

[From: *Wired* 8(5), May 2000; pp. 234-235.]

Universal Translators- A look at the hubs for machine translation R&D worldwide.
Compiled by Carl Zimmer (zimmer@panix.com)

**Natural Language Laboratory, School of Computing Science, Simon Fraser University
Burnaby, British Columbia**

Researchers are devising "relaxed grammars" to extract the sense of transcribed speech through a method known as "partial parsing," which deciphers chunks of language rather than breaking down the structure of entire sentences. Their work is being integrated into closed-captioning technology for real-time TV translations.

**SRI
Menlo Park, California**

SRI, a not-for-profit R&D center and Telia, a Swedish telecom, have built a speech-to-speech translation system that takes only a few seconds to turn Swedish into English, and vice versa. Currently, you can talk to the program only about booking airline flights. But the underlying software could be applied to other situations where the lexicon is relatively finite, such as in hospitals or banks.

**Integrated Wave Technologies
Fremont, California**

IWT has designed a new tool for law-enforcement personnel called the Voice Response Translator. The portable computer attaches to a police officer's belt, receives English questions or commands through a lapel microphone, and then broadcasts a translation. Rather than translate word for word, the device automatically recognizes 150 prerecorded phrases that range from "If you speak Vietnamese, raise your right hand," to "Police! We have a search warrant!"

**Information Sciences Institute, University of Southern California
Marina del Rey, California**

Example-based MT machines must be trained with a steady stream of text input. The ReWrite project at ISI seeks to break free of this constraint by programming computers to compare original and translated texts, then use that knowledge to teach themselves to understand the new language, using only original texts. ISI also uses techniques borrowed from cryptanalysis, where machines approach languages as codes to be cracked.

**Laboratory for Applied Research in Computational Linguistics, Université de Montréal
Montreal, Quebec**

This Canadian research lab is improving software to assist human translators. Its program, TransType, lets a translator begin typing a translation, then pause so the computer can take over, tapping its own databases to complete words and phrases.

**Dragon Systems
Newton, Massachusetts**

In 1997, Dragon Systems and Darpa outfitted US soldiers in Bosnia with a computerized Serbo-Croatian phrase book. Dragon is now testing its AudioMining technology, which transcribes audio files into time-indexed searchable text.

**Spoken Language Systems Group, Massachusetts Institute of Technology
Cambridge, Massachusetts**

The SLS Group has created a telephone-based weather guide that they've expanded to other apps, such as airline schedules, traffic reports, a dining guide, auto classifieds, and voice-driven Web searches. The technology is based on a

conversational-speech-recognition platform designed to overcome problems that commonly plague other systems, including restricted vocabulary, atypical pauses, and ambient noise.

**Language Technologies Institute, Carnegie Mellon University
Pittsburgh, Pennsylvania**

LTI's translation system, Diplomat, learns by example; it can make rough translations of a foreign language after only two days of lessons. By the end of 2000, LTI and Lockheed Martin will have built wearable translation computers for field-testing by the US Army. Also under development at LTI is the knowledge-based accurate translation (KANT) system for technical documents.

IBM

Armonk, New York

IBM is developing machine translation components that will let servers automatically translate Web pages between English, German, Italian, Spanish, Portuguese, French, Korean, Japanese, and Chinese. Other ongoing projects include a program for transcribing news broadcasts in real time, with intelligent software that filters out ambient noise and accompanying music.

AT&T Labs

Florham Park, New Jersey

AT&T researchers are developing a Very Large Vocabulary Speech Recognition database, which can handle more than a million words - the world's largest for a continuous-speech dictation system. The company has partnered with Advanced Telecommunications Research in Kyoto, Japan, to design and build an automated telephony system for translating queries between English and Japanese.

Systran Software

Soisy-sous-Montmorency, France

Founded in California in 1968 by Georgetown linguist Peter Toma, Systran is one of the oldest MT companies in the world. It started by supplying the US military with software that translated Russian documents. More recently, Systran developed Serbo-Croatian-to-English software for US forces sent to the former Yugoslavia. Right now, Systran software can translate text between 16 language pairs, and more are being added.

Lernout & Hauspie

leper, Belgium

L&H is the Microsoft of the translation world. It markets translation software, provides customized MT packages, and sells speech recognition technology for dictation. The company also invested \$100 million to set up 10 industrial R&D parks in Israel, South Africa, Norway, and other countries.

German Research Center for Artificial Intelligence

Saarbrücken, Germany

The German Research Center for AI leads a consortium of 23 universities and corporations for the \$80 million Verbmobil project. Funded by the German government and industrial partners, the project is developing software for translating German to and from Japanese or English. Verbmobil is limited to travel planning, but researchers hope to convey the emotional tone of speech into translations for tasks like expediting requests from irate airline passengers.

Prompt

St. Petersburg, Russia

Prompt (Project MT) introduced the first commercial MT system in Russia and the first English-Russian text translation software to run on Windows. Prompt's R&D remains strong, with newer products for translating Russian into English, German, French, Spanish, and Italian - and vice versa.

**Laboratory of Computational Linguistics, Institute for Information
Transmission Problems, Russian Academy of Sciences
Moscow, Russia**

The Laboratory of Computational Linguistics ETAP-3 prototype is a multifunction system capable of machine translation, interfacing with databases using natural language, and assisting in the resolution of syntactically ambiguous translations. LCL researchers are also using ETAP-3 to bring Russian-language capability into the UN's Universal Networking Language.

**Research Center of Computer & Language Information Engineering,
Chinese Academy of Sciences
Beijing, China**

The Chinese Academy of Sciences is the largest in a crowded field of research institutions on Chinese-language MT. Among its many projects is a collaboration with IBM to merge its Chinese-to-English system with Big Blue's speech recognition software.

**Media Research Laboratories, NEC Corporation
Kawasaki, Japan**

NEC is perfecting a Japanese-English speech translator that can run on an ordinary PC. The system is limited to conversations about travel. Last fall, NEC showed off a prototype that offered real-time responses to travel queries using Windows NT and a Pentium II. The company hopes to release a commercial version in a few years.

**Universal Networking Language Center, United Nations University
Tokyo, Japan**

The UN is trying to decentralize machine translation with the Universal Networking Language. Developers will create UNL-embedded Web pages with special software. The user's corresponding browser add-on will then spot this text and request a translation from a UNL server. Researchers unveiled the system - which can translate between 15 languages - in April. By 2006, it should be available in every language of the UN's 185 member states.

**NTT
Kyoto, Japan**

In 1998, NTT released Altflash, a system for translating between Japanese and English and the first of its kind that doesn't rely on a human translator to check for mistakes. Altflash is limited to financial reports, but NTT is trying to develop systems that can handle other kinds of text and speech.

**Spoken Language Processing Group, Advanced Telecommunications
Research
Kyoto, Japan**

In December, ATR, supported by industry and government funding, demonstrated its latest prototype: Two people met on a sidewalk, dialed up ATR's computer on their cell phones, and conducted a conversation (limited to travel planning) simultaneously in Japanese and English. ATR is now working on commercial apps for real-time TV translations.