



## It's the next best thing to a Babel fish

- 26 October 2006
- From New Scientist Print Edition. [Subscribe](#) and get 4 free issues
- Celeste Biever

Advertisement

Imagine mouthing a phrase in English, only for the words to come out in Spanish. That is the promise of a device that will make anyone appear bilingual, by translating unvoiced words into synthetic speech in another language.

The device uses electrodes attached to the face and neck to detect and interpret the unique patterns of electrical signals sent to facial muscles and the tongue as the person mouths words. The effect is like the real-life equivalent of watching a television show that has been dubbed into a foreign language, says speech researcher Tanja Schultz of Carnegie Mellon University in Pittsburgh, Pennsylvania.

Existing translation systems based on automatic speech-recognition software require the user to speak the phrase out loud. This makes conversation difficult, as the speaker must speak and then push a button to play the translation. The new system allows for a more natural exchange. "The ultimate goal is to be in a position where you can just have a conversation," says CMU speech researcher Alan Black.

In October 2005 Schultz and her colleague Alex Waibel demonstrated the first automatic translator that could pick up electrical signals from face and throat muscles and convert them into text or synthesised speech - a technique called sub-vocal speech recognition. This ran on a laptop and translated Mandarin Chinese to English or Spanish, but it could only translate around 100 words, each of which had first to be spoken into the system by the user, to "train" it on their voice.

Now the team has developed a system that can recognise a potentially limitless lexicon. Their secret is to detect not just words but also the phonemes that form the building blocks of words. The system then uses these to reconstruct the word. To translate from English to another language, the user only has to train the system on the 45 phonemes used in spoken English.

The researchers use software that has been taught to recognise which phonemes are most likely to appear next to each other and in what order. When it encounters a string of phonemes it is unfamiliar with or has only partially heard, it uses this knowledge to come up with a range of sequences that make sense given the surrounding phonemes and words, assigns a probability to each one, and then picks the one with the highest probability.

The system still has some way to go. Faced with a sequence of words it has never heard before, it picks the right phoneme sequence only 62 per cent of the time. This nevertheless ranks as "a very significant achievement" according to Chuck Jorgensen, who is working on using sub-vocal speech recognition to control robots at NASA's Ames Research Center in Moffett Field, California. "This is showing that the technology is really within reach."

Schultz's team plan to attach the phoneme recognition software to their prototype Spanish or German translators, once they have improved its accuracy.

From issue 2575 of New Scientist magazine, 26 October 2006, page 32

[Close this window](#)