Speak and spell

Voice-recognition software has been around for ages, but for various reasons few of us exploit its benefits. With results from the latest packages nudging 100 percent accuracy rates and low-cost alternatives on the market, Wendy Brewer helps you decide if it's time to talk

Microsoft promises its much-hyped tablet PC will change the way we interact with technology, replacing the need for a keyboard with a more natural paper and pen system. But it's the power of speech that will truly set us free from our desks.

However, while other technologies have improved apace speech recognition has struggled to keep up, hamstrung by inadequate processing power as well as the complexities of the English language. But with today's multi-gigahertz speed machines and low-cost packages sporting accuracy rates that leave little room for criticism, there's never been a better time to talk to your PC – or so the software makers would have you believe.

To help you make up your mind whether or not you are ready to pack away your keyboard and limber up your vocal chords, over the next few pages we'll take a closer look at the variety of software available and predict the likely advances in the coming years.

Getting vocal

Voice-recognition applications are all around us, from automated telephone services to voice labels on our mobile phones that allow us to ring people without having to touch the dialpad.

Many language learning programs also rely on voice recognition, allowing users to repeat and record phrases in order to hear and monitor their pronunciation as they develop their skills. But such technology is expensive to develop. Intense competition coupled with the need for deep pockets to fund innovation has left just two players: IBM with its ViaVoice range and ScanSoft's Dragon NaturallySpeaking software line.

The daddy of voice recognition, IBM has been researching this area since the 1970s, first introducing it to consumer products in the early 1990s. But the real breakthrough didn't happen until 1997 when IBM released its ViaVoice brand.

Toby Manners, IBM's program director, says the launch meant finally being able to speak in a continuous, natural way. Dictation could now be accomplished with no hesitation between words, says Manners.

Competitor ScanSoft purchased the ailing Lernout & Hauspie, which previously owned the Dragon titles, back in December 01 and has since focused its energies on improving the speech facilities its predecessor started.

New pretender

So with accuracy rates in these packages hitting 97 percent and natural rather than robotic speech now possible, what's stopping the mass uptake of voice recognition? Users who stick with their trusty Qwerty keyboards often do so either through laziness (because voice-recognition software takes time to train) or because the software must be used in a quiet environment, making the noisy offices most of us work in less than ideal places.

Now software giant Microsoft is strutting its stuff in the sector, hoping to coax speech-recognition software into the mainstream via its Office XP suite. Working on the premise that if users have the tools they'll use them, both the PC and tablet versions of Office XP include voice-recognition capabilities.

"To ensure the gradual takeup of this technology by both users and the IT industry, it's imperative that we make it available within products for those who choose to use it," says Neil Laver, Microsoft's product manager.

But many users are unaware a voice-recognition facility exists within Office XP and the company openly admits that, for the tablet PC at least, voice recognition will remain a secondary application with most users opting for digital ink and handwriting recognition.

More widespread adoption of Office XP's speech facilities is further hampered by the US English voice-recognition engine which poses inherent problems for UK English speakers. General language differences such as the command 'fullstop' rather than 'period' detract from the feeling of natural speech, providing a less tailored feel.

"The speech recogniser is tuned to US English due to the relative size and demand from the US market," says Laver. "Other international versions will be developed in time. However, the US recogniser produces reasonable results for those with non-US accents."

Many users are unaware a voice-recognition facility exists within Office XP and the company openly admits that, for the tablet PC at least, voice recognition will remain a secondary application with most users opting for digital ink and handwriting recognition

Speech recognition in practice

NaturallySpeaking's interface is slightly more intuitive than ViaVoice's. Its accuracy centre is also really easier to use, with toolbars and menus well organised and accessible. The playback voice allows you to listen to notes you've already dictated and sounds far more natural than IBM's robotic alternative.

Dragon's headset works well with the software to cut out any background noise and ignores long pauses or unnecessary 'ums' and 'ahs'. ViaVoice, on the other hand, has an infuriating tendency to insert small words like 'it' and 'um' if you mumble. To be fair, though, ViaVoice does include a key-control feature so you can turn the microphone on and off quickly when you need to start or stop recording.

