

## **MACHINE TRANSLATION IN NORTH AMERICA**

### **1. The Period 1947-66**

In a very real sense machine translation (MT) was born in North America, although the concept is really transatlantic, the result of the collaboration of two people: Warren Weaver in the USA and Andrew D. Booth in the UK. The first suggestion of using electronic computers to translate human language arose in the course of a discussion between Weaver and Booth in New York in 1947, although there are some claims that the idea can be traced to 1946. Alan Turing had also imagined that computers could be used for translation, as he indicated in his report for the National Physical Laboratory in 1948.

Weaver had had experience with the new computer technology during World War II when he worked with computers in cryptography. In March 1947, he wrote a memorandum to Norbert Wiener at the Massachusetts Institute of Technology suggesting that computers could be used for translation; Wiener was unreceptive to the idea. Weaver then contacted Booth who had had similar ideas for computer translation. The impetus for machine translation came from a memorandum Weaver wrote on July 15, 1949, to some 200 of his acquaintances proposing the possibility of computer translation of human language. Weaver conceived of the problems of machine translation as similar to those for computer cryptography. Although some of Weaver's views were certainly naive, his memorandum is significant, in that it essentially launched MT as a field of research not only in the USA, but also in the world.

In the two years following Weaver's memorandum, machine translation research was begun at a number of places: the Massachusetts Institute of Technology (MIT), the University of Washington, the University of California at Los Angeles (UCLA), the US National Bureau of Standards in Los Angeles, and the RAND Corporation in Santa Monica, California. This early work was almost entirely devoted to studying word-for-word translation. In January 1950, Erwin Reifler, a sinologist at the University of Washington, circulated mimeographed copies of his *Studies in Mechanical Translation*, n° 1, a series of

## MACHINE TRANSLATION IN NORTH AMERICA

studies in MT, which, among other things, introduced the notions of ‘pre-editor’ and ‘post editor’ to the field. These functions were conceived as ways to assist the computer system in producing accurate translation despite the linguistic problems inherent in the translation process. Pre-editors were to prepare texts for translation, including, in one suggestion, marking each word for its grammatical category; post editors were to correct the results of the translation process. As Reifler’s work indicates, the linguistic problems for machine translation became apparent very early.

In May 1951, Yehoshua Bar-Hillel became the first person appointed solely for machine translation research, when he was hired by MIT to study the possibilities of and to plan future research in machine translation. By the end of 1951 Bar-Hillel had produced a survey on MT: ‘The state of machine translation in 1951’ (Bar-Hillel 1951). This report raised a number of issues that are still significant in machine translation research, for example, the role of post editing and the feasibility of fully automatic translation.

The first machine translation conference in the world was organized and presided by Bar-Hillel at MIT, June 17-20, 1952. This conference was attended by eighteen people, seven of whom were from MIT. Booth was the only person who was not from the USA. Although no formal conclusions came out of the conference, it was agreed that research should proceed in two directions: (a) studies of word frequencies and translations and (b) work on operational analysis of syntax with a view to programming. Proceedings were not published at the time, but the substance of the discussions and a number of the papers were subsequently published. The conference marked the beginning of public awareness of machine translation.

The 1952 conference led to the establishment of a machine translation research effort at Georgetown University by Leon Dostert, who had attended the conference and presented a paper. Dostert intended to approach the problems of machine translation experimentally, rather than theoretically, by developing a small system to see if translation by computer was actually feasible. His approach to the issue heralded what is called the ‘empirical’ approach to machine translation. In conjunction with IBM, the Georgetown University group put on a demonstration of a small-scale Russian-English system on January 7, 1954. Although the

## MACHINE TRANSLATION IN NORTH AMERICA

system was quite small and the demonstration used a carefully selected Russian text, this public experiment helped stimulate US government funding for MT research.

In 1954, Harvard University founded a research group under Anthony Oettinger, who in the same year presented the first doctoral dissertation in the field. The Harvard group concentrated on the development of a large-scale Russian-English dictionary to be used in its word-for-word system. Throughout the period of this research at Harvard the group produced a series of important research reports in MT.

