

## MT AT TEXAS: THE EARLY YEARS

W. P. Lehmann  
Linguistics Research Center  
The University of Texas at Austin  
Austin, TX 78712  
lehmann@mail.utexas.edu  
Tel: (512) 471-4566

On my return from directing the Georgetown English Language Program in Ankara Turkey during the years 1956-1957, Professor Léon Dostert suggested that I undertake machine translation, with concentration on German, since he was doing Russian. Having spent World War II at the predecessor of NSA dealing with translation of Japanese messages, where forerunners of the computer were lined up by scores, I had not been without interest in Weaver's suggestion that the newly developed instrument be employed for MT. But other obligations interfered with more than merely keeping abreast of reports on activities. Dostert sweetened his suggestion with a grant of \$10,000, a sum large enough, in the days when professorial salaries scarcely reached half as much, to employ a good group of graduate students.

The work consisted chiefly in learning what was going on and planning. The University's computer, an IBM 650, was preempted by physicists, for what it was worth. Our best informed member in computational areas was Nick Hopkins, a graduate student in archeology, who among other capabilities had learned on trips to Mexico how to make sandals by cutting strips from old tires and keeping them on with pieces of string. Most of the other members were linguists well acquainted with German. The best result of the year was an award of \$300,000+ from the U.S. Army Signal Engineering Laboratories to investigate the feasibility of MT. The Army was planning to develop a computer, Moby Dick, to be located at division headquarters; MT was to be one of its functions.

The grant provided means to employ full-time programmers as well as linguists. One of these, Gene Pendergraft, had worked with Marvin Minsky, and soon became the chief systems specialist. We set up three groups: linguists, programmers, and mathematicians. Sponsorship by the Signal Corps gave us the advantage of using the IBM 709 at Fort Huachuca. Programs and translation algorithms were developed, prepared by our keypunchers for entry, and taken out to Arizona to be checked. Our progress was recorded in quarterly reports. The third, dated 31 January 1960, and the ninth, dated 31 July 1961, state the theoretical basis of all our further work, which led to the production of the METAL translation system. Following the theoretical position of Charles Sanders Peirce, we treat language as a system with three components: syntactics, semantics, and pragmatics. Programs were designed to manage the linguistic rules and were kept distinct from the grammar and the lexicon.

Through personal contacts, conferences, and publications we kept in close touch with other groups and profited by their conclusions. Leon Dostert generously sent down his specialists to inform us of progress at Georgetown. Victor Oswald at UCLA credited C.V. Pollard in our department with developing syntactic rules for scientific German, as did Yehoshua Bar-Hillel (1964:159), at a time when language teaching concentrated on morphology. Erwin Reifler of the University of Washington developed the notion of sublanguages, basically an application of pragmatics that recognized the presence of distinct types of language; MT could concentrate on technical language rather than attempt to control such sublanguages as that of literature. Our progress was recognized by further grants, and by the founding of the Linguistics Research Center in 1961.

The conferences kept us in close touch with members of an expanding technology. Léon Dostert remained the operational father of MT. Known at the highest levels of government through his wartime duty as French interpreter for General Eisenhower, when funds were running out Dostert would simply call on Allen Dulles, then head of the Central Intelligence Agency, who would turn him over to his staff with the injunction to “give Léon what he needs.”

Victor Yngve must be credited with producing the first appropriate computer language, COMIT, and with producing a journal for the field. Yehoshua Bar-Hillel, a logician who might be called the theoretical father of MT, insisted on FAHQT (fully automatic high-quality translation), then became highly pessimistic and vocal about its realization. Curiously, one of the major proposed shortcomings was the need for post-editing, a regular procedure among translation agencies. His unfortunate insistence, clear still in the “Feasibility Study on Fully Automatic High Quality Translation” produced by the Center in December 1971, resulted in part from ignorance of translation procedures, as did the devastating National Academy of Science report of 1966 by the engineer John Pierce—otherwise known as the “ALPAC Report.” By contrast, the far-sighted supporter Zbigniew L. Pankowicz, of the Rome Air Development Center, may be credited with keeping the technology from vanishing by providing the funding for LOGOS and the Center.

In recollection, the measures required to achieve results hardly seem credible. Access to the computer was through punched cards; to cut costs, Dostert established a keypunching center in Frankfurt, Germany, where labor was cheaper. Rules and algorithms were transferred to tapes, which whirred crazily in the computer room. In the absence of computer languages, coding was done initially in machine language, then in Fortran. When the University acquired a top-notch computer, the CDC 1604, we used it on Sunday mornings, because testing our programs required full use of the machine. Somewhat later the Center acquired its own, a lesser CDC, so that our systems personnel no longer had to defile the Sabbath. But the pitiful success provided as much assurance that MT was feasible as did Newton’s apple for his views on gravitation.

Yet Bar-Hillel’s vigorous negativism, coupled with unethical activities, succeeded in undermining success. As Bar-Hillel stated (1964:9-10), funds for MT were siphoned off, not least in the institution that was his home, so that the amount invested was larger than that applied. Another damaging effect was produced by “demonstrations” that had been pre-cooked. Even the self-sacrificing support provided by Pankowicz could not offset the devastating effect of the shameful ALPAC Report of 1966, which eliminated almost all research in this country and abroad. Fortunately, with the support of Pankowicz, the Center was able to continue its work until Siemens found that their human translators could not keep up with the demands of quantity and time. Their funds and that of other U.S. agencies maintained research at the Center so that METAL became an operative system, which Somers has credited as “a success story in the development of MT” (Whitelock: 1995:198).

#### REFERENCES

Bar-Hillel, Yehoshua. 1964. *Language and Information: Readings*. MA: Addison-Wesley.

Whitelock, Peter, and Kieran, Kilby. 1995. *Linguistic and Computational Techniques in Machine Translation System Design*. London: UCL Press, See my review, *Machine Translation* 12:261-69, 1997.