

CMU's InterACT Center will reverse Babel

BY SEAN CONROY

the tartan.org 26/9/05

MONDAY, SEPTEMBER 26, 2005

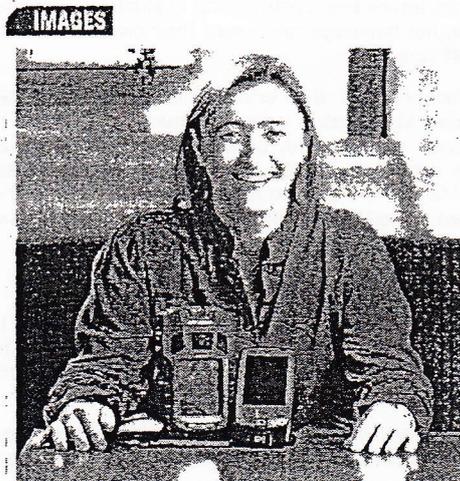
A British doctor helping tsunami victims in Thailand wants to give advice to a patient. In the past, he'd tote along a translator; but today, he speaks into his PDA — and in seconds, it's translating into Thai aloud.

A soldier in Iraq needs to talk to the local warlord, who doesn't speak a word of English. He issues his statement into a box in his hand, and it repeats to the Iraqi in Arabic. He responds in Arabic, and his words are likewise translated into English.

Sound fantastic? Not anymore. The recently birthed "International Center for Advanced Communication Technologies" (InterACT) jointly hosted by the Language Technologies Institute, Carnegie Mellon, and the Fakultät für Informatik, Universität (TH), in Karlsruhe, Germany, is a facility devoted entirely to multimodal human and human-machine communication.

How It Works

InterACT's goal is to create a comprehensive software system that can translate audible speech between any two languages. This extraordinary task breaks down into three parts: transforming what people say into text (automatic speech recognition), translating this text into another language (machine translation), and synthesizing translated text into audible sound (speech synthesis).



Tanja Schultz, principal researcher at the Language Technologies Institute, with the Phraselator and PDA used for language translation. (photo by Jiaqi Tan)

Primary analysis is acoustic: The computer recognizes the sound of unique phonemes, or syllables. Secondary analysis is vocabulary: The computer combines phonemes and integrates them into words. Final analysis is language: The computer looks at three-word "trigrams" to check for logical consistency. Once the system can recognize sentences, it must be "trained" to recognize actual speech. Linguists spend many hours transcribing speech into text. And that requires someone who is familiar with both English and the second language requested — as well as the writing systems for each. In a case involving colloquial Arabic, a writing system had to be invented, requiring almost half a year of hard work.

What follows is a tour of the most exciting projects currently underway at the InterACT center, based on several personal interviews with Tanja Schultz, one of the center's principal researchers.

The TansTac Project

Four weeks ago, DARPA selected InterACT's Two-Way Speech Translation System as one of three finalists in their search for a device to provide a better speech translation system for the U.S. military.

At the beginning of this summer, DARPA issued a challenge to InterACT: Could they build a two-way speech translator between Iraqi Arabic and English in only 100 days? The military had an urgent need for Iraqi translators, and "they wanted to get this into the field ASAP." InterACT accepted the challenge, but soon ran into trouble. The 100 days were ticking away, and they still had no data.

The translation texts and accompanying audio between Iraqi Arabic and English were taking forever to build, since there was no writing system for Iraqi-Arabic. When the data finally arrived, it was August 2: only 23 days before the deadline. Along with several other CMU related spin-offs, InterACT's 7-member team immediately went to work.

Despite time and data limitations, though, InterACT's device performed so well at the evaluation that the team was flown down to Fort Huachuca, Ariz., for a final, comprehensive evaluation.

InterACT used VoxTec's Phraselator, a ruggedized PDA, as hardware.

Schultz gave The Tartan the opportunity to test the result, something she called the Speechalator.

I activated the microphone and said "Hello"; the PDA thought for a few seconds, and then loudly announced: "Marhaba". Theoretically, an Iraqi would then respond in Arabic, and his words would be translated into English. 9.2005 11:12

up" and "get out of the car" were perfectly recognized.

Although it is uncertain whether InterACT's software will ultimately be deployed to Iraq, there is no question that the center is one of very few institutions that are able to produce such a device.

The SPICE Project

Schultz was eager to tell The Tartan about "her baby," a speech-processing project called SPICE whose goal is to produce an Interactive Creation and Evaluation toolkit for new languages. Those who know English and another language can create their own speech recognition and translation system with the SPICE software.

Schultz explained that the program will "bridge the gap between technology expertise and language expertise." It takes about half a year for a human to build a basic language database, and with so many languages in the world, InterACT could be forced to spend the majority of their time building these databases. SPICE asks for public to help to solve the problem by contributing to a new public-domain of language databases. So far, she has had three visiting researchers use the system to build the first-ever speech recognition and translation systems in Afrikaans, Vietnamese, and Bulgarian.

The Dolphin Project

Are dolphins as intelligent as we are? InterACT is trying to find out. In collaboration with the Wild Dolphin Project and the Naval Research Center, Schultz and her CMU colleagues Alan Black and Robert Frederking are working on a project for the Wild Dolphin Project to translate dolphin cries, or 'dolphones,' into English.

Since dolphins communicate at frequencies up to ten times the audible limit of the normal human ear, the InterACT group needed to first build special microphones and computers.

The next step was to record the dolphin sounds. Last year Frederking spent 10 days on a boat in the Bahamas swimming with dolphins and collecting data. Schultz's team then worked on deconstructing dolphin whistles into what she calls "dolphones." Although the project is still in an early stage of development, Schultz and her collaborators have already built and tested a software package that "takes the signature whistle of a dolphin as input and outputs the identity of a dolphin." With this system, the members of the Wild Dolphin Project could identify dolphins near their boats in open ocean without actually seeing them.

Schultz hopes to eventually be able to work with marine biologists in deciphering what the various dolphin sounds mean. Perhaps one day we will be able to speak "dolphinish."

Non-Audible Speech Recognition

Ever had to "leave the room" to make a phone call, so as to avoid disturbing everyone else? InterACT's Non-Audible Speech Recognition solves this by recognizing speech based on face muscles, and not on any audible sound.

A set of electrodes is placed on the subject's face, and the resulting "myoelectric signals" from the articulatory face muscles are analyzed, training the computer to correlate motion with words. Schultz demonstrated a test of the system, a phone conversation between a "normal" party and one who spoke inaudibly. As the subject mouthed the words, the computer immediately recognized them and spoke into the phone. It was as if the other party was talking to Valerie the Robot. Amazingly, the Non-Audible Speech Recognition project is "merely" a master thesis.

One day Schultz and her grad students may finally develop a machine capable of virtually instantaneous translation between all languages. But is that a good thing? Nobody wants to talk to a box, and one wonders if this could ultimately result in less communication.

Schultz disagrees: "Some day, the box will disappear from our sight — it will do its magic in the background — and we can just concentrate on communicating with the human in front of us."

Her work is of great importance to those who need to communicate right now. And the doctors in Thailand are grateful.

If you are interested in helping out with the SPICE Speech Processing Project, visit the project website at www.cmuspice.org. Learn more about InterACT at www.is.cs.cmu.edu.

If you are interested in helping out with the SPICE Speech Processing Project, visit the project website at [SLANT 12]www.cmuspice.org[/SLANT 12]. Learn more about InterACT at [SLANT 12]www.is.cs.cmu.edu[/SLANT 12].