In March 1954, the first issue of *Mechanical Translation* was issued, under the editorship of William N. Locke and Victor Yngve, both of MIT. In October 1956 the first international conference was held in Teddington in the United Kingdom, with some 30 people attending from the United States, the United Kingdom, and Canada, and papers from the U. S. S. R.

The period from 1954 into the early 1960s was marked by considerable activity in machine translation, although many of the projects combined such research with other computational issues, for example, information retrieval and automatic indexing. Some of the research groups that had started in the late 1940s and early 1950s continued their work into the 1960s, including the University of Washington (1949-62), RAND Corporation (1950-60), MIT (1951-65), Georgetown University (1952-63), and Harvard University (1954-64). These groups were joined throughout the 1950s by others, including the University of Michigan (1955-62), the University of California at Berkeley (1958-64), the University of Texas at Austin (1958-), Wayne State University (1958-72), Ramo-Wooldridge Corporation (later Bunker-Ramo) (1958-1967), the National Bureau of Standards, again, (1959-1963), IBM Research Center, Yorktown Heights, New York (1958-66), and others. The particular approaches varied considerably from site to site, with some emphasizing research in lexicons over research in structure, whereas others preferred to investigate syntax over dictionary work. The language pairs worked with varied somewhat less; most of the projects dealt primarily with Russian-English, with some exceptions, for example, The University of Texas.

The machine translation research at Georgetown University is among the most notable

## MACHINE TRANSLATION IN NORTH AMERICA

in the history of this work in the USA. Lasting for some eleven years (1952-63), the project was led by Leon Dostert and included such researchers as Paul Garvin, Michael Zarechnak, Peter Toma, A. F. R. Brown, Ariadne Ljukanov, and Muriel Vasconcellos. The group tried a number of approaches to the problems of MT, including the General Analysis Technique—later called the Georgetown Automatic Translation or simply GAT—which was the basis for much of their later work. Although much of the work focused on Russian-English, other language pairs were employed. The Georgetown project's GAT was installed at EURATOM in Ispra, Italy, and at the Oak Ridge National Laboratory, Oak Ridge, Tennessee, in 1963 and 1964, respectively. The influence and importance of the Georgetown work in machine translation cannot be overstated.

In 1960 a research group for MT of scientific titles was established at the Universidad Nacional Autónoma de México, under the directorship of Sergio F. Beltran. The project started with Russian-Spanish, but later expanded its research into English-Spanish, French-Spanish, German-Spanish, and Italian-Spanish. Much of the work focused on the development of dictionaries. The group continued its work into the late 1960s.

Whereas most research focused on the theoretical issues of machine translation, the idea of commercially viable MT systems was in the minds of some researchers. Notable among those who looked for such commercial applications was Peter Toma, a researcher who had worked on the Georgetown project before starting his own MT company, Computer Concepts, Inc., in 1962. The company's AUTOTRAN Russian-English system became operational in July 1963.

Machine translation work in North America was in a period of considerable optimism for the first decade of its existence. Many projects were well-funded, primarily by the USA government, especially the intelligence and military agencies; researchers were highly confident that solutions to the vexing problems of MT would soon be found. In light of later developments it is now clear that researchers were overoptimistic, underestimating the difficulties involved in the syntax and semantics of natural languages that make the whole task of translating such a challenge.

Yehoshua Bar-Hillel published his *Report on the State of Machine Translation in the*

## MACHINE TRANSLATION IN NORTH AMERICA

*United States and Great Britain* in 1959. This report was followed in 1960 by his well-known ‘The present status of automatic translation of languages’ (Bar-Hillel 1960), which included his assessment of MT in general along with specific comments on research efforts in the USA, Great Britain, the USSR, Italy, and Israel. Bar-Hillel’s papers presented a very pessimistic view of machine translation. Of the groups discussed in the USA, only the National Bureau of Standards group (Washington, DC), led by Ida Rhodes, received anything approaching praise; the group at The University of Texas at Austin was mentioned but not assessed. The most far-reaching part of Bar-Hillel’s paper was his claim that the goal of fully automatic high-quality translation (FAHQT) by computer was unattainable. Using his now famous example, *the box was in the pen*, Bar-Hillel reasoned that no software or program available at that time or in the future could disambiguate such a sentence or any like it. In his view MT researchers should strive for either fully automatic, low-quality translation or partly automatic, high-quality translation. Bar-Hillel’s views convinced a number of people that machine translation was not viable. His influence on MT is still felt in the late twentieth century, although there have been notable achievements in the resolution of the issues his papers raised.

