Chapter 7: Projects and groups in Japan, China and Mexico (1956-1966).

7.1: Projects in Japan.

Research on MT in Japan evidently began about 1956. The major groups up to the mid-1960's were those at the Electrotechnical Laboratory in Tokyo, at Kyoto University and at Kyushu University (Josselson 1970, Nagao 1982). Nearly all the research in this period was devoted to translation from English into Japanese. There were major inherent difficulties in the direct analysis of the Japanese character script, in particular, because of the lack in written Japanese of any indication of word boundaries.

Probably the best known of the earlier efforts in MT in Japan was the research on English-Japanese translation which began in 1958 at the Electrotechnical Laboratory in Tokyo, initially under H. Wada[1] and later under R. Tadenuma (Takahashi et al. 1960, Nishimura 1970, Sakamoto 1970, *CRDSD* 5-15, 1959-69). There was a lack in Japan at that time of general-purpose computers with sufficient storage capacity for MT needs; the decision was made to build a special-purpose machine, the YAMATO, with a relatively large magnetic drum store but no multiplication or division mechanisms. As its corpus in the initial experiments the project selected textbooks for beginning students of English in Japanese junior schools. The deliberate simplification of these texts (just 2000 words, no relative pronouns, no relative adverbs, and no use of present perfect tense) was considered an advantage in the first trial. The system had four dictionaries: English word dictionary (arranged in frequency order), an idiom dictionary (also in frequency order), a syntax dictionary (a table of 20 groups of grammatical categories, referring to syntax subroutines), and a Japanese word dictionary. The inadequacy of direct word-for-word translation in the case of English-Japanese MT was to be solved by syntactic transfer routines. These were initiated by an iterative program for the identification of basic syntax patterns (i.e. category sequences such as noun + verb + noun). Problems of homography had not been tackled. At a later date it was proposed to increase storage capacity by incorporating a photoscopic disc (ch.4.2). It would appear that the YAMATO system lasted until shortly after 1970.

There was also some work in the Electrotechnical Laboratory on Japanese-English translation, apparently on much the same lines. The problems of dealing with Japanese input were detailed by Yamada (1964), who had previously worked at CETA in Grenoble (ch.5.5). He described a program for segmentation and syntactic analysis of Japanese. The program scanned character by character identifying prefixes, roots, suffixes, etc. in a single pass, checking mutual compatibilities of adjacent elements and leaving complete resolution of structures until the Japanese verb was analysed at the end of the sentence (Japanese being a verb-final language.) The parser was modelled on the Harvard predictive syntactic analyzer (ch.4.9)

Research at Kyushu University, Fukuoka, under Katsuo Ohno and Tsuneo Tamachi, also involved the construction of a special-purpose machine, the KT-1 (*Electronics Weekly* Sept 1960, Nagao 1982, *CRDSD* 8, May 1961) Initial research began in 1955, with experimental testing starting about 1960. The research was apparently mainly theoretical in nature, inspired in large part by the work at MIT under Yngve (ch.4.7) There was some investigation of a German-Japanese system but most emphasis was apparently placed on the foundations of an English-Japanese system (*TA* 2(4), Dec 1961), particularly the creation of an English-Japanese dictionary and on the formalisation of English and Japanese grammar (basically on the dependency grammar model). In essence, the system was of the 'syntactic transfer' type (as at MIT) with evident problems in the resolution of homographs and polysemes (Bruderer 1978)

---

[1] For his memoirs see: H. Wada: 'Memoirs of a survivor', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 377-385.

A highly theoretical approach was adopted also by Itiroo Sakai, of the First Research Centre of the Defence Agency, who investigated a method of syntactic analysis based on algebraic operations (Sakai 1962) Intended as 'universal' and "applicable to any language of any syntactical nature" the method was somewhat akin to the categorial grammar of Bar-Hillel (ch.3.4). For example, the combination of the 'word categories' v-st (a verbal stem) and v-st/p/ (a possible suffix to a verb stem) yields p (a predicate); the combination of /n/n + p/ (an 'adjective predicate') and n (a noun) yields n. The basic mechanism was a matrix table lookup combining every category with all others in the sentence. Research on the model continued until about 1968 (*CRDSD* 15, 1969), and the approach was also investigated in the United States by Ascher Opler of the Computer Usage Company, New York, for a French-English system to be run on the IBM photoscopic machine (ch.4.13 above). It is not known whether there were in fact any implementations of Sakai's approach.

At Kyoto University MT research seems also to have been primarily of a theoretical nature. Josselson (1970) reports the work of Toshiyuki Sakai on a "translation procedure from English to Japanese, the principal feature of which is the use of four grammar tables. Analysis is done from the end of the sentence to the beginning, applying these four tables recursively". The most substantial work at Kyoto, however, was to begin about 1968 (ch.18 below)

As in many other countries, Japan witnessed a number of short-lived, small-scale projects during this period. There are reports of MT studies by Seiichiro Ohnishi at Nagoya University involving statistical analyses, by Osamu Watanabe at Niigata University on syntactic studies for English-Japanese and Japanese-English systems, and by T. Obonai and T. Sinohara at Nihon University, Tokyo, also on English and Japanese (*CRDSD* 13, Nov 1964). No doubt there were others.

In general terms, it is evident that up to 1965 Japanese MT research was pursued on much the same lines as in the United States. However, the shortage of powerful general purpose computers in Japan encouraged the investigation of special-purpose machines and perhaps the emphasis on theoretical studies. Research on the specific problems of dealing with Japanese script had been certainly started, but as yet results were still rather tentative.

7. 2: Research in China

In the late 1950's and early 1960's there were apparently a number of MT groups active in China (*TA* 1(1), April 1960). Groups were set up in 1958 and 1959 at various Peking institutes of the Academy of Sciences of the People's Republic of China to investigate Russian-Chinese and English-Chinese systems. There were other groups elsewhere in China, one at Kharbin for the study of French-Chinese and German-Chinese translation. There is little information on the methods used by the groups, and naturally even less on their success or progress. The Russian-Chinese system was designed for translating mathematics texts and was apparently based on the approach of the Soviet group under Lyapunov at MIAN (ch.6.2) It would seem that at least a prototype was programmed and ran with a small dictionary.

7.3: National University of Mexico

A research group was set up in February 1960 under Sergio F.Beltran at the National University of Mexico (Universidad Nacional Autonóma de Mexico, Centro Electrónico de Cálculo). The first experiments were conducted on a Russian-Spanish system for translating titles of scientific and technical articles and books (*CRDSD* 7, Nov 1960). The project was later extended to investigation of systems for English-Spanish, French-Spanish, German-Spanish and Italian-Spanish. Most effort was devoted to the compilation of dictionaries, for which various statistical studies of the vocabulary of titles were undertaken. The project collaborated with the UNESCO Centro de Documentación Científica y Técnica, established in Mexico City to serve all Latin America. For the compilation of the Russian-Spanish dictionary the Mexican group established

close links with the University of California group at Berkeley (ch.4.10) By November 1962 (*CRDSD* 11, Nov 1962) it was reported that the group had completed the adaptation of the Berkeley RUSDIC (Russian-English dictionary). The group continued until at least 1966 (*CRDSD* 14) with plans for "Russian-Spanish machine aided translation".