Real Time Language Translation

Imagine eliminating the language barrier between native speakers of different languages. A 911 operator who gets a call from someone speaking in unknown language would be able to immediately translate the call and offer assistance in that person's native tongue, or a relief worker would be able to more efficiently meet the needs of those they serve by understanding what was communicated. Or a soldier would hear shouts of warning from partner soldiers of another country in a non-native language. With the advancement in real time language translators, all of the above is becoming a reality both through computer-aided textual support systems and through wearable devices that translate and provide real time language support. The benefits to elearning applications are unbounded. Being able to cross the language barrier when educating a global workforce or being able to offer just in time learning in that worker's native language not only would result in greater knowledge retention but could also be life saving in dangerous situations.

Just in Time Language Translation technology is currently available through a number of providers and is being used in a variety of different industries and environments. The following are some examples of its current application.

- The Consortium for Speech Translation Advanced Research (CSTAR) has developed and demonstrated hand held devices for tourists that provide geographical direction and speech translation of six different languages. (Carnegie Mellon, 1999)
- The Office of Naval Research has developed a fanny pack type wearable computer that not only translates the language but also is able to understand the context in which the term or idiom is used and how this influences the resulting translation. (Cleere, 2001) The US Army successfully used a similar device during operations in Bosnia. (Bass, 1998)
- IBM is developing a device that will be able to use visual cues such as the movements of the lip and mouth to understand and translate speech. While currently in beta testing and not yet ready for market, this type of device could be used within online customer relationship management applications allowing call center personnel to understand the unspoken mood of the customer by interpreting the body language. (Schwartz, 2001)

- IBM's WebSphere Translation Server is currently being used by many online enterprises to translate languages over the web at about 500 words per second. It can currently translate English, German, Spanish and Italian and translate English to Chinese, Japanese and Korean. (Schwartz, 2001)
- Ford currently use "Ernie" a web enabled software to support its automotive technicians. The software uses a natural language interface to not only understand conversational questions, but it is also able to follow conversation flow so when the human speaker refers to "it" in a following sentence, Ernie knows what "it" means. (Johnson, 2001)

Translating Body Language

When communicating with a speaker from a different country using translation software it is often difficult for the user to follow the conversation due to the disjunction between the speaker's facial movements and the words spoken. Being able to read the speaker's lips greatly enhances speech recognition. Researchers at Carnegie Mellon University in Pittsburgh are developing a device that might meld the two and allow users to hear what is being spoken in their native language and the body language simultaneously. This speech-synthesizer technology will translate spoken worlds from one language into another, even with the speaker and listeners are on opposite sides of the earth. They have developed "face translation" devices where the faces of the speakers are morphed by applying computer animation to the speaker's eyes and lips on a static photograph of their faces. This allows the user to see (for instance) Japanese speaker's lips move as if they were speaking English. (Guernsey, 1999) While not yet ready for market (further refinement of the visual display is needed) this type of device might eliminate miscommunication between different speaker of different languages as well as cultures that occur when the emotion and visual cues are stripped from spoken phrases.

Meaning is often lost when the user is unable to associate the body language of the speaker with the conversation. Penn State researchers have developed a prototype system that allows visitors to locate campus buildings and other points on campus by talking to a computer controlled map that response to hand gestures as well as speech. (Hale, 1999) Being able to read gestures and body language is essential to understanding the complexity and full meaning of the intention of a communicator's spoken words. Utilizing technology like Penn State has implemented can greatly enhance current language translation systems.

Further Challenges

Despite the considerable progress made in recent years, many outstanding research issues remain. For instance, speech recognizers are quite sensitive to the acoustic and linguistic properties of the data, and in particular to mismatches between the training and the real usage conditions. (Coretex.com) Translation systems must also cope with regional accents and homophones – words that sound the same but have different meanings. Mandarin Chinese, for instance, is the most complicated language to translate due to the fact that the language has four distinct tones that change the meaning of a word. (Fisher, 1999) The improvements that are necessary for efficient voice translation are clearly illustrated in the following example:

