MACHINE TRANSLATION AND THE LANGUAGE BARRIER

Muriel Vasconcellos

Pan American Health Organization
and Georgetown University, Washington, D.C.

1. The Overall Context

Translate by machine? "Impossible!" say some. Others jump quickly to the conclusion that "All our problems are solved!" In reality, Babel is far from being conquered, but there is definite progress in the technology of machine translation, which has been in use in different parts of the world since 1963.

1.1 The Concept of MT

For starters, it's important to be clear about the term machine translation--or MT, as it is often called. The most common definition is: translation generated by computer, with or without human interference (Hutchins 1988a:227, Lawson 1988:106). When MT is combined with human intervention, the process is called human-aided machine translation, or HAMT. Its reciprocal, machine-aided human translation (MAHT) is traditional human-initiated translation in which the computer is enlisted as an aid--mainly in the form of word processing, terminological data bases, and text-critiquing software. These aids, even though they are the defining characteristics of MAHT, can also be used in the revision of MT, of course, and MT is even more effective when it is integrated with desktop publishing and other software resources.

Many people wonder how MT works and whether it's really possible for a machine to translate.

The basic mechanics of MT are not that mysterious. In any MT system the computer uses three different sets of data: the input text, the electronic dictionary/ies that permanently reside in the computer, and the program that executes the logic, or algorithm. The input text has to be in a form that is "readable" by the computer--in other words, on a magnetic tape or disk--or lend itself to optical scanning (OCR) in cases where it is practicable to use this technology. Once the text has been entered, the first module of the program checks the individual words in the text against the source dictionary--an alphabetized set of records that correspond to the words of a given source language or sublanguage (the source and target dictionaries may be combined in a single data set). One by one, the words of the input text are compared against those in the dictionary. For each word that is matched, a complete record is retrieved. The record will include a series of codes indicating: the possible functions of this word; for each function, the corresponding syntactic and semantic characteristics; the relationships to other words with which this word occurs in collocation; and, depending on all this information, either the translation gloss itself, with the accompanying target codes, or an index pointing to this data in a separate dictionary. The information from the dictionary is then copied into a temporary working area, and the program sets about to apply its logic.

It is this logic that represents the linguistic and computational sophistication of the system. Usually it has three components: analysis, transfer, and synthesis (or generation) of the translation. In some systems, still under development, the transfer component is expanded into an interlingua--a full artificial language which incorporates all the linguistic knowledge considered to be universal, or at least common to a large number of languages. The idea of this approach is to cut down on the parts of the analysis and synthesis that are specific to different languages in order to facilitate translation from "many to many."[1] With these building blocks, the computer is able to generate a text, or semblance thereof, in the target language.

The success of MT will depend in large part, of course, on the robustness of the algorithm and the dictionaries, but in a broader sense it also depends on the objective that has been set, the mode and extent of human intervention, the type of text, the needs of the client, the setting in which it is installed, and, naturally, the cost.

1.2 Objective

In the early years of MT, it was thought that with sufficient research one day it would be possible to produce a high-quality translation without any human intervention at all. It was assumed that fully automatic high-quality translation, or FAHQT, could in fact be achieved. Although the products of the earliest systems were far from this dreamed-of goal, with development of their dictionaries and corresponding logic, the quality did improve. Before long these systems, when their dictionaries were large enough, could be used to give a rough idea of the content of a technical text. But the output still fell short of the FAHQT ideal. As the complexities of formalizing the translation process began to be more evident, insistence on FAHQT gradually began to yield to the recognition that for many purposes human intervention in one form or another--in other words, HAMT--would be necessary.

1.3 Types of Human Intervention

There are three points of possible human intervention in the MT process: before, during, and after.

Pre-editing can be of two kinds. In the first, an existing natural language text is revised with the idea of eliminating structural or lexical ambiguities before the job is submitted for translation by the computer. The disadvantage of this approach is that it is difficult to anticipate which are the structures and words that are going to be ambiguous for the machine. As a result, it has not been widely used. However, the situation is beginning to change with the introduction of text-critiquing software such the the SMART Expert Editor (Smart 1988, Walraff 1988) and Critique (Lippmann 1986, Walraff 1988), which, among other applications, can be used for automatic analysis of the input text. These programs recognize certain types of ambiguities and bring them to the attention of the human pre-editor, who can introduce changes before the text is submitted for automatic processing.

In the second type of pre-editing, the input text is written especially for the machine. It may be a new version of an existing text, as in the case of TITUS II (Ananiadou 1987:187; 3.2.4 below), or it may be an entirely new text, drafted from the outset according to pre-established rules and vocabulary. The latter approach is useful for documentation that will be appearing in several languages--for example, technical manuals for products to be shipped to foreign markets. Figure 1 shows a text in Multinational Customized English, a restricted English developed by Xerox Corporation (Ruffino & DeMauro 1986), and its translation into Portuguese by SYSTRAN (see 3.1.1 below). The text in Portuguese, even though it is displayed alongside the final artwork, has yet to be postedited (Russo 1988).

