

FROM THE HORSE'S MOUTH: A PRACTICAL GUIDE FROM JAPAN'S MACHINE TRANSLATION GURU

"Machine Translation - How Far Can It Go?" by Makoto Nagao (translated by Norman D. Cook, Oxford University Press, 1989) Price: UK£25.00 (US\$41)

By Andrew Joscelyne

In Japan, the author of this book is popularly known as *kakai hon'yaku-san* - Mr. Machine Translation - an epithet he deserves for his single-minded dedication to running MT projects, training MT engineers, coordinating across-the-board work in the field and liberally sharing it with others, whether with the French MT group at Grenoble in the past or via international computational linguistics organizations today.

Nagao's students from his Kyoto University engineering department were responsible for the research that led to the first kanji wordprocessor, and most Japanese work on MT systems has been inspired by his courses on designing semantically-based transfer systems.

In explaining what MT is to the Japanese business community - the original 1986 target readership for this short work - Nagao pursues the implications of his own research to its logical conclusion: making operational translation systems available to Japanese manufacturers with vast export markets. As an answer to the title's own question, Nagao's account suggests a reasonably positive "far enough to be worth investing in."

This English version aims to satisfy the need for a good overview of MT from the point of view of genetically disparate languages, rather than trying to address

the more practical problems that Nagao's original public was confronted with.

Japanese and English are presented as the epitome of dissimilar languages, posing translation problems that a solely European perspective cannot confront.

In his Preface, for example, Nagao argues that pivot language architectures for MT systems as developed for Indo-European languages (e.g., the Eurotra program) will fail to handle the specifically Japanese - or Thai? - markers for respect and politeness if they do not have a sufficiently rich conceptual structure in the interlingua.

This would tend to exclude the rationale for Esperanto-type pivots for Japanese-European MT systems, since the interlingua simply doesn't include culture-specific politeness concepts. It will be interesting to learn if BSO's corpus-driven DLT translation rig, which makes use of Esperanto as an interlingua, proves him wrong.

NO EASY MATTER

The body of the book consists of a straightforward account of precisely what the title suggests. Combining a historical introduction to MT efforts with brief descriptions of the overall architecture of such systems as TAUM Meteo and Eurotra, Nagao prepares the way for four chapters on the real problems involved in MT design.

This takes us into some fairly detailed language analysis, complete with tree diagrams and conceptual networks. The exposition is clear, the "I want an apple" type examples are simple and effective for Nagao's purpose. And he is at pains to suggest that building systems to account for the complexities underlying language understanding is no easy matter, even though it can gradually be achieved in specific textual areas such as technical documentation or scientific papers (his own favoured field and the source of several authentic MT examples in the book).

For the Japanese salaryman who



knows a bit of English, the examples of sentence structure used to illustrate the problems and processes are probably fairly easy to follow - they cover such features as the relatively free word order of Japanese, cryptic gender and number distinctions, and the thematic markers *wa* and *ga*, which do not get realized at the morphemic level in English.

For the English reader, however, the publishers have decided that the linguistic - and more particularly semantic - "distance" between English and Japanese makes for a more "interesting" perspective on MT. As a result, every example in the slim book ends up filling half a page: the Japanese sentence is given in the original script (kanji and kana), with a capitalized transliteration, a literal translation, and finally a translation into acceptable English.

The presumably intended effect is that of actually showing how a source sentence would gradually be processed by a machine until its target language sense and structure emerge - a neat idea, but repeating it each time strikes me as redundant.

Problems and solutions for sentence structure eat up the bulk of the space devoted to the mechanics of MT, and by the time Nagao reaches the general discussion of software development, dictionaries, and other "peripheral" aspects, he offers only cursory coverage, leaving the general reader no doubt pretty mystified about what a program that converts sentences into a sequence of tree structures is all about.

OUT OF THE CLOSET

An intriguing reference to "interactive" MT systems - the user being questioned by the system in order to clarify his/her intentions in writing - is the only insight we get into actually handling these systems, even though the example - based, I believe, on research by Jun-ichi Tsujii, now at the University of Manchester Institute of Science and Technology (UMIST), England - is atypical, since most MT screen editors provide simple menu-driven utilities for dictionary updating and rule addition.

Something more on hardware would doubtless have given the reader a clearer picture about running actual rigs and how they engage with human operators. Nagao, however, refrains from mentioning a single proprietary Japanese brand, so we get no details on such ergonomic features as runtime orders of magnitude, throughput ranges, and the like.

Nagao the trained engineer is realistic about the future of MT. He claims that linguists in Japan just aren't interested in machine approaches and that as a consequence, the spread of MT systems is a technology-driven affair. In a recent interview, he claimed that in two or three years, the quantum leap in Japanese workstation use would make MT so widespread in that country (see LT14 for a report) that the business and technical communities would be "forced to use it" - with the knock-on effect of modifying

technical writing habits to fit the machinery.

As with the rapid development of the Japanese wordprocessor market since 1980, a gradual drop in the price of MT systems is likely to lead to an explosion of purchases, increase competitiveness, and stimulate customer demands for technical improvements in a spiral of interactive needs linking users and developers.

The leading French technology magazine *Science et Techniques* once told its readers that Makoto Nagao gave the best available introduction to MT in an

article he wrote directly in French some ten years ago. For all its brevity, this book also organizes a mine of information about the subject into as orderly a whole as you could want before seeking further details elsewhere. However, ignore the over-academic bibliography at the end, and stick to Electric Word reports.

MT is finally coming out of the closet, in large part due to the efforts of such steadfast but pragmatic devotees as Makoto Nagao.