Net contender
ViaVoice wins hands down when it comes to surfing the web. Designed to work with Windows XP, it quickly grasped the concept of shortcuts and web addresses. Its control commands, such as 'delete this paragraph', work flawlessly and the help section with its sheer depth of information simply cannot go unmentioned.

Speech recognition in XP
Microsoft has built speech-recognition facilities in to Office XP and, in doing so, has come up with the ideal solution for anyone who needs an occasional break from the keyboard but does not need the power of a dedicated voice-recognition package.

The software is perfectly adept at document dictation (scoring around 85 percent in the accuracy league), although you will find yourself using the keyboard far more frequently with Microsoft's package because you need to manually switch between dictation and command modes using the keyboard.

Background noise is a real problem too, rather limiting where you can use the voice application. Any slight murmur is recorded as gobbledegook.

System requirements
• Dragon NaturallySpeaking 6.0 **400MHz Pentium II**; **Windows 98/ 2000/Me/XP**; **128MB RAM**; **300MB hard disk space**; **USB port (for use with bundled microphone)**; **16bit sound card.**
• ViaVoice 10.0 **300MHz Pentium and 256K L2 cache**; **64MB RAM for Windows 98, 128MB for Windows XP**; **510MB hard disk space**; **USB port (for use with bundled microphone)**; **16bit sound card.**
• Office XP Standard **400MHz Pentium II**; **128MB RAM**; **USB microphone**; **normal requirements for OS installation.**
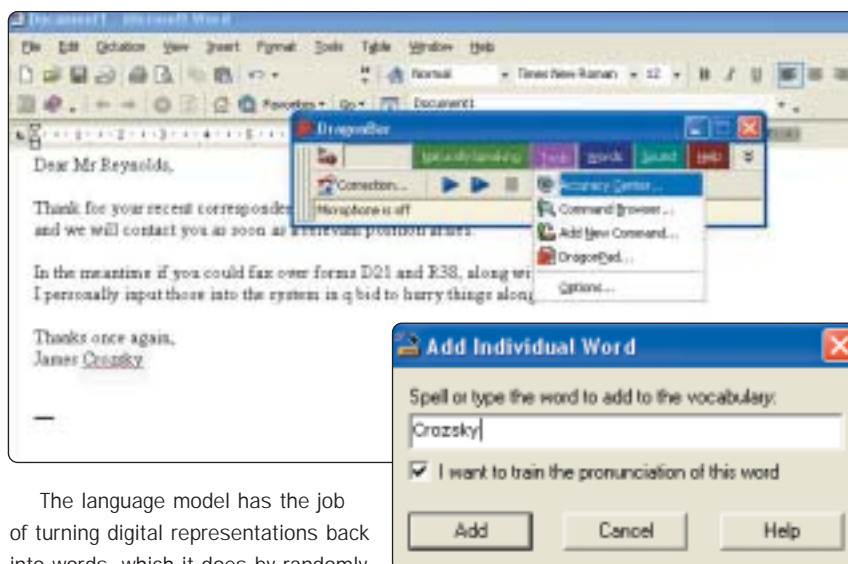
How does it work?
Speech-recognition software converts spoken words into commands which are then interpreted into actions by a PC. The computer isn't able to understand what you are saying – that sort of comprehension is entering the altogether different realm of artificial intelligence.

Instead the microphone converts what is said into an analogue signal and communicates that to the sound card. Therefore the better the sound card on your PC, the higher the accuracy rates you will achieve.

Next, an ADC (analogue to digital converter) transforms the signal into a stream of digital data. The software then uses its so-called 'acoustic model' to analyse voice sounds and convert them into units of sound called phonemes.

At this stage the acoustic model removes any interference, such as background noise. It then reduces the data into pitches and sounds, analyses it and turns the words into digital representations of phonemes.

The language model has the job of turning digital representations back into words, which it does by randomly checking its internal dictionary. The problem, however, is that many words sound similar – such as their, there and they're – but have different meanings. Therefore the program uses trigrams, grouping the word with the two before or after it to determine which is correct in the given context.