In interactive editing, the computer calls on the human user during the translation process to resolve ambiguities that are identified by the program, which presents various alternatives from which the translator/editor is expected to make the most appropriate choices. This mode can be advantageous for the translation of inflected languages, because when the correct decision is taken prior to synthesis into the target language, noun-adjective and subject-verb agreement is generated automatically--compared with changes made after the automatic process has been completed, which often require a series of flexional adjustments throughout the sentence. As a complement to interactive editing, depending on the purpose of the translation, it is also possible to postedit the resulting product. When this is done, the time spent on the second pass can make the process more costly--sometimes too costly to be worthwhile. The first interactive system on the commercial market was TransActive, from ALP Systems (Weaver 1988; 3.1.4 below), which was introduced in 1982. Recently this mode appears to be gaining greater acceptance (Hutchins 1988b), and it is being considered as part of several systems still under development (for example, Tomita 1986, Ben-Ari et al. 1988).

Of the three options for human intervention, postediting is by far the most widely used. The posteditor, who is usually a professional translator, corrects the machine output after the automatic phase has been completed. This task is much more efficient when it is done directly on the screen using some type of word-processing software (Vasconcellos 1986, 1987a, 1987b, McElhaney & Vasconcellos 1988). The alternative, in which the posteditor writes the corrections by hand and an operator enters them onto the magnetic version, is considerably slower. It is estimated that a posteditor produces between 4,000 and 8,000 words a day (Magnusson-Murray 1985, Vasconcellos 1985:119) and with practice can even attain 10,000 words on some texts. This output is two to five times greater than the 2,000-word standard established for human translators in the United Nations--whose work, moreover, is dictated and has to be transcribed by additional personnel.

1.4 Types of Text

After a while it became clear that the systems which focused on a single type of discourse gave more reliable results. Specialized systems--also called sublanguage systems because the input language is restricted in terms of structure and vocabulary--began to be developed and used for particular applications. These systems provide highly usable, and uniform, results, and they have the added advantage that they require less investment in research and development. A well-known example is METEO, which has been translating weather forecasts in Canada since 1977 (Chandioux 1988, Chandioux & Guérard 1981, Thouin 1982; 3.2.2 below). Restricted input, as pointed out earlier, is considered to be a form of pre-editing. The more restricted the input text, the less need there will be for human intervention downstream. This is one of the big advantages of sublanguage systems.

In contrast, the nonspecialized, or general, systems attempt to deal with any kind of text. In the beginning they were used mainly to gather information--for example, on advances in technology, as is done at the U.S. Air Force, which since 1970 has been using SYSTRAN to translate technical literature from Russian into English (Bostad 1987:129; 3.1.1 below). Later, as MT dictionaries become more developed, this type of system came to used for translations in general. Versatility is the goal, and the best of them have earned the name try-anything systems (Lawson 1982:5). Their success depends on a series of factors associated with their specific application.

1.5 The Needs of the Client

The quality of a machine translation--or any translation, for that matter--must largely be judged in relation to its purpose. For example, in the catch-up spirit of post-Sputnik, the U.S. Government undertook to scan large quantities of material from Russian, and the linguistic refinement of the translation mattered little or not at all. For this purpose SYSTRAN served--and continues to serve--very effectively. It managed to transmit the essential content. On the other hand, at the Commission of the European Communities (CEC), where SYSTRAN has been in use since 1981 (3.1.1 below), until recently everything underwent either a "rapid" or a full postedit. Many of the Commission's texts serve as a basis for diplomatic discussions, and some of them enter directly into law in the member countries, thus requiring very careful translation. For this purpose MT is not always very useful (Wagner 1985 and p.c. 1988). It is now true, however, that some raw MT is being delivered directly to requesting offices (Pigott p.c. 1988).

The urgency of the translation is another factor that contributes to the acceptability of MT. For example, at the Pan American Health Organization the SPANAM system (3.2.7 below) was once used to translate twenty reports in a single day, which had to be scanned by the monolingual rapporteur of a meeting, and on other occasion it provided a rough translation of a long document needed to brief a consultant who was going to travel the next day. In these cases MT, albeit of less-than-perfect quality, made an important contribution.

MT is in fact being used more and more for the transfer of information that previously was not being translated and which perhaps never could be translated by traditional means because of either tight deadlines or the high cost in relation the benefit to be gained--for example, the information contained in data bases. In such cases it is rendering a cultural and social service, if not an economic and political one as well.

1.6 The Setting

The effectiveness of MT will depend in large part on the circumstances of the environment in which it is installed: the form of the input text, the existing hardware, the facilities available for manual intervention, and the attitude of the staff who will have to be using the system daily, as well as of the clients who receive the final product.

It is very important--almost indispensable--that the input texts be already in machine-readable form. Sometimes, for example at the U.S. Air Force, the volume, the priority