

Phone call translator app to be offered by NTT Docomo

An app offering real-time translations is to allow people in Japan to speak to foreigners over the phone with both parties using their native tongue. NTT Docomo – the country's biggest mobile network – will initially convert Japanese to English, Mandarin and Korean, with other languages to follow. It is the latest in a series of telephone conversation translators to launch in recent months.

Lexifone and Vocre have developed other products. Alcatel-Lucent and Microsoft are among those working on other solutions. The products have the potential to let companies avoid having to use specially trained multilingual staff, helping them cut costs. They could also aid tourism. However, the software involved cannot offer perfect translations, limiting its use in some situations.

NTT Docomo unveiled its Hanashite Hon'yaku app for Android devices at the Combined Exhibition of Advanced Technologies (Ceatec) show in Japan earlier this month, and plans to launch it on 1 November. It provides users with voice translations of the other speaker's conversation after a slight pause, as well as providing a text readout. "French, German, Indonesian, Italian, Portuguese, Spanish and Thai will be added for this application in late November, raising the number of non-Japanese languages to 10," the firm said in a statement. "Fast and accurate translations are possible with any smartphone, regardless of device specifications, because Hanashite Hon'yaku utilises Docomo's cloud [remote computer servers] for processing." The caller must subscribe to one of Docomo's packages to be able to use it.

NTT Docomo will soon face competition from France's Alcatel-Lucent which is developing a rival product, WeTalk. It can handle Japanese and about a dozen other languages including English, French and Arabic. The service is designed to work over any landline telephone, meaning the company has had to find a way to do speech recognition using audio data sampled at a rate of 8kHz or 16kHz.

Other products – which rely on data connections – have used higher 44kHz samples which are easier to process. Alcatel-Lucent uses a patented technology to capture the user's voice and enhance it before applying speech recognition software. The data is then run through translation software before being run through a speech synthesiser.

The firm said all this could be done in less than a second. However, it has opted to wait before the speaker has stopped talking before starting the translation after experiments carried out with workers at insurance company Axa suggested users preferred the experience.

"We are still working on improving the system," Gilles Gerlinger, the product's co-founder, told the BBC.

"You can do conversations with one person, but we want to allow conferences with 10 people and four different languages, and the system would provide translations in every language needed.

"We also have a project called MyVoice which can have a synthetic voice that sounds like your real one."

Mr Gerlinger suggested that his firm would make money from the product by renting servers with the necessary software to big businesses, and charging smaller ones a fee for the amount of time they used the service.

<http://www.bbc.co.uk/news/technology-20004210>