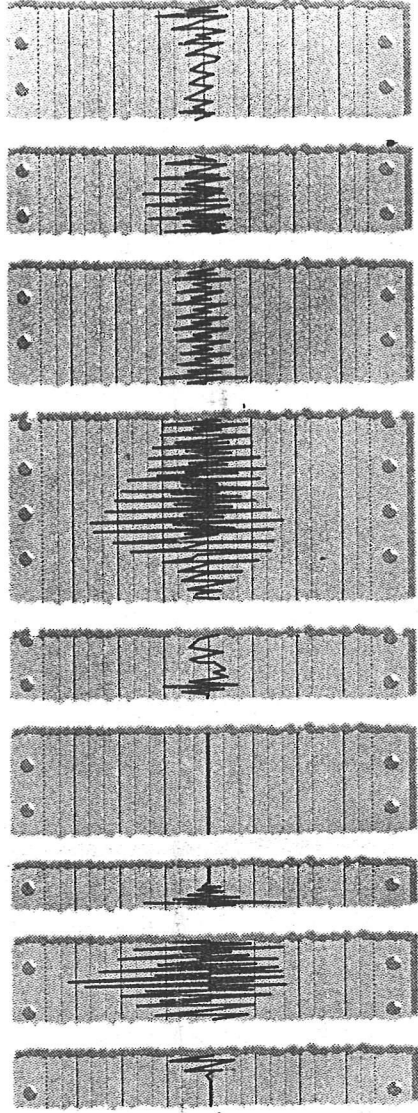
**1** A person speaks into a microphone attached to a computer.



**2** The computer converts the speech into a waveform.

**3** As the speech is converted, each part of the waveform is matched with the waves from a set of basic sounds, called phonemes. Phonemes are the sound-building blocks of a language. There are about 48 phonemes in the English language.

w a a t d t t a y m i h z



**4** Next the computer matches the phonemes with words that exist in the English language. The computer then evaluates which combination of the words makes the most sense, using grammar rules and information about what things people are more likely to say.

Guess: **WHAT I MISS IT.** Evaluation: Words possible but sentence doesn't make sense.

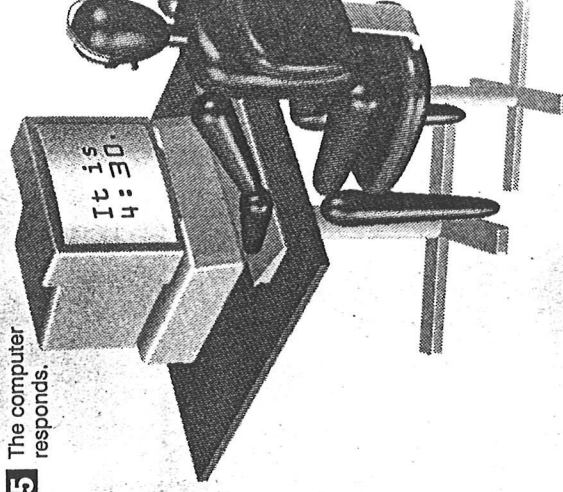
Guess: **What tie Ms. it.** Evaluation: Same as above.

Guess: **What thyme is it?** Evaluation: Sentence is possible, but not very likely.

Guess: **What time is it?** Evaluation: Words and sentence are possible, and the phrase is common. This is the most likely candidate.

SOURCE: Arthur McInair, research programmer, Carnegie-Mellon University

**5** The computer responds.



The Philadelphia Inquirer / ARI

# Computers understand English

uses a voice-activated computer made by Dragon Systems. The technology will improve, she knows, but she's already in love with "Dragon."

"What is sort of lovely about the Dragon is that, first of all, writing is such an isolating thing, and having RSI is such an isolating thing, and then you come to this Dragon and suddenly you have this companion who writes with you. And I feel very bonded to this spirit. I just love him."

Although she knows it's just a coincidence, the machine's mistakes often come out sounding sassy.

"He says funny things all the time and now he has started to become my unconscious and has started to comment as I write."

"I was giving a demonstration, saying, 'Now the Dragon will show off what he can do. As... you... can... see... he... is... very... smart... and... also... very... dumb,'" she said into the microphone.

"When I said *dumb*, Dragon typed *not*."

It was the kind of eerily appropriate mistake that leads some people to think of their computers as human.

"The other day I wanted to say the word *teensy*," Winer said. "He did not want to say that word. It was perhaps too undignified for him. So I said, 'Begin spell mode: Tango. Echo. Echo. November. Sierra.' When I got to Sierra, he [displayed] *teensy*, but he also offered *teensy-weensy*. I wanted to pinch his little computer face."

