

---

## **Translation & Computers**

### **Future Directions for JE MT**

The U.S. Government, which turned a deaf ear to progress in machine translation for approximately 20 years, is finally getting back

into the act. Prompted by the growing commercial implications of Japanese technology and the awareness that Japan is heavily committed to the development of MT, the National Research Council's Office of Japan Affairs held a one-day Symposium on Japanese-to-English Machine Translation at the National Academy of Sciences in Washington, D.C., on December 7, 1989. There was a dual irony to this initiative: not only the date chosen for the event, but also the fact that NRC was precisely the Government agency that published the famous ALPAC Report which in 1966 managed to relegate MT research to two decades of doldrums.

Despite the omens, however, the Symposium was an unqualified success. It was charged with energy and provided just the right forum for an exhilarating exchange of views on the issues that face both the capture of information in Japanese and the further development of MT--not only JE but in general. The success was due in large part to excellent organization on the part of NRC's Martha Harris, director of the Office of Japan Affairs, who selected speakers from across a wide range of disciplines and perspectives and then focused their presentations by furnishing them ahead of time with an insightful set of "stimulation questions." These valuable questions were worth the price of admission. Less important than the ultimate answers were the ideas that they brought to the fore.

#### **State of the Art**

The first panel was aimed at assessing the current status of MT technology not so much from the static view of output quality but more from dynamic perspectives such as different levels of consumer needs, trade-offs between time and precision, and the "hidden" demand that

MT might meet. The stage was set by Superguru Makoto Nagao, who pointed out that all criteria are elusive and subjective--that not only quality but usability may be viewed quite differently by ordinary people. The real demands for translation are yet to be clarified. The usability of MT will vary depending on the ultimate purpose of the translation and the characteristics of the particular source text.

The panel also examined the problems that remain to be solved before viable MT systems can be put into widespread use. Professor Nagao called for automatic self-learning mechanisms and grammatical rules and parsers that take multiple criteria into account simultaneously, and he emphasized the need for research on anaphora, ellipsis, and topic and focus. The search for these solutions may be aided, he believes, by automatic analysis of bilingual corpora, automatic collection of terminology, and the construction of computer-based thesauri.

Other panelists emphasized the importance of knowledge bases for the "understanding" of text and the development of statistical weighting formulae to assist in automatic decision-making. There was discussion of whether the bilingual process of translation might be divided into tasks that can be done by monolinguals, and strong views were presented on both sides of the fence.

#### **Market Prospects and User Needs**

The next panel addressed the issue "Who needs, wants, or will pay for JE MT 10 and 15 years from now?" It attempted to estimate the size of the U.S. and global markets for MT and consider how the potential, untapped demand might be measured, bearing in mind that the availability of effective MT in and of itself may well generate more demand.

Underlying these questions was the more basic one, which kept coming up throughout the day's discussions, of how much information now existing in Japanese is actually needed by consumers in the United States. Is there a vast quantity of critical and urgent importance, or is the information which is publicly available already stale and repetitious? The answer appeared to lie between these two extremes, but perhaps the real issue is how to distinguish between them.

The panel also attempted to characterize the types of MT that would be cost-effective, and it looked at the problem of how further MT development can be funded.

Congressman George E. Brown, Jr., joined the Symposium just before lunchtime and addressed the last-mentioned issue, namely the funding of technological development. One of the big problems is continuity and the need for decision-makers, both corporate and congressional, to shift attention from year-end profit lines to steady support for the attainment of goals that can only be achieved in the long term.

The third panel looked at how MT is currently being used and by whom, what the real experience has been, and what, from the users' perspective, most needs to be done to improve the quality of MT systems. What are their top priorities? It addressed the involvement of users in the development process and examined mechanisms for communicating the needs of users to developers. Participating on this panel were ATA members Maria Russo, who discussed the Systran experience at Xerox, and Alan Melby, who emphasized the need for a range of options for translators.

#### **R&D Policy**

Implicit in the discussions throughout the day was the need for

rational decisions about future MT developed based on input from all sides-- MT developers, providers, and users. This point was made in particular by the vendors--Tom Seal, president of ALPNET, and Bud Scott, representing Logos. The solution is a trisectoral approach that will involve the public institutional sector, the private entrepreneurial sector, and the user community. The idea of a consortium was proposed, as was an international and an American regional association for machine translation. Professor Nagao circulated a detailed prospectus for the International Association for Machine Translation, the text of which will be published in this column in the near future.

It was pointed out that MT is a fragile technology whose development has been impeded by lack of a coordinated policy and lack of steady resources to keep development teams going. The U.S. Government was called on to re-examine its commitment to MT technology--not with the gusto or naiveteé of the 50's and early 60's but with a pragmatic acknowledgment that times have changed. This is important both because of the need to capture current scientific and technical information in foreign languages and because of the need to launch U.S. products overseas in multiple markets simultaneously. The Government can help to: focus investments in areas of highest priority, avoid unnecessary duplication, share information, identify resources, develop technical criteria, and, possibly, reward excellence.

#### **Maybe MT Is Not the Answer After All**

Every effort was made in the course of the Symposium to give due weight to the arguments of those who felt that MT should be embraced with extreme caution. Martin Kaye emphasized that current MT does not take into account what really hap-

pens in translation. James Unger saw the MT problem as insurmountable, given the need for inferencing, and called for increased spending on education for translators. A more moderate view was taken by Takehiko Yamamoto and myself, both of us MT developers and producers of translation, who recognized the need for inferencing but felt that the human user can be trained to effectively fill the gaps. Yamamoto likened MT to a Japanese garden: the stones make for the tangible structure, but their meaning is given by the spaces between them.

- Muriel Vasconcellos, Chair  
Committee on Translation and  
Computers