Although Bar-Hillel’s reports had significant influence on people outside of the MT research community, he had little influence on the MT community itself. The reports did have the effect of polarizing the MT researchers into two camps: those who had FAHQT as a goal and therefore felt the need to defend their position, and those who did not have FAHQT as a goal and generally ignored Bar-Hillel’s position. The unfortunate result of these responses to Bar-Hillel’s paper was that the real shortcomings of machine translation did not get addressed.

In spite of Bar-Hillel’s views the machine translation community continued to work on various approaches to the problems associated with computer translation of natural language. Whereas earlier systems had focused on the lexicon as the basis for the system, from 1960 on researchers focused more and more on syntax. A number of approaches to syntax were examined, including use of Chomsky’s formal grammar by the MIT group and the stratificational approach proposed by Sydney Lamb. Such approaches influenced the later

## MACHINE TRANSLATION IN NORTH AMERICA

approaches to machine translation.

The US House of Representatives Committee on Science and Astronautics compiled a report on MT in 1960. This *Research on Mechanical Translation* report was quite encouraging in its views on the prospects of machine translation and recommended support for the whole variety of MT efforts, from the highly theoretical to the very practical. In effect this report gave MT research in the USA an official approval.

With the widespread interest and efforts in machine translation research, particularly in the USA, there was naturally quite a bit of overlap. There was some cooperation between groups, but closer links were deemed necessary. This situation led to a series of national MT conferences: UCLA, February 1960; Princeton, July 1960; Georgetown, 1961; Princeton, 1962; and Las Vegas, 1965. The Association for Machine Translation and Computational Linguistics (later simply the Association for Computational Linguistics) was formed at the 1962 conference. There is some dispute about the particular value of these conferences, although they clearly represented steps in the evolution of the field.

Machine translation came under attack in 1961 in a book by Mortimer Taube entitled *Computers and Common Sense*. In this book Taube stated the view that MT was impossible, because there seemed to be no real way in which to formalize human language precisely enough for computers. His position was simply that all research in computer translation should cease. The book was generally ignored by MT researchers, but it had a negative effect on the public perception of such research.

In the period of 1962-64 several USA groups were successful in producing operational systems. Both the IBM and the Georgetown systems for Russian-English were installed at USA government sites. Although their output was not of the highest quality, these systems were deemed to be serviceable enough for their purposes and were well-received by the users.

Theoretical work did not fare as well as did the more pragmatic. Researchers such as Oettinger and Yngve published rather pessimistic papers on the possibilities for ever achieving FAHQT. In his 'Implications of Mechanical Translation Research' (*Proceedings of the American Philosophical Society* 1964) Yngve even suggested that machine translation

would probably never be able to get beyond the ‘semantic barrier’. The situation in 1964 may be viewed as one of many unfulfilled promises for the field, whereas the research enterprise had become ‘a multimillion dollar affair’ (Bar-Hillel 1960).

A growing lack of confidence in machine translation research led to a request by the director of the National Science Foundation that the National Academy of Sciences set up a committee to study machine translation. The National Academy of Sciences formed the Automatic Language Processing Advisory Committee (ALPAC) in 1964. This committee consisted of seven members from a variety of interested fields. None of the members of the committee, including the two MT specialists, was convinced of the feasibility of machine translation from the outset.

The ALPAC studied the issue of machine translation from a number of standpoints, specifically, existing demand, supply, and cost of translations; demand and availability of human translators; quality of some MT output; and cost of post editing. Their study was generally limited to Russian-English translation. Two experimental and two production systems were evaluated using a specific Russian text.

The conclusions of the ALPAC, published in its report, *Language and Machines: Computers in Translation and Linguistics* (1966) were that funding for MT research should be decreased and funding for computational linguistics increased. Most machine translation researchers at the time believed that the ALPAC report was highly biased and shortsighted. However, the fact remains that much money had been put into machine translation research without a related amount of success. Whatever may be said about the ALPAC report’s objectivity, there is no doubt about its effect on machine translation research in the USA and in the world. By 1968 the MT research effort in the USA had been reduced by approximately 70 percent. Funding in the USA, which had amounted to many millions of dollars between 1956 and 1965, had virtually disappeared.