For now, the technology is hard work. "The first review that I wrote [using Dragon] took me 2½ days. This usually would have taken me a half-day," said Winer, one of four reporters at Newsday now using voice-activation.

"It takes a different part of your brain than it would to express yourself through your fingertips.... I get tired really fast and then I start to slur my words and he starts to mishear me and we both get crabby and have to take a nap."

■ This year, consumers have gotten a sampling of some of the simple tasks that voice-activation can already do. There is Sprint's new voice phone card, boosted by Candice Bergen, that lets you dial a number by simply saying, "Call home."

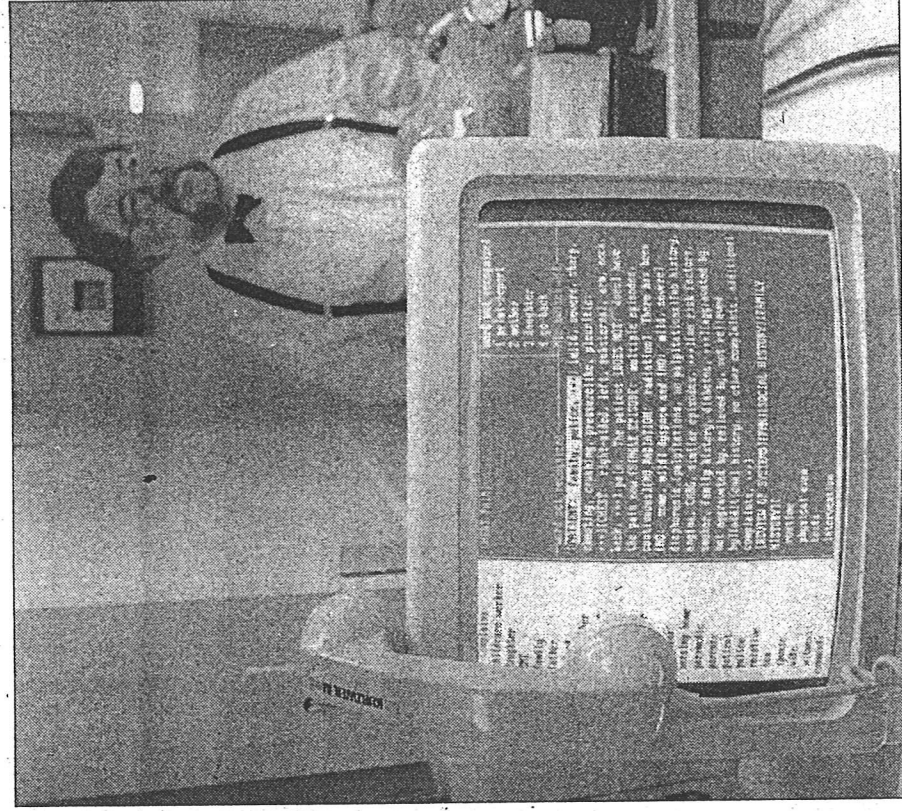
And there's the device that lets you program your VCR by telling it to "Record Six. 10 p.m." The commercial, with Monty Python's John Cleese, uses a process designed by Voice Powered Technology, the California company that last fall came out with the Voice Organizer, a voice-driven date and phone book.

Then, in some parts of the country, automated teller machines are recognizing people's voices rather than their secret numbers.

Voice-activated cellular phones are now an option for cars built by Rolls-Royce and Lincoln. And Sanyo is testing a car stereo that your voice can tell to turn on and change channels. Besides being fun, the products are being touted as safety features.

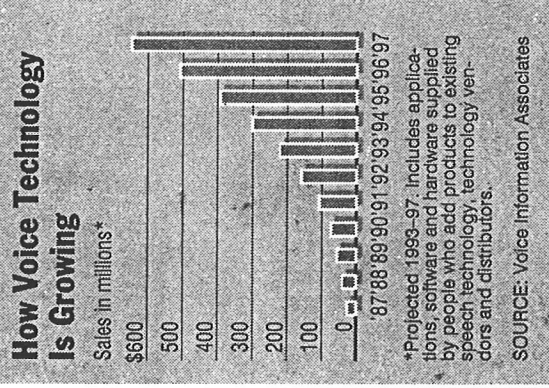
Compton's Encyclopedia, already on compact disc with an enormous amount of data, will soon come out with a voice-activated version.

And voice-powered computer games are just around the corner.



