Machines break the language barrier

As overseas trade expands, computers quickly translate reams of technical information

Newton, MA—Human translators, adept at interpreting the nuances of literary writing, often founder when faced with scientific definitions. So it should be no surprise that machines are translating more and more routine engineering and business information.

Already, software packages for PCs, minicomputers, and microcomputers efficiently translate product manuals, patent data, brochures, and manufacturing instructions. A few instant audio-to-audio devices have emerged for simple conversion translation.

The next 10 to 15 years, however, promise a much wider role for machines in everyday, worldwide interpretation. The European Economic Commission estimates the machine translation market has the potential now for $10 billion in business, with a 10% annual growth rate. The product of everyone's dreams: a hand-held computer that instantly translates conversations or elaborate writing with 100% accuracy.

Can computers meet the task?

Skeptics insist that machines will never match the performance of humans—especially when you consider there are 72 major languages and more than 1000 dialects. Not only are there many thousands of word-for-word definitions to match, there are fundamentally different ways to present ideas.

For example, the gulf between English and Japanese—as separate as languages can get, yet among the most profitable to bridge—pits the word against the kanji. These ideographic characters stand for anything from a word to an idea. There are 2000 of them, many with only subtle distinctions.

Adding to the complexity, many languages use tradition, context, or inflection to distinguish between seemingly identical statements. An infamous machine translation error demonstrates the difficulty. When asked to translate the adage "Out of sight, out of mind," from English to Russian and back again, an early machine translation program produced: "Invisible insanity."

Intermediary language used

Currently, two forms of machine translation exist, but only one can avoid these kinds of potentially disastrous mistakes. In one method, text is replaced word for word from one language to the next, with some errors. The other more accurate method replaces entire concepts with similar concepts in another language by using an intermediary mathematical language of ideas.

Rather than depend on memory banks to substitute one word with a synonymous version from a second language—an elaborate search-and-replace routine—the intermediary language attaches mathematical representations to whole concepts or thoughts. This way, computers can almost "understand" what's intended. As a result, the program chooses words or symbols that best communicate the original concept or thought structure.

Once a document is entered into the computer, for example, the process reduces the original language to the mathematical forms and creates an intermediary language of concepts. This middle step is based on artificial intelligence and parallel processing research. From there, the information can be processed back into sentences or symbols of the target language. The two-step process requires more computational power, but the results are impressive.

Schematic of thought

An intermediary language called Interlingua has been developed by researchers at Carnegie-Mellon University, Pittsburgh. So far, the technique has laid the foundation for an accurate machine translator for Japanese to English, and vice versa.

"Interlingua is to language what a blueprint or schematic is to a circuit," explains Jaime Carbonell, director of the Center of Machine Translation at Carnegie-Mellon. "It's more important to have the schematic representation of
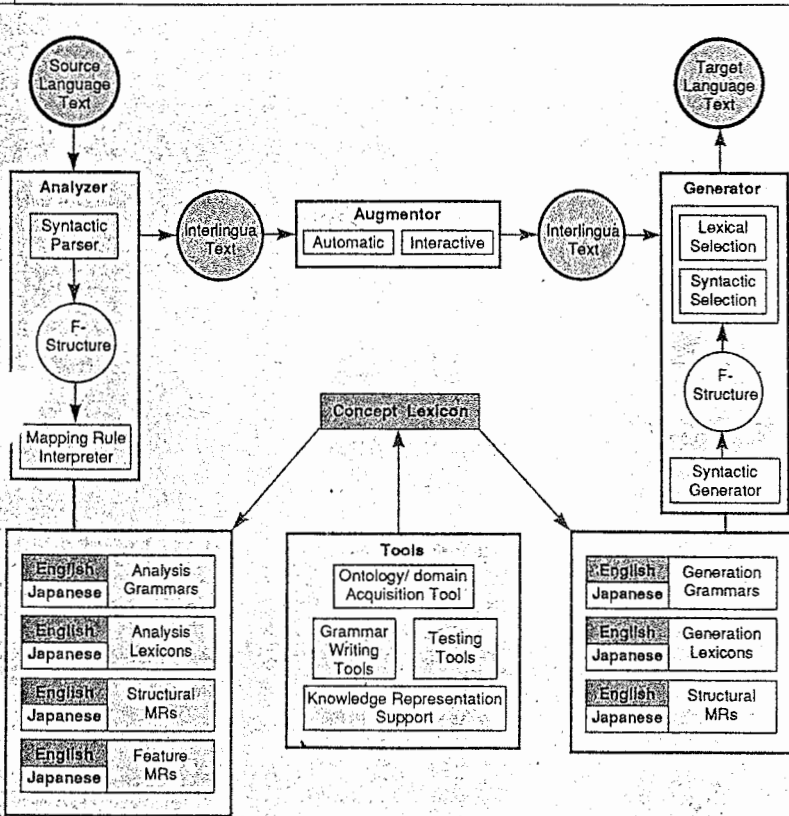
Carnegie-Mellon's complex machine *translation process converts human concepts to knowledge-based computer structures to avoid mismatched words.*
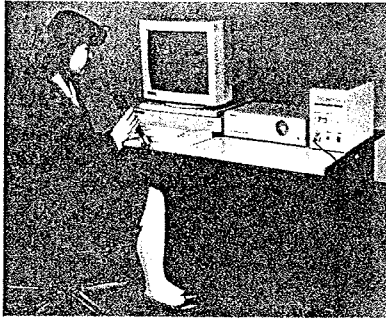
Desktop speech recognition *and voice generation translation could be reduced to a hand-held, yet highly accurate, interpreter in 10 years.*

