

July 1988

TACnology Corner

Getting Together to Get it Together

Our readers will be interested in hearing about two conferences last month on the subject of computers and translation: the Second International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (Carnegie Mellon University, Pittsburgh, June 12-14), and the Conference on Teaching Computers and the Humanities Courses (Oberlin, June 16-18). Herewith are some highlights.

Carnegie Mellon

The meeting in Pittsburgh addressed a broad variety of theoretical approaches to different MT problems, most of which rely heavily on semantics and knowledge-based decision-making. The climate was open-minded: there was a general recognition that no single linguistic theory is sufficient to solve all the problems that need to be dealt with. This attitude was coupled with a call for increased cooperation between research projects, with pooling of resources where possible, and an appeal for greater attention to the lessons of history so that past mistakes won't be repeated.

Aspects of a number of systems, both venerable and experimental, were presented. There was a strong Japanese presence, with 10 speakers reporting on developments that involve Japanese and English.

An area in which Carnegie Mellon is quite active is "real-time interpretive MT," which involves speech recognition in the source language and speech synthesis in the target--with MT, of course, in the middle. Topics related to this technology were presented in a

special panel, and there were two demonstrations of speech recognition. In the first, a short phrase was spoken in Japanese, translated, and the result synthesized as English speech. The lexicon and syntax were highly constrained, and not all the speakers who tried the system could be "understood." The two successful outputs were "I have a headache" and "I have a pain in the back." The second demo involved the recognition and synthesis of English only. The recognizer was trained on a set of about 40 sentences and proved to be quite speaker-independent. It gave good results even with foreign and Southern accents.

The meeting concluded with a roundup by Professor John Hutchins, author of *Machine Translation: Past, Present, Future*, who identified some

"In certain highly structured types of discourse it may be possible to generate texts spontaneously in multiple languages...."

broad trends in the evolution of MT, most of which had come to light in the course of the conference. In terms of languages being developed, there is now a strong emphasis on Japanese, and the direction has shifted away from *into* English to *from* English into other languages. The emphasis on Japanese is making semantics and knowledge bases more important than ever. He feels that the quality of output has improved over the last 30 years, although not by very much since 1976. The improvement is more in the range of texts being translated than in translation quality as such. In any case, the "try-anything" approach appears to be yielding to sublanguage systems focused on specific subject areas. Whereas in the beginning emphasis was on scientific texts, today MT is used at least as much if not more for commercial and administrative applications. With an increasing volume

of machine-readable text already available, MT becomes feasible as a source of "information-only" translation. In certain highly structured types of discourse it may be possible to generate texts spontaneously in multiple languages, i.e. without any source text to begin with, from a common data base. Interactive MT is being seen as a possible tool for disambiguation--and perhaps not all ambiguities need to be resolved, since they may be intended in the original. Of course, hardware and programming languages are infinitely more sophisticated than they were in the early history of MT, and we now have high-order languages such as Prolog and Lisp that are specially geared to decision-making for natural languages. And finally, he said, MT undertakings are becoming international, with teams from different countries joining together in large common efforts.

Oberlin

Perhaps a majority of the participants at the conference on computers and humanities courses were from the fields of English, foreign languages, and/or linguistics. Other were doing related work on cognition from the perspectives of psychology, philosophy, and artificial intelligence (AI). The meeting's interdisciplinary character made it both refreshing and stimulating.

A special panel was devoted to the teaching of courses on computers and translation. The first paper (mine) emphasized the increasing need for preparation in this area within academic translation programs. After identifying three types of competence that need to be developed (word processing, terminology management, and machine translation), I outlined briefly the course I have been giving at Georgetown University for the last 8 years. Ingrid Meyer, from the University of Ottawa, then discussed the more advanced of two courses that she gives in M(A)T. She described ways in which the translator can be prepared for involvement in machine translation

and identified levels of increasing complexity and commitment at which this participation can take place. Josette Coughlin, whose course at Georgia State University currently focuses on the use of MT and MAT as tools to speed up the translation assignment, called for a course for all graduate-level language specialists in the theory and use of grammar-, semantics-, and knowledge-based computer systems.

A number of the participants at Oberlin commented on the tendency for humanities students to be "self-selected runaways" from technology-refugees who are now in their last redoubt, face-to-face with their most feared demon. This impression was confirmed by statistics on word processing competence presented by Behar and Hackett in a population of undergraduates at Dowling College in Oakdale, NY: among entering students in the business program, 72% already had some knowledge of WP and 94% expected to use computers in their future work, whereas of the liberal arts majors only 22% had had any experience with WP and a considerably smaller proportion (60%) foresaw a possible role for computers in their career. With regard to the translation program at Georgetown, I reported that the number of entering students with experience in WP is steadily increasing, yet in the fall of 1987 the proportion was still only 27% (M. Bowen, p.c.).

Incidentally, according to the results reported from an early 1987 poll of humanities computer-users, 80% of the respondents were working on IBMs, not including clones. It was pointed out, however, that the Macintosh has made strong inroads into the market in the last year. Some heavy-duty users prefer AT&T computers because of the UNIX operating system.

TACnology Corner welcomes news of courses on computers and translation, especially those geared to the would-be or practicing professional, and looks forward to covering the subject in greater detail in future columns.

--Muriel Vasconcellos, Chair
Committee on Translation and Computers