

Your word is their command

SPEECH RECOGNITION PROGRAMS ARE BECOMING SO CLEVER, THEY CAN NOW UNDERSTAND EVEN THE MOST ARTICULATED CHALLENGED OF USERS. ROGER GANN AND ADELE DYER TESTED THE **LISTENING ABILITIES** OF THE LATEST LEADING PRODUCTS.

THE DAYS WHEN you had to speak in staccato dalek tones to your computer as if you too were a machine, are long gone. Continuous-speech products let you talk naturally, if slowly and carefully, to your PC. The obvious advantage of this is that it is faster. Most people can type at only 40 words a minute at best, but they can talk at 200 words a minute. Most continuous-speech recognition systems make it possible to dictate up to 140 words a minute when the system is fully trained. Even if you factor-in the time it takes to correct the dictated text, you can still input text faster than you can type. The downside is that continuous speech products are extremely power hungry. Their development has been delayed as much by the lack of sufficiently powerful processors in office and home PCs, as by the difficulties of coding.

Dragon Systems NaturallySpeaking Preferred

The latest release of Dragon Systems' NaturallySpeaking is available in four different versions, all supplied with a noise-cancelling microphone headset. The Standard, Preferred, Professional and Point & Speak versions feature Dragon's new BestMatch technology, which boasts improved recognition accuracy and support for natural language commands in Word 97. This technology recognises commands even if you phrase them in different ways. All versions have an active vocabulary of 62,000 words and space for a further 54,000 user-added words.

All versions offer integration with both Word 97 and WordPerfect 8, which is good news, although support for WordPro as well would have been even better. You can dictate directly into pretty much any other Windows applications, too. Dragon's entry-level continuous speech package, Point & Speak, sells for £50; it only lets you correct mistakes using the keyboard, not your voice, however. The Preferred and Professional versions allow you to store dictation on a MiniDisc for later processing — an option dubbed NaturallyMobile. They also offer text-to-speech synthesis and audio playback of words dictated.

Installation follows the usual well-trodden path, with wizards leading you through the audio set-up and then the voice enrolment. This is a two-stage process: a short initial calibration is followed by a 30-minute spell of reading out loud. Fortunately, you are given a choice of



Most people can type at only 40 words a minute, but can talk at 200 words a minute

Features to look for in voice recognition software

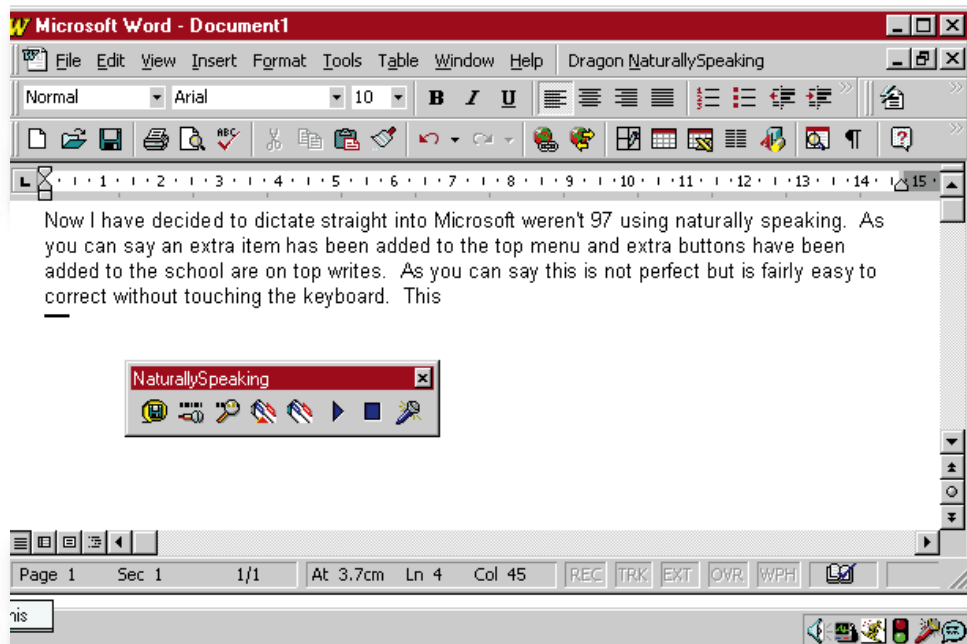
- A standard dictionary of around 30,000 words and a back-up dictionary of up to 260,000 words.
- Most packages let you add up to 64,000 of your own words, or

- to add specialist dictionaries — a legal one, for example.
- You can dictate into a WordPad-like environment, various versions of Word or WordPerfect, or any

- Windows application.
- Correction facilities: these need to be seamless if you don't want to spend more time correcting text than dictating it.
- Command and

- control facilities for formatting text and managing applications.
- Some packages let you set up macros for repetitive tasks.
- Most packages have text-to-speech facilities:

- they can read back what you said, or synthesise speech from text imported into the program.
- The majority of packages let you enroll more than one user.