Speech-to-speech translation *research by Carnegie-Mellon's Hideto Tomabechi and Margaret Zabludowski (right) helps make interpretation almost instant.*

the circuit than a photograph when you try to duplicate it on another job."

Logos Corp., Dedham, MA, has commercialized the intermediary language method. Using what it calls Semanto-Syntactic Abstraction Language, the firm has developed software packages for several language sets, mostly European, which use Unix to run on IBM and Wang computers. Logos programs, which cost between $5000 and $8000 a month to lease, can be tailored to specific industries by using special glossaries and subroutines.

Logos's biggest customer is the U.S. Department of Defense, but the majority of its clients are large, high-technology companies—Hewlett-Packard, Wang, IBM—looking to branch into foreign markets. Some firms prefer machine translation for security reasons, since trade secrets can be leaked when outside translation firms are used.

Like Logos, Worldwide Communications, Chicago, has marketed machine translation. It offers two packages: MacroCat, which runs on eight pairs of languages on minicomputers, and Micro-Cat, which operates on seven language pairs on microcomputers. Both systems use the word-for-word method of translation, and are used primarily by manufacturers for long instruction texts.

Advanced Products and Technologies, Redmond, WA, plans to sell a handheld, electronic interpreter for European languages this year. The $2000 product does not use Interlingua and is limited to small talk. Travelers and businessmen in foreign countries, as well as some hospitals with foreign-speaking patients, are targets for sales.

Generation two: voice recognition

Voice activation has also sparked research at such firms as Digital Equipment Corp., GTE, and Wang, as well as the National Institute of Standards and Technology (NIST). Working with the Defense Advanced Research Projects Agency, NIST researchers are developing improved algorithms and software for phonetically based recognition of speech as well as ways to measure the performance of automatic speech recognizers.

Some of the biggest names in Japanese industry—Fujitsu and Matsushita among them—are investing heavily in their own machine translation technology. Indeed, the machine translation research at Carnegie-Mellon is funded in large part by these companies.

Matsushita of Osaka, parent to Panasonic, Technics, and Quasar, has developed a Japanese-to-English interpretation system on a 32-bit workstation that can audibly recognize speech from any speaker. Previously, high-speed speech interpretation was possible only with a mainframe computer using a single "pre-registered" voice.

Fujitsu Ltd., Tokyo, has also developed a voice-activated machine translator for PCs based on the firm's ATLAS translation method. Similar to Interlingua, it can only go from Japanese to English. However, Fujitsu has had success with a Japanese-to-Korean machine translator that circumvents the intermediate language, thanks to the similarity between Japanese and Korean.

Markets demand 100% accuracy

Logos claims its programs are the most advanced on the open market. Using a human editor to screen out ambiguities, reduce language to direct sentences, and eliminate confusing terms, its programs can attain 90% accuracy, says Jon Cave, marketing director.

"The best measure of accuracy is how much productivity improvement you gain," says Cave to his critics. "Our customers are accustomed to two to three times more productivity than human translators."

For many, that's still not good enough. "Machines don't do a sufficiently good job," says Dennis Hennessy of American Translators International, Stanford, CA. "Strictly technical stuff comes out okay, but verbiage comes out wooden. The further away you get from the technical, the worse it gets."

In the future, however, translation firms may become the machine translation industry's best customers. "The optimal use of the machine translation product will be for a translation service," predicts Robert Sprung, president of Harvard Translations, Boston.

Sprung believes that machines are appropriate for a limited number of large-batch translation jobs, and that their currently limited role may soon broaden. He is also confident that machines will not put him out of business.

"There will never be a perfect translation machine," he insists. "You're always going to need a sophisticated editor, and that will make the role of editing all the more important." □

—Dana Gardner, Associate Editor