## **2. The Period 1966-79**

Shortly before the ALPAC report was published, the Canadian government began supporting

## MACHINE TRANSLATION IN NORTH AMERICA

machine translation by funding three projects: one at the University of Saskatchewan, one at the University of Montreal, and one at Cambridge University. All three of the projects were for English-French translation.

The project at the University of Saskatchewan was directed by Kathleen Booth. It started in 1965 and ended in 1972. The system used a direct translation approach with a bilingual dictionary put together using statistical analyses of a 20,000-word corpus. These analyses determined the most probable grammatical categories for English and their most frequent French equivalents. In 1971 the system was tested; the results of the test indicated that the output was usable for post editing.

In 1965 the Canadian Research Council also set up Centre de Traitement Automatisé des Données Linguistiques (CETADOL) at the University of Montreal, under the directorship of Guy Rondeau. CETADOL's initial work focused on problems of natural language processing, English morphology and syntax, and grammatical classification. In 1970 the group was renamed TAUM (Traduction Automatique de l'Université de Montréal); the goal of the newly named group was an operational English-French system. The first prototype for this system was developed under Alain Colmerauer's leadership and work continued on the prototype under the direction of Richard Kittredge until 1977 and the completion of the TAUM-METEO system.

In 1976 the TAUM group had begun work on an MT system for translating weather forecasts. The system, called TAUM-METEO, was delivered in 1977 and was in daily use from then until 1990, when it began to be replaced by a system that automatically generates English and French weather reports from numerical data. TAUM-METEO is easily the most successful machine translation endeavor in the history of the field.

The TAUM group next worked on a project to develop a system for the translation of aircraft maintenance manuals. This project, TAUM-AVIATION, was directed by Marcel Paré throughout its course (1977-1981). TAUM-AVIATION was terminated in 1981, because it was decided that it would never become a cost-effective production system. With the end of the TAUM-AVIATION project, the TAUM group was disbanded.

On the commercial front, Peter Toma continued his work in the development and



## MACHINE TRANSLATION IN NORTH AMERICA

marketing of production machine translation. In 1968 he formed Latsec, Inc., in La Jolla, California. His AUTOTRAN subsequently gave way to SYSTRAN and SYSTRAN II. SYSTRAN replaced IBM Mark II system at Wright-Patterson Air Force Base, Dayton, Ohio, in 1969, where it has remained in use ever since. SYSTRAN was also used by the US National Aeronautic and Space Administration (NASA) for the US-USSR Apollo-Soyuz space project (1974-75). SYSTRAN also was used in Europe at EURATOM and later in the European Community. From a single system that could translate from Russian to English, SYSTRAN now consists of a number of language pairs and is installed at a number of international sites, although it is essentially still a direct translation system.

In 1969, Bernard Scott founded the Logos Development Corporation to develop an English-Vietnamese MT system under the sponsorship of the US Air Force. The result of this work was a direct-transfer system, LOGOS I, which combined features from both the direct and transfer approaches to MT. An evaluation of LOGOS I in 1972 indicated that the output was useful once post edited. Subsequently, in the 1970s, Logos Corporation continued its work in machine translation for such pairs as English-Russian (LOGOS II) and English-Persian. Plans for other systems were considered, but not carried out.

The Pan-American Health Organization (PAHO) entered the machine translation business in 1976 when it contracted three consultants, Bedrich Chaloupka, Giuliano Grugnoli, and Allen Tucker, for the development of a Spanish-English machine translation system. Chaloupka, Grugnoli, and Tucker had worked together previously in developing a small-scale Russian-English system at Xonics in 1970-76. Responsibility for the project rested with Muriel Vasconcellos who had been involved with the Georgetown project in the 1950s. In 1979, Marjorie León joined PAHO as a full-time MT researcher. The result of the work was SPANAM, which became operational in 1980. At this stage SPANAM was a direct translation system, seen as a practical approach to the computational translation of all sorts of Spanish language documents.