▲ **NATURALLYSPEAKING**
INTEGRATES WITH
WORD 97, APPENDING
ITSELF TO THE MENU

sample scripts to read out, one of which is a rather amusing view of computers and Comdex. Crunching the voice data took less than 15 minutes, so the whole enrolment process is commendably swift, and there are some helpful multimedia clips, too, showing how to use the program all the way through. As in IBM's ViaVoice 98, there are several ways to get your words on to the screen. One method is via a WordPad lookalike, which you can dictate into, and then copy the words across or save as a text or RTF file. You can also configure the program from here. Alternatively, you can fire-up Word 97 or WordPerfect 8. If you're using Word, an extra menu item appears, from which you select NaturalWord, which takes just a few seconds to load. An extra toolbar containing the NaturallySpeaking buttons is then appended to your toolbar.

Once NaturallySpeaking is loaded, a speech bubble on the taskbar indicates whether you can dictate into the current active application, and a little microphone icon shows whether the mike is switched on or off. We tested NaturallySpeaking after only 30 minutes of training, and initial results were impressive. Not only did it post a low rate of mistakes when compared to ViaVoice 98 and FreeSpeech98, but it also registered words noticeably faster than its rivals. As you talk, a small yellow speech bubble appears next to the cursor, and you can check

as ViaVoice, NaturallySpeaking has a vocabulary builder which searches text files, to look up unusual words to add to your vocabulary. Adding special words is particularly easy: all you do is point NaturallySpeaking at a list of words —

Adding special words is easy: you point NaturallySpeaking at a list of words and the program processes most words automatically

such as Word 97's CUSTOM.DIC file — and the program processes most words automatically, so you don't have to say them. Even so, it can't cope with all words — FDISK, for example.

NaturallySpeaking Preferred, however, isn't quite as clever as ViaVoice 98. It won't automatically format numbers and telephone numbers — you have to say "Pound sign 45" to get £45. Dragon Systems still uses its clunky MouseGrid method of manoeuvring your mouse. And a control centre on permanent display would be helpful.

PCW DETAILS

NaturallySpeaking

★★★★★

Price £149 (£126.81 exc VAT)

Contact Dragon Systems UK 01242 678575

www.dragonsystems.com

System Requirements Windows 95, NT 4, P133MHz, 60Mb hard-disk space, 40Mb to save recorded speech, 15Mb to install text-to-speech, 32Mb RAM.

Good Points Fast, good level of accuracy, easy to use, short enrolment.

Bad Points Not as sophisticated as ViaVoice 98.

Conclusion On a par with ViaVoice 98 for overall accuracy, but with a noticeable speed advantage.

your words here before they are transferred to your application. And you don't have to stop dictating when you want to edit or make corrections. If you make a mistake, you just say "Scratch that", to delete the last thing you said. You can also select words and navigate throughout your document simply by saying them. The Preferred version has a vocabulary of 42,000 words which you can bump up to 230,000, although you may have to increase your memory substantially to support this many. In the same way

IBM ViaVoice 98 Executive

Both the £140 Executive Edition and the £50 Home Edition of ViaVoice 98 have the same 32-bit recognition engine and support continuous speech recognition. They also offer direct dictation into Word 97 and the IBM WordPad lookalike, SpeakPad. ViaVoice 98 offers a natural language command for Word 97, too. It has a 56,000 word vocabulary and can absorb a further 64,000 manually-added words.

