Machines not lost in translation

By Ann Harrison

Faced with daunting translation problems in war and disaster zones around the world, the U.S. military is refining a handheld voice-translation device that will soon be used by police and emergency-room doctors back home.

The palm sized PDA-like Phraselator lets users speak or select from a screen of English phrases and matches them to equivalent pre-recorded phrases in other languages. The device then broadcasts the foreign-language MP3 file and records reply dialog for later translation. Unlike other machine translators, the Phraselator does not require that users train it to recognize their voice, and it produces human rather than synthesized speech.

Phraselators have recently been used by the U.S. military in tsunami relief operations. The voice module for humanitarian assistance now offers 2,000 phrases in Hindi, Thai, Indonesian and Sinhala such as: "Are any of your family members missing?" "We have medical supplies." And, "Has anyone tested this water?"

Navy doctor Lee Morin generated the idea for the Phraselator during Operation Desert Storm when he loaded Arabic language audio files onto his laptop and clicked on phrases to help communicate with patients. Morin brought the idea to developer Ace Sarich, vice president of VoxTec, a division of Marine Acoustics.

VoxTec landed seed money from DARPA to build a rugged, weatherproof, handheld translator. About 2,000 Phraselators are now deployed in Iraq and Afghanistan, where the device was first field tested in 2001.
Sarich says it helps remedy the chronic shortage of human translators who are often reluctant to work in the line of fire in a war zone.

"The problem with reliable translators is that they have to be knowledgeable in English and the target languages and not have their own political agenda," says Sarich. "Sometimes the military forces are frustrated because the translator does not want to offend people, but the military forces want to get their point across."

According to VoxTec, the Phraselator is a "cost-effective means of bridging the cross-cultural communications divide."

Sarich says military forces in Iraq use the device to provide information and issue commands at checkpoints, on patrol and inside detention facilities. Sample phrases include: "Get out of the vehicle." "Everyone stop talking." "Put your hands on the wall." "Space your feet." "We must now search you."

About five months ago, the U.S. Navy began developing a version of the Phraselator coupled to 70 highly directional phased-array speakers that broadcast a clear voice 300 to 400 yards, warning people to stay away from Navy ships.

"For homeland security, port patrol or general law enforcement, it is usually a one-way conversation and your responses are actions or physical affirmations," says Sarich.

Phraselator voice modules are typically stored on 128-MB secure digital cards that contain up to 12,000 phrases in four or five languages. The Phraselator Force Protection module now used by the U.S. military translates phrases into Dari, Pashto, Urdu and Arabic.

A toolkit allows soldiers to build their own custom language modules or download phrase modules from the Phraselator web portal, a database that currently contains more than 300,000 phrases.

The Phraselator is now being tested by law enforcement officials and corrections officers in Oneida County, New York, and in 10 other states. The device is also being evaluated in hospital emergency rooms and county health departments, where it is used to issue a set of standard diagnostic questions such as "Show me where it hurts."

The latest Phraselator model, the P2. was refined based on feedback from U.S. soldiers. It has a longer battery life, a directional microphone and an expanded library of phrases. The P2 still translates just one way from English to about 60 other languages, but it is inching toward full two-way voice translation.

According to Phraselator software developer Jack Buchanan, the accuracy of translating voice into text is above 70 percent. But the middle step of translating that text into a foreign language text before outputting the data again as voice is technically difficult.

"Taking into account cultural differences and context issues is an extremely hard problem," says Buchanan, who believes that developing something close to *Star Trek's* "universal translator" will be

harder than building the *Enterprise*. "When you are coming in and giving food to a village, how you would say 'hello' is totally different than if you are a military person at a checkpoint holding a gun pointed in their direction."

According to Buchanan, the Phraselator is now being programmed to translate limited two-way conversation where responses correspond to a specific domain of words like numbers, colors or dates.

The next generation of the devices will also feature pictures, allowing the user to ask, "Have you seen any of these people?" or "Have you seen these weapons?" The Phraselator is advertised on its website as an interrogation tool, but Sarich says it is inferior compared to human interrogators.

Douglas Jones is a researcher at MIT's Lincoln Laboratory which is helping the U.S. government develop baselines to measure the effectiveness of translation systems. Jones says he expects speech-to-speech machine translators to achieve incremental progress in limited domains and gradually expand two-way translation capabilities.

"The current level for text translation is about level two, which means that people are able to get basic facts out of a machine-translated newspaper article, but can't necessarily read between the lines," said Jones.

In 2003, DARPA estimated that open-domain, multi-task and unconstrained dialog translation was still five to 10 years away. But the research group developing IBM's MASTOR, or multilingual automatic speech-to-speech translator system, says its DARPA-funded bidirectional voice translator is a year or two from deployment.

According to Yuqing Gao, a member of the IBM team, MASTOR skips the small incremental steps and uses algorithms to extract the concept from each sentence and match it to a comparable sentence in another language.

"We have been working on Chinese for medical domains because Chinese is the most popular language and the potential number of users is huge," says Gao, who notes that the biggest challenge is analyzing emotional speech. "When people are very emotional or depressed, the speech signal is quite different, it's a very important step and without that function the usefulness can be limited."