In a very real way the ALPAC report (1966) reduced interest in the theoretical issues of machine translation in the USA for some time after its publication. All funding was not eliminated however; projects continued at the University of California, Berkeley (until

## MACHINE TRANSLATION IN NORTH AMERICA

1975); Wayne State University (until 1972); and at The University of Texas at Austin.

From the standpoint of theoretical research, an important figure in the post-ALPAC era was Zbigniew L. Pankowicz, a US civil servant at Rome Air Development Center (RADC), who had the vision to see the usefulness of machine translation in spite of the statements in the ALPAC report.

The MT research at the Linguistics Research Center (LRC), The University of Texas at Austin, is the longest running machine translation research effort in the USA. Although work in MT research had begun at The University of Texas in 1959, the LRC was founded in 1961, under the directorship of Winfred P. Lehmann, as a research center for various computational linguistic projects, not just machine translation.

Funding from 1968 to 1970 was somewhat diminished at the LRC, but research continued throughout the period. In 1970, with funding from RADC, work on an interlingual German-English system was begun. The principal researcher for this work was Rolf A. Stachowitz. The system was initially called simply the Linguistics Research System, but this was changed in 1973 to METALS, then METAL. Although there is still some discussion about what this name stands for, the general view is that it is a shortened form of metalinguistics. METAL has been the name for the LRC's machine translation system ever since. The work on an interlingual system was interesting, but not entirely successful and the project was abandoned in 1975.

After a hiatus, research in machine translation began again at the LRC in 1978. RADC again funded a project, this time to investigate the feasibility of an operational machine translation system. This funding was augmented in 1979 by Siemens AG, the German electronics company, which was interested in an operational MT system to translate large volumes of texts. The project manager initially was Jonathan Slocum; Winfield S. Bennett was the primary linguist for the Siemens-sponsored work.

From 1970 to 1974 Yorick Wilks, who previously had been involved in machine translation at the Cambridge Language Research Unit, worked on a prototype English-French system at Stanford University, California. Wilks's system used a semantic grammar and commonsense inference rules, based on semantic formulas in which preferred, rather than

obligatory, features were represented. This experimental work at Stanford ended when Wilks returned to Europe in 1974.

In 1970 the Translation Sciences Institute was established at Brigham Young University (BYU), Provo, Utah, under the direction of Eldon G. Lytle. Funding was provided by BYU and the Church of Jesus Christ of Latter-Day Saints (LDS). The goal was to produce an interactive MT system for translating English texts into a wide variety of languages. The analysis was due to be done interactively; the multilingual synthesis automatically. In 1979 the sponsors decided that the system offered little chance of commercial viability and work on ITS itself essentially ceased. However, Alan K. Melby continued to work with the system from the standpoint of workstation design. Melby's work with MT work environments, especially from the standpoint of the user, continues to provide useful and significant information to the machine translation community.

The Weidner Communications Corporation (WCC) was founded in 1977, also in Provo, Utah, as a competitor to the Translation Sciences Institute. This corporation was a completely private enterprise, without funding from either BYU or LDS. The corporation moved its headquarters to Northbrook, Illinois, (near Chicago) in 1982, but left the research group in Provo. Bravice International, a Japanese translation company, became the full owner of WCC (then called World Communications Center) in 1988. WCC's primary goal was the development of computer-aided translation (CAT), rather than automatic systems. The system, which has a number of language pairs, runs either on an IBM PC (MICROCAT) or on a DEC MicroVax II (MACROCAT). Although WCC is still active in Europe (under the corporate title Weidner Translation Europe Ltd), activities in the USA ceased in the late 1980s.

### **3. The Period 1980 to the Present**

On one hand, splitting a discussion of machine translation in North America at 1980 is artificial because a number of projects that continued into the 1980s actually began in the 1970s or earlier. On the other hand, however, 1980 is a useful boundary because it marks the

## MACHINE TRANSLATION IN NORTH AMERICA

beginning of a change from the post-ALPAC era to the modern era of MT research.

At the beginning of the decade there were a number of commercial systems either at production stage or under development. On the academic scene there were few projects, the most notable of which was the continuation of work on METAL at The University of Texas at Austin, and the work at Yale.

