

[*Financial Times*, 30 October 1996, p.17]

Challenge of translating the spoken word

By VANESSA HOULDER

A machine that can translate a complex conversation is usually seen as the stuff of fantasy. But a prototype of a pioneering speech translation system, unveiled in Germany last week, represents a step, albeit an early one, in that direction.

The Verbmobil system is designed to help Japanese and German managers who can understand a little English to agree on the time and place of a meeting. It can translate 2,500 German words and 400 Japanese words into English. It provides the 'correct' translation in 75 per cent of cases.

That may seem a modest achievement. But it represents a significant advance in some of the most difficult aspects of independent speech recognition, language processing and synthesis.

'Verbmobil is, right now, the largest, most innovative speech-to-speech translation system there is in the world,' says Reinhard Karger, the project manager.

Unlike conventional dictation systems or those that process written language, the Verbmobil system has to deal with the intermediate sounds of speech such as 'ers and ums' or sneezing.

The system also has a speech-rhythm module that works out phrasing from breaks and intonation.

It can also distinguish the meaning of an ambiguous word from the emphasis in the sentence. For example, the German indefinite article can be translated as 'a' or 'one' in English. By analysing the stresses in the sentence, the system can work out the difference between, say, 'to meet for one hour' and 'to meet for an hour'. Other ambiguities are handled by reference to the context of the speech.

The project, which has received DM64.9m (Pounds 28.20m) in sponsorship from the German government and DM31m from industry, has involved 150 researchers and engineers from 29 universities and companies in Germany, the US and Japan - including Alcatel, Daimler-Benz, IBM and Philips. The project is run by Professor Wolfgang Wahlster at the German Research Centre of Artificial Intelligence in Saarbrücken.

The system is composed of 24 communicating modules. The acoustic modules record and digitise the speech, while the recogniser modules use statistical language models, lexicons and key-word spotters.

The linguistic modules use complementary forms of analysis, known as deep and flat. The deep analysis searches for strings of words that are grammatically correct, and analyses their meaning by referring to a module containing a lexicon of phrases. The flat analysis deals with non-grammatical sentences by extracting words and phrases that appear relevant to the conversation.

The transfer module then translates the phrases into their structurally correct English equivalent, which is synthesised into speech. The modules developed for the translation system can be used in other devices. For example, Mercedes-Benz has used the speech recognition technology in a device for car telephones that recognises phone numbers. The next phase of the project aims to extend Verbmobil's vocabulary to 10,000 words and expand its scope to deal with travel planning as well as meetings. It will also be developed to translate into German and Japanese. Researchers on the Verbmobil project are confident that it will lead to viable products. But the development of a machine that can handle a free-ranging conversation is not yet on the horizon.

Vanessa Houlder