

# Talking Back

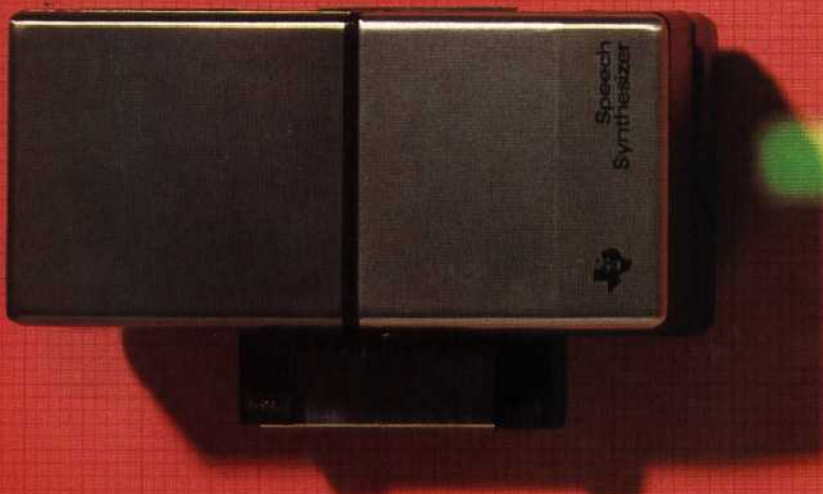
**It used to be pure science fiction. Now with a speech synthesiser your computer really can talk to you. And it needn't sound like a Dalek, either**

While the science of speech recognition has yet to be fully developed, the generation of electronic speech has been mastered. Until recently, however, the computing power and memory capacity needed to produce human-like utterances were substantial. Now, with the aid of a suitable add-on, almost every home computer and electronic toy is capable of talking back. The rapid advances in technology and the fall in the cost of computer components have made the talking computer commonplace.

IAN MCKINNELL

When people talk, sounds of three distinctly different types are produced. The first is 'voiced' or vowel-type sound — *oo, ar, ee* and so on. These

are produced by the vibration of the vocal chords in the throat, the frequency of this vibration determining the vowel sound. The second is the 'fricative' or unvoiced sound, such as *ss, sh, t* and *ff*. Here air from the lungs bypasses the vocal cords and the frequency of the sound is controlled by the positioning of the lips and tongue. The third 'sound' is silence or — to be more precise — the gaps occurring within words like *six, eight* and so on. You may not realise that there are gaps in these words, but if you try to pronounce them slowly you will realise that it is impossible to run smoothly from the sound of *i* into *x*.



## Building Blocks Of Sound

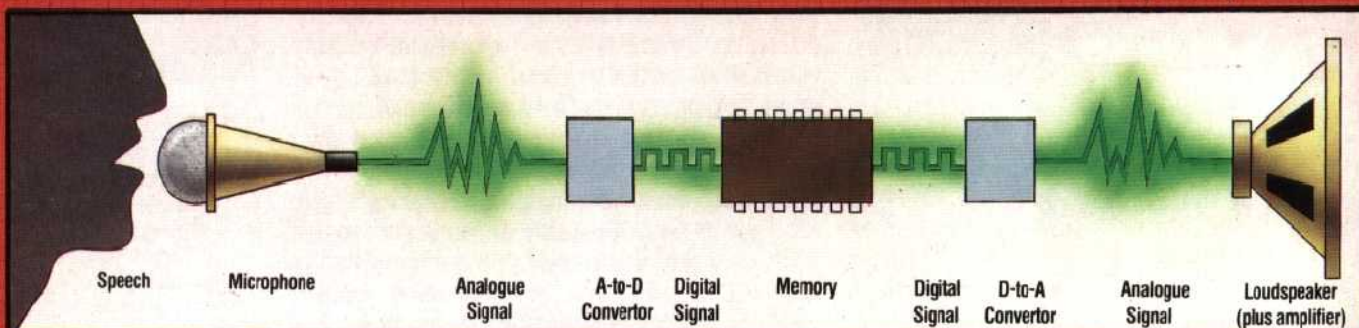
There are two ways of generating speech-like sounds electronically. The first, until recently the most common, is that of synthesis by rule. By analysing the frequencies contained within speech it is possible to devise a system of rules that allows us to create any given sound from its components. For example, the word 'too' could be defined as so many milliseconds of the mixture of frequencies that make up the sound *t*, followed immediately by the *oo* frequencies.

These individual building blocks are called 'phonemes' and by using them in various combinations any word can be constructed. The individual characteristics of a human speaker tend to be lost when speech is generated in this way, but the words can be recognised and understood. Because the rules for generating the phonemes are built into the equipment itself, the user is able to enter a list of the phonemes into the system. These are reproduced through a small speaker. With a little practice it is possible to generate complete sentences instantly by calling up sequences of phonemes, which can usually be stored in BASIC strings.

The second method of speech synthesis relies

### The Flow Of Sound

Speech can be digitised and stored in memory, either RAM or ROM. Electrical output from a microphone is passed through an analogue-to-digital convertor. The output from this chip is a digital pattern of 1s and 0s. The speech can be recreated using a digital-to-analogue convertor, an amplifier, and a loudspeaker



KEVIN JONES