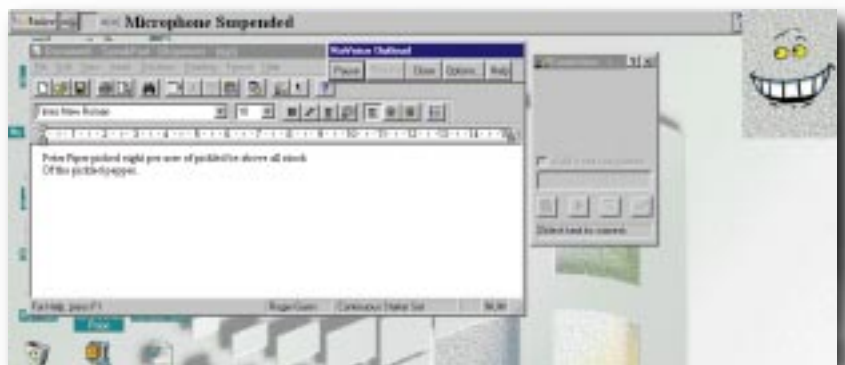
New to this release of ViaVoice 98 is the ability to automatically recognise and format numbers, dates, times and prices. The Executive version also enables you to dictate into most Windows applications, and features full system-wide command and control. As a bonus, you also get a specialised vocabulary, Legal Topic.

Installing the package is very simple. Online help is of a high order and you get useful multimedia demonstrations. On launching the software for the first time, you have to set up the microphone and the audio levels. Then a three-minute wizard establishes you as a user by prompting you to say a few words and sentences, and by building a personal profile. At this point, your computer is ready to start listening.

Initially, enrolment consists of dictating 100 sentences, which takes less than 20 minutes. However, accuracy won't be very good until you complete the full enrolment of 476 sentences, which takes about an hour and a quarter: and it takes another hour or two for the program to crunch that data. Although the program is stated to run on a Pentium 166MMX, it ran quite slowly on a non-MMX P166 with 48Mb, so you might

consider a P200 as a more realistic entry level. ViaVoice 98 Executive comes bundled with the Andrea NC80 microphone. This is a rather basic headset, and the microphone was a little reluctant to stay adjusted.

In theory, ViaVoice 98 is backwardly compatible with the previous version. However, when I tried upgrading my two ViaVoice installations on a Windows 98 and a Windows NT 4 PC, the program wouldn't even load. The package offers the option of migrating previous enrolments, but I was unable to do this. IBM's technical support suggested hacking the registry to expunge all trace of previous installations, which is something



▲ **WITH VIAVOICE 98 YOU CAN DICTATE TO ANY WINDOWS APPLICATION**

the software should have handled automatically. But, in practice, losing previous enrolments isn't a big problem, because the latest version samples at 22KHz, twice the speed of the previous version. As a result, migrated previous enrolments will be distinctly lo-fi and less accurate.

ViaVoice 98 presents a clean, simple interface. The VoiceCenter Taskbar, which sits at the top of the screen, displays the user's name, a simple

How does speech recognition work?

All speech recognition packages work in the same way. A microphone picks up the sound of your voice and then feeds the analogue sounds back to your sound card. The better the microphone, the more background noise it will filter out, increasing the chances of the word being recognised accurately. Minimum sound requirement is a good 16-bit sound card – the sound chips in many

notebooks simply aren't good enough. **The sound is compared to a database of sounds stored on the system.** Whole words are not identified immediately, but are broken up into their constituent parts, or phonemes – English has about 80 phonemes from which all words are constructed. Phonemes are analysed individually, then in groups, and finally as words, to make sure a

correct identification has been made. This statistical analysis eliminates various possibilities according to the likelihood of one phoneme following another and making up a word. **Most packages work with a number of pre-recorded sets of phonemes.** They will ask you to identify yourself as either male or female and over or under 16. From this – and from the analysis

of words recorded during the training session – the system decides which set of phonemes it will use. **Language modelling is also used to second-guess what the word might be** according to where it comes in the sentence. **Finally, contextual modelling looks at the words in sequences to find the right match.** For example, the system might decide

whether it is "to", "two" or "too", depending on the words before and after it – for example, "walk to London", "walk two miles" or "walk too far". **The majority of packages can carry out contextual modelling on only three words at a time.** But as processor power increases, they will be able to carry out grammatical analysis of full sentences, and so improve accuracy.

microphone on/off button, a volume metre and a ViaVoice menu button. With the mike active, you simply say "Dictate to word": Word 97 is then loaded and you can start dictating.

IBM has always loved its totally arcane error messages and has continued to uphold tradition in the case of ViaVoice 98. Whenever a problem crops up, you receive an incomprehensible and thoroughly unhelpful message, along the lines of "SMAPI RC=60".