The importance of training
To carry out these tasks efficiently it is essential the software recognises a user's speech patterns. It will use previously constructed sentences to determine context so the more recorded data stored the more accurate it becomes. But within

10 to 15 minutes it should have grasped the tone of the user's voice.

Both packages encourage users to read through a series of passages when installing the software. Although NaturallySpeaking was fairly easy to train, ViaVoice did test my patience a little. If it does not recognise a word it will highlight it, stop the text and bleep incessantly until it hears the response it's waiting for.

NaturallySpeaking has a better approach, encouraging the user to continue reading whether or not the software misses a word. It then picks up that word in a later passage. For the reader, the fewer the interruptions the more natural the process feels.

IBM offers a quick and easy setup option with ViaVoice 10.0 and, although it produces lower accuracy rates to begin with, it picks up new words quickly while the software is being used and cuts out this initial frustration.

What's available?

There are several different versions of NaturallySpeaking 6.0, from the £645 version designed for integration with Lotus Notes and other business applications to the more affordable £150 solution for small businesses. All are designed to be used with everyday applications such as Word, WordPerfect and Excel.

The best option for home users who wish to sample speech recognition before splashing out is the budget £47 version. With it you can dictate into Word and navigate the internet but you will have to buy a headset or free-standing microphone as none is supplied.

IBM ViaVoice 10.0 (www.ibm.co.uk) comes in three versions. At £95 the Pro USB edition for business users is a good deal cheaper than Dragon's same-level offering, while IBM's £38 Windows Standard edition is aimed at cost-conscious small businesses and home users.

Who's using it?

People with illness or disability have been reaping the benefits of voice recognition for many years. However, it is not yet a keyboard alternative for all complaints. "Voice recognition is absolutely fantastic [but] it is not generally a good one for visually impaired people. If their hands work okay touch-typing it is by far the best

Technically speaking

So just how far has voice recognition come over the last few years? We repeated our dictation test, first carried out in May 99 using the then current version of Dragon's NaturallySpeaking Preferred and IBM's ViaVoice 98. This time we used Dragon's NaturallySpeaking 6.0 and IBM's ViaVoice 10.0 (USB edition) and tested the mettle of Office XP's built-in speech-recognition features.

The highlighted words below indicates the errors made by the software and, looking at the number of errors, the packages are 50 percent more accurate now than they were in 1999. And you get what you pay for – the integrated Office XP speech-recognition engine made double the number of mistakes as the standalone versions.

What we said:

You will shortly be receiving our quarterly financial report and Money Talk, our new monthly newsletter. Please visit our newly launched international section on our Web site at http://www.okeefeint.com.

Sincerely yours,
Sebastian Sokoloff

Sebastian Sokoloff

What we got:

Dragon NaturallySpeaking Preferred
He will shortly be receiving our quarterly financial reports and money talk, our new monthly newsletter. Please visit our newly launched international section on our WebSite at http://://WWW.O K. E. E. F. E. INT.com..

Sincerely yours,
Sebastian stucco off

Sebastian stucco off

IBM ViaVoice 98
U.S. surely be receiving our quarterly financial report and the money talk, our new monthly newsletter. Please visit our newly launched international section on our web site at H T T P:/ / W   W O K E F E I anti dot,.

Sincerely yours combine's
Sebastian cycle off

Sebastian cycle off

Dragon NaturallySpeaking 6.0
You will shortly be receiving our quarterly financial reports and money talk, our new monthly newsletter. Please visit our newly launched international section on our Web site at http://WWW.O.K.EEF.E.INT.com.
Sincerely yours,
Sebastian Stucco Off

Sebastian Stucco Off

IBM ViaVoice 10.0
You will shortly be receiving our quarterly financial report and money talk coming our new monthly newsletter. Please visit our newly launched international section on our web site at http://WWW.OKIfaint.com.
Sincerely yours,
Sebastian Soccer lot

Sebastian Soccer lot

Office XP
Do will shortly be receiving our quarterly financial retort and money talk coma our new monthly newsletter. Please missing our newly launched international second on our website at httpcolon//www.okifeint.com.
Sincerely yours,
Said action suck on off

Said action suck on off

← Voice recognition has been around for many years, starting life as a solution for the disabled

approach," said Bill Fine, spokesman for non-profit group AbilityNet. "It is often a terrific solution for those with hand and arm problems, including those with RSI [repetitive strain injury]," he adds.

