Bringing down the language barrier... automatically

May 2 2008

Progress being made by European researchers on automatic speech-to-speech translation technology could help the EU tackle one of the biggest remaining boundaries to internal trade, mobility and the free exchange of information – language.

With 23 official languages, European institutions spend more than a billion euros a year translating documents and interpreting speeches. Companies trading across the EU's internal borders spend millions more just to understand their business partners.

The situation, unparalleled anywhere else in the world, makes Europe a natural market for automatic translation technology, and, logically, a leader in the development of systems that can help speakers of different languages communicate.

"There is an evident need for this sort of technology in Europe and elsewhere in the world… it saves time and costs over human translation," explains Marcello Federico, a researcher at FBK-irst in Trento, Italy.

But no one has been able to develop an automatic translation system that comes anywhere close to the capabilities of a human translator or interpreter. Internet translations are a case in point, littered with punctuation errors, misplaced words and grammatical mistakes that can make them almost unintelligible.

Other systems can only translate certain predefined words and phrases,

so-called 'constrained speech' that suffices for a tourist booking a hotel or checking flight times but is next to useless if you want to understand a news bulletin.

Federico led a team that sought to achieve something far more ambitious. Working in the EU-funded TC-STAR project they tackled what is perhaps the biggest human language technology challenge of all: taking speech in one language and outputting spoken words in another.

First in speech-to-speech translation

"For humans, translation is difficult. We have to master both the source language and the target language, and machine translation is significantly more difficult than that," Federico notes. "To our knowledge, TC-STAR has been the first project in the world addressing unrestricted speech-to-speech translation."

For such a system to be able to translate any speech regardless of topic and context, three technologies are used, all of which are still far from perfect. Automatic Speech Recognition (ASR) is used to transcribe spoken words to text. Spoken Language Translation (SLT) translates the source language to the target language. Text to Speech (TTS) synthesises the spoken output.

The TC-STAR research partners developed components to handle each of those tasks, creating a platform that has brought the state of the art of translation technology a step closer to matching the performance of human translators.

One of their key innovations was to combine the output of several ASR and SLT systems in order to make the transcription and translation phases considerably more accurate than comparable systems.

Based on the BLEU (Bilingual Evaluation Understudy) method, a way of comparing machine and human translations, evaluations of the quality of translations improved by between 40% and 60% over the course of the project, while up to 70% of words were translated correctly, even if they were not placed in the right position in a sentence.

From speeches to Chinese news bulletins

The 11 partners – including big telecom and entertainment companies, such as Nokia, Siemens, IBM and Sony – worked with recordings of speeches from the European Parliament, which they translated between English and Spanish. They also worked with radio news broadcasts, which they translated from Chinese to English.

Though the system still cannot match the accuracy of a human translator or interpreter, Federico is convinced that, with further research a commercially viable automatic speech-to-speech translator will be feasible within a few years, at least for some simpler language pairs.

In the meantime, components developed in the TC-STAR project have been made available under an open source license. The project has also led to at least one spin-off company and a follow-up initiative.

Called PerVoice, the spin-off is offering remote-automated transcription services for companies and public bodies.

"It saves them time and money to have minutes of meetings or town council sessions transcribed automatically," Federico notes.

The follow-up project, JUMAS, focuses on developing a similar transcription system to record court trial proceedings.

Source: ICT Results