As stated above, the work on METAL began in 1978 with funding from Rome Air Development Center. This funding was augmented in 1979 by funding from Siemens AG, and since 1980 Siemens has provided all the funding for METAL research. The framework for the system was developed almost entirely at The University of Texas at Austin until 1985; starting in October 1985, the development of parts of METAL framework has been shared among the various METAL sites in Belgium, Germany, Spain and Denmark, along with work on particular language pairs. METAL is being marketed in Europe currently; there is some promise that it will be marketed in North America in the early 1990s.

Although in its current form METAL uses a transfer approach the system has taken on some characteristics of an interlingual system in the past few years, specifically in the use of a language-independent intermediate representation. METAL is a highly modular system that allows the combination of any of the source language analyses with any of the target language syntheses. METAL is unique in combining theoretical and practical considerations in one system.

From 1978 to 1982 there were two MT project at Yale University using the conceptual dependency representations developed by Roger Schank. Schank's conceptual dependency representations seek to represent semantic relationships that express not only the surface structure of a string but also other information that is implied in it. In the first project Jaime Carbonell, Richard E. Cullingford, and others developed a basic interlingual system using 'scripts' to translate an accident report. In 1982, Steven L. Lytinen and Roger Schank experimented with machine translation using Memory Organization Packets (MOPS), which were representations of general, rather than specific, situations. The system was called MOP-STRANS.

From the late 1970s, LOGOS has continued development of its MT system, this time

## MACHINE TRANSLATION IN NORTH AMERICA

using German-English, initially with assistance from Siemens AG. In 1982 Logos Corporation introduced the Logos Intelligent Translation System for German-English; this was followed in 1984 with the introduction of their English-German system. Work is in progress for other language pairs, for example, English-French and German-French. The newer systems use more of a transfer approach than did the earlier direct translation LOGOS systems.

ALPS (Automated Language Processing Systems) was formed in 1980 by Eldon G. Lytle and other people from the BYU Translation Sciences Institute (TSI). Their work was seen as a continuation of the TSI work in interactive machine translation. As it was marketed in 1984, the ALPS system offered the user three possible levels for assistance in translation, ranging from access to the user's own dictionary to interactive translation systems for up to eight language pairs, all involving English as the source or target language. Since 1988, ALPS (now ALPNET) has shifted its emphasis in the machine translation market to providing translation service and translation aids in the form of various packages, for example, Transactive (an interactive translation program), and Translation Support System (a multilingual word processor with a variety of tools for the user). In this they have been most successful.

SYSTRAN has not particularly changed in the 1980s, but has found somewhat wider distribution. Gachot SA has acquired the ownership of SYSTRAN in the USA and Europe, whereas a Japanese company owns the Systran Corporation of Japan. In North America, SYSTRAN remains the MT system of choice for the USAF Foreign Technology Division at Wright-Patterson Air Force Base, Dayton, Ohio; for Xerox; and for General Motors of Canada. It also has a significant market in Europe.

Smart Communications, Inc., founded in the 1970s by John M. Smart, offers an editor, which is designed to assist in technical writing, and a translator, which is designed to translate texts in a restricted source language grammar and lexicon. Without making claims for any particular sophistication, Smart Communications offers tools for multilingual documentation in a variety of subject areas.

Executive Communication Systems, Inc. (ECS), was founded in 1984 to develop

## MACHINE TRANSLATION IN NORTH AMERICA

natural language software, especially MT software, including an MT tool kit. Until 1989 ECS was owned by Bravice, the Japanese firm that also owned WCC. Their systems use Lexical Functional Grammar (LFG). Over the past six years ECS has worked on a number of MT projects, including a Japanese-English system, an English-Chinese system, bidirectional English-French, English-German, and English-Spanish prototypes, and most recently a bidirectional English-Korean system.

1983 saw the start of a project for the creation of an MT model based on AI called TRANSLATOR at Colgate University (Hamilton, NY). The principal investigators for this project were Allen Tucker and Sergei Nirenburg. TRANSLATOR was an interlingual system that was designed to accommodate pragmatic and discourse-structure information, and to represent the expert knowledge of a human translator.