ViaVoice 98 is modelless — in other words, you can mix dictated text and commands. All you have to do is make a slight pause, and ViaVoice 98 knows that your next word is a command rather than text. This can take some getting used to, as most people do not dictate in one fluid stream. If you're not sure what commands are available, a "What can I say?" command conjures up a long list for you to browse through. Easy navigating is one of ViaVoice 98's strong points.

The way you correct the dictated text is a big improvement on earlier releases of ViaVoice. You can opt to have a correction window permanently on display, and when you speak the words you want to correct, or double-click on them, they will appear in the window. You then

listen to the words, and either dictate them again, or select from a list of alternative suggestions in the window. These words can be added to the program's vocabulary automatically. As before, ViaVoice 98 allows you to analyse documents to pick out words that aren't in its vocabulary.

Like NaturallySpeaking, ViaVoice 98 Executive includes a text-to-speech synthesiser which can read a document back to you. The voice is fine, but Robbie the Robot has the edge on it.

PCW DETAILS

ViaVoice 98 Executive

★★★★★

Price £140 (£119.49 exc VAT)

Contact IBM Speech Systems 01705 492249

www.software.ibm.com

System Requirements Windows 95, 98 or NT 4. P166MMX, 256Kb L2 cache, 180Mb hard-disk space, 32Mb RAM with 95/98 (48Mb with Word 97), 48Mb RAM with NT 4 (64Mb with Word 97).

Good Points Many new features, high degree of accuracy, very easy to use.

Bad Points Upgrade feature doesn't work.

Conclusion Offers state-of-the-art recognition rates. Once IBM sorts out the migration issues, it will be a tough choice between this and NaturallySpeaking 3.0.

Lernout and Hauspie Voice XPress Plus

At the time of writing, the UK version of Voice XPress Plus was not available. Lernout and Hauspie was unwilling to let PCW loose on the US version, and insisted on demonstrating it to us. Hence, we didn't get the chance to test it as thoroughly as we would have liked.

The UK version of Voice XPress Plus, which lets you dictate straight into Word 7 and Word 97, will be in the shops by mid-September. This version will be followed soon afterwards by Voice XPress Pro, which allows you to dictate into any Windows application. When you first load the package, it puts 30,000 words into active memory and stores a further 230,000 words on the hard

disk. Using a PII 233 with 128Mb RAM — a much more powerful machine than the one we used to test the other packages — enrolment took a full 50 minutes and processing that data took 41 minutes. After enrolment, the recognition rate seemed good — about 90 percent accurate

Vocabulary Training

Enter a word or phrase you want to add, train or remove. You can also enter a voice command to train.

TeleTalk

Add

Train

Remove

Help

☐ Consider this phrase as a single vocabulary word appropriately.

Close

▲ TEACH VOICE
XPRESS PLUS NEW
WORDS

when speaking quite slowly. When we

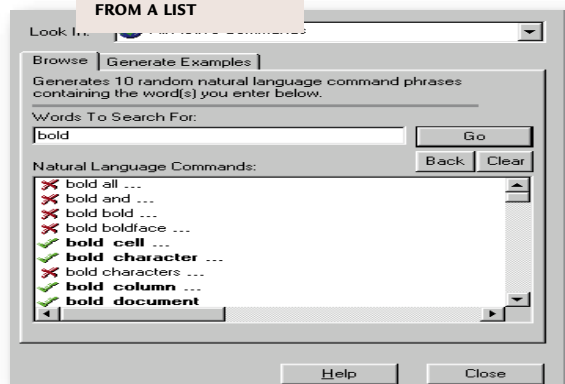
stepped up the pace, it was able to catch a good 80 percent accurately. As the demonstration was quite short, we could not see how much it would improve over time. Even in the short time we had to evaluate Voice XPress Plus, it did not appear to be as well thought out as other packages. For example, you still needed to swap from dictation mode to command mode, in order to format text, and some of the correction methods seemed unnecessarily cumbersome.

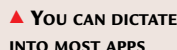
PCW DETAILS

Voice XPress Plus

System Requirements Windows 98, 95 or NT 4, P166MMX, 40Mb RAM (98/95), 48Mb RAM (NT 4), 130Mb hard-disk space.

▼ PICK THE COMMAND
YOU WANT TO USE
FROM A LIST





164 • PERSONAL COMPUTER WORLD • OCTOBER 1998