Fine says speech-recognition software has not sufficiently progressed to a point at which the disabled market can use it as a sole command centre.

"Almost all our clients who use voice recognition also use their keyboard and mouse as an alternative," he said. Traditionally, visually impaired users have required the aid of a sighted person, but things are progressing.

"Given the lower technical literacy in the visually impaired population due to the costs and other barriers attached to acquiring PCs, voice-recognition systems are one of the emerging technologies that will play a key role [in the future of this software]," said Royal National Institute for the Blind spokesman Steve Tyler.

Future speak

The future of voice recognition looks promising. Car manufacturers Honda and BMW have tipped it as the future of driving, negating the need for knobs and buttons. You'll be able to control your stereo, heating system and lights by talking to them, minimising the amount of time you take your eyes off the road.

Mobile phone companies have already latched on to voice recognition possibilities

The future of voice recognition looks promising. Car manufacturers Honda and BMW have tipped it as the future of driving, negating the need for knobs and buttons

and a number of household appliances can also be controlled by speech. "In the future we envisage users being exposed to voice recognition on a wider range of devices, such as the tablet PC and handheld devices," said ScanSoft's spokesman.

Its £130 NaturallySpeaking Mobile version comes bundled with a voice memo recorder – an obvious business application – but voice recognition is set to become more mainstream and, says ScanSoft, "more people will use speech-recognition technology as the main data input method for these devices."

IBM has already released its ViaVoice package for PDAs and is concentrating its efforts on multimodal devices. "Multimodal interaction allows you to access data from databases to inventory, financial and sales information in ways that are convenient to you, instead of forcing you to adapt to technology," said IBM's Lisette Kwong.

Talking the talk

Voice-recognition software has come a long way in the past couple of years with accuracy rates now reaching around 95-97 percent (see *Technically speaking* on page 109 for a performance comparison). Interfaces have been designed for ease of use but training the software, especially with ViaVoice, is time consuming and often frustrating.

The idea that you put on a headphone and start speaking to your computer is way off the mark. But the software is impressive if you're prepared to spend time on the initial training process and, once you

get the hang of the program, it becomes almost second nature to shout commands, which is certainly quicker than typing.

So is it time to give your keyboard a rest? It depends on what you use your PC for and whether you're prepared to put in the time training the software. It can have great benefits if you're a hopeless touch-typist; those with disabilities have an obvious use for voice recognition too.

Despite Microsoft's best efforts, for most of us in the foreseeable future talking to our PCs will remain confined to verbalising frustrated outpourings at the latest blue screen of death. ■

Top tips

1. Softly, softly **Use the software in as quiet an environment as possible. Although the headsets are designed to cut out noise, a peaceful environment makes for a better recording.**
2. Stay the same **Try to position the microphone in the same place every time you use it, otherwise it may take a while to recognise your voice.**
3. On the edge **Position the microphone at the edge, rather than front, of your mouth for best results.**
4. Test first **Use the playback features to test the clarity of the recording before you start dictating a document so that you can immediately locate and resolve any problems.**
5. Pause for thought **Complicated words should be spelled out but make sure you don't leave large pauses between letters – it just confuses the software.**
6. Save phrases **Add multiple-word phrases to the allocated menu. It's especially useful to do this with addresses and complicated names.**
7. Lead by example **Dictate as many of the initial training passages as you can; the more examples of your voice it has stored the more accurate the software will be.**
8. Spec it out **Speech-recognition programs really test the processing power of your PC so invest in more power if you're running anything below a 400MHz processor.**