The machine translation research into knowledge-based systems at Carnegie-Mellon University(CMU) is one of the most significant research projects currently in the field. With Jaime Carbonell as director and Masaru Tomita and Sergei Nirenburg as principal investigators, CMU's Center for Machine Translation is committed to creation of a functional knowledge-based machine translation (KBMT) system. The KBMT-89 system is interlingual with a knowledge base used to assist in the analysis of the source language. In February 1989, work began on a new system known as DIONYSUS. Although the current work involves primarily English and Japanese, it is conceptually a multilingual system, because the meaning representation of the source language, which is the result of analysis, may be generated into any number of target languages.

In 1982 the Pan-American Health Organization (PAHO) began work on ENGSPAN, its in-house English-Spanish system, with partial support from the US Agency for International Development (USAID). This effort, which represented a major nonmilitary interest in MT, was led by Marjorie León and Muriel Vasconcellos. ENGSPAN became operational in 1984. The system features that were developed for ENGSPAN have since been retrofitted into SPANAM. Both systems have developed considerably since they first became operational and no longer resemble the GAT system, as SPANAM once did. SPANAM and ENGSPAN are in almost daily use, translating documents in a wide variety of subject areas for PAHO. ENGSPAN is

installed at USAID, the International Center for Tropical Agriculture in Colombia, and the International Rice Research Institute in the Philippines.

In the late 1980s, Pierre Isabelle and Elliott Macklovitch began work at the Centre Canadien de Recherches sur l'Informatisation du Travail on CRITTER, a system for English-French and French-English translation of agricultural texts. More than being simply a limited domain MT system, CRITTER is intended to be a testbed for a theoretically well-motivated model for translation. CRITTER is designed to be a reversible system, that is, the single system will handle both English-French and French-English.

Work on an experimental interlingual English-Chinese system, called XTRA, was conducted at New Mexico State University (NMSU) by Xiaming Huang. This work was begun at the University of Essex, UK, by Huang and taken by him to NMSU. Huang's system incorporates definite clause grammar, case grammar, and Yorick Wilks's preference semantics.

#### **4. Future Prospects**

The interest in machine translation in North America certainly seems to be on the increase. Whereas there was little research in the field at the beginning of the 1980s, there are now several large projects, notably at the LRC (The University of Texas at Austin), and Carnegie-Mellon University, and a number of significant smaller academic projects. Additionally the market for machine translation in North America seems to be opening up, which will stimulate both the development and marketing of existing and future production systems.

Future work on machine translation in North America will certainly focus more and more on semantics, either in semantically driven systems or in semantic components for syntactically driven systems. It is safe to assume that work with knowledge-based systems will continue, as will investigations into AI-based systems. Use of various grammar formalisms in models for machine translation will increase as theoretical linguistics continues to influence the MT community. The computational and linguistic technologies appear to have reached the stage at which production interlingual systems at last seem

## MACHINE TRANSLATION IN NORTH AMERICA

feasible. Perhaps the only limiting factor is the availability of funding to support a variety of MT research efforts in North America.

---



## BIBLIOGRAPHY

- Bar-Hillel Y 1951 The state of machine translation in 1951. *American Documentation* 2: 229-37
- Bar-Hillel Y 1960 The present status of automatic translation of languages. *Advances in Computers* 1: 92-163
- Hutchins W J 1986 *Machine Translation: Past, Present, Future*. Halsted, New York
- Hutchins W J 1988 Recent developments in machine translation. In: Maxwell, et al. (eds.)
- Lehrberger J, Bourbeau L 1988 *Machine Translation: Linguistic Characteristics of MT Systems and General Methodology of Evaluation*. Benjamins, Amsterdam
- Maxwell D, Schubert K, Witkam T (eds.) 1988 *New Directions in Machine Translation: Conference Proceedings, Budapest, August 18-19 1988*. Foris, Dordrecht
- Nirenberg S (ed.) 1987 *Machine Translation: Theoretical and Methodological Issues*. Cambridge University Press, Cambridge

---

Source : "Machine Translation in North America", dans E.F.K. Koerner et R.E. Asher (dir.) *Concise History of the Language Sciences: from the Sumerians to the Cognitivists*, New York, Pergamon, 1995, p. 445-451.