> Consider an actual spoken dialog between two Spanish speakers trying to agree on a time for an appointment. The following example shows a manually produced careful transliteration of the utterance, the way it was actually spoken by the speaker:
> "...sí sí el viernes diecinueve puedo sí porque sabes me voy de viaje d hoy la verdad así es que este mes es muy viajero me voy el día seis de viaje y estoy hasta el doce así que el día diecinueve me viene muy bien francamente..."
>
> Running this utterance through a commercial text translation system, the following translation results were obtained. (Note, that this would even assume perfect speech recognition):
>
> *yes yes on friday nineteen can yes because know I go me of trip D today the truth such is that this month is very traveler I go me the day six of trip and I am until the twelve as soon as the day nineteen comes me very well outspokenly* (C-Star, 2001)

While it is evident that more research and development is needed to create more effective speech recognition and translation devices, a variety of industries have shown that it can be very effective in enhancing worker's on the job learning and communication. Eliminating the language barrier between communicators will create a global environment of more efficient communication and information processing for many global enterprises. Voice and text translation technologies will enable enterprises to quickly roll out elearning applications globally without thought to translating the product into

different languages. Armed with systems that bundle wearable computing with instant voice recognition and translation, global field forces can be instantly prepared to respond to a variety of situations both domestically and in foreign language environments.

References:

Bass, Thomas. 1998. Dress Code. *Wired.com*. April Archives. www.wired.com/wired/archive/6.04/wearables_pr.html.
Provides an overview of the technology presented at two wearable computing conferences hosted by MIT Media Lab and the Institute of Electrical and Electronics Engineers Computer Society.

Carnegie Mellon. "Carnegie Mellon Scientists Orchestrate an International Video Conference Demonstrating Spontaneous Speech-to-Speech Translation in Six Languages" Press Release. 22 July 1999. http://www.c-star.org/main/english/cstar2/.
Carnegie Mellon announces a videoconference where they demonstrated their speech-to-speech translation technology and its uses in wearable devices.

Cleere, Gail. "Wearable translators". Press Release. 18 January 2001. *Office of Naval Research* www.eurekalert.org/pub_releases/2001-01/OoNR-Wt1801101.php.
Announces the Office of Naval Research's funding of a contextual voice translation device that would be outfitted with an additional military dictionary so that the system can comprehend military as well as civilian terminology.

Cortex. Retrieved via the World Wide Web on November 28, 2001. http://coretex.itc.it/public/frames/f_overview.htm
Describes the efforts of four speech recognition laboratories that have join their efforts to address some of the outstanding speech recognition and translation research issues.

C-Star3. Retreived via the World Wide Web on November 28, 2001. http://www.c-star.org/cstar3/orga_contact.html
Presents research findings and general information about C-STAR, a voluntary international consortium of laboratories around the world that are devoted to improving voice translation.

Fisher, Lucy. 1999. "Conversation Piece". *Time.com* 11 October. Vol. 154: 15.

Fisher explores the state of real time translators –how reliable they are, the technologies involved and the research institution and enterprises involved in their development.

Guernsey, Lisa. “Toward Better Communication Across the Language Barrier”. *New York Times.* 29 July 1999. www.nytimes.com/library/tech/99/07/circuits/articles/29tran.htm. Describes a day long video conference held at Carnegie Mellon University where their face translation software was debuted for laboratories in Germany, Italy, France and Korea.

Hale, Barbara. “Computerized Map responds to speech and gestures”. Press Release. 19 August 1999, *Penn State*. www.eurekalert.org/pub_releases/1999-08/PS-Cmrt-190899.php Describes Penn State’s newest development in computerized gesture and speech recognition software that they have embedded in a campus map.

Johnson, Amy. “Web Tool Makes for Good Conversation”*.* *Computerworld*, 26 February 2001. p60.
Johnson describes the features of Ford’s new “Ernie” software that utilizes natural language interface and logical conversation tracking to answer naturally phrased verbal questions from ford technicians over the web.

Schwartz, Ephraim. “IBM Ramps up Speech Technology Products and Research”. *InfoWorld*. 25 June 2001.
Schwartz examines IBM’s voice translation and speech recognition technology that have been recently announced or are still in development.