The Philadelphia Inquirer / MICHAEL S. WIRTZ

**Dr. Reese says that the voice-activated computer has revolutionized the time it takes to get out patient reports. "It used to take two days to get a dictated report back from the typist. Now we have it in two or three minutes."**



The Philadelphia Inquirer

"We are in a transition state right now," said William Meisel, editor of Speech Recognition Update News. (The field even has its own newsletter.) "In 1992 the sales of voice-activated products were in narrow niches. This year, speech recognition is making major inroads."

The arrival of text generation by talking is the result of research over two decades, much of it at Carnegie Mellon University in Pittsburgh. The first voice-powered personal computers, in the mid-1980s, had limited vocabulary (1,000 to 5,000 words), cost about \$11,000, and were used by people with handicaps so debilitating that they could not work otherwise. This group has often pioneered new technologies.

Today, the major companies producing large-vocabulary voice-recognition systems for the mass market are IBM; Dragon Systems of Newton, Mass.; and Kurzweil Applied Intelligence of Waltham, Mass.

All voice systems work by converting sound into the digital language of 0's and 1's that computers understand. This is done by means of a digital signal processor (DSP) chip. Five years ago the DSP was expensive, inaccurate and placed a tremendous burden on the computer. Today it is more accurate, does not need a super-sensitive microphone, and is so cheap that it is becoming standard equipment in some PCs.

"The holy grail in speech recognition has been ... unlimited vocabulary, true independence and continuous speech. The Pentium chip has allowed us to have everything except the continuous speech," said Bernard Bradstreet, CEO of Kurzweil.

Continuous speech — the ability to talk at a normal speed, rather than one ... word ... at ... a ... time — is the technology's last hurdle.

"If we don't have [continuous speech] in the Pentium generation, then it will be in the following," Bradstreet said.

The computer's inability to understand speech at a conversational speed frustrates many users.

Hildreth Noronha, the owner of Eurosoft Inc., a Baltimore language translation firm, first saw the IBM Personal Dictation System at the American Translators Association meeting in November. Noronha, who was working 12 to 14 hours a day typing his translations into a computer, immediately saw how the \$995 software could give him relief from

the keyboard.

Now, he calls up the German on his computer screen, translates it in his mind and dictates the English into a microphone. The English appears on a dictation window on his screen.

Noronha, who was born in India, said he has trained the system to understand his accent. "I don't roll my R's enough, but it adapted to me after a short while," he said.

But the need to hesitate between words throws him off, he said, and he looks forward to when he can speak to his computer at a more conversational speed.

Dr. Charles Reese, who runs the emergency rooms at the Medical Center of Delaware's two hospitals, Wilmington and Christina, uses a Kurzweil program specifically designed for emergency medicine.

Already, the voice-activated computer — installed 18 months ago — has revolutionized the time it takes to get out patient reports. And, he says, they're more accurate.

"It used to take two days to get a dictated report back from the typist. Now we have it in two or three minutes."

Emergency-room physician Ross Megargel, who had just treated a patient with chest pain, recently offered a demonstration.

"Chest pain," said Megargel into a telephone plugged into his computer. A form came up on the screen, with blanks to fill in and lists of choices. At the blank, PATIENT STATUS, Megargel said: "Alert and cooperative, is in no distress." The computer typed his words and moved on. At XRAY, he said: "Chest negative." At MEDICATIONS, Megargel said, "Motrin," and gave a few details about dosage. The computer automatically typed a lengthy, consumer-friendly explanation for the patient.

About three minutes after Megargel sat down at the computer, the heart patient's medical history, diagnosis and discharge instructions were finished. The electronic record stayed in the hospital's database and a printout went home with the patient.

The system is expensive — \$45,000 — but eventually pays for itself, Reese said. And once the doctors get used to pausing between words, they like it.

The annoying problem of continuous speech has already been solved in the lab, according to Jaime Carbonell, head of the Center for Machine Translation at Carnegie Mellon.

Now, Carbonell is working on another breakthrough: a system using artificial intelligence so smart that it can understand the ambiguities of a conversation and translate accurately from one language to another.

Take the word *bug*. It means one thing to a person with the flu, another to a computer scientist, and yet another to an entomologist. Carbonell and his team envision voice-activated computers, programmed for specific situations, that would understand and translate words appropriately.

"As we become more and more of a global village, it will become important to break down language barriers," said Carnegie senior researcher Alex Waibel.

One day, he said, computers will even be able to recognize gestures and read lips.

Inquirer staff writer Gayle Sims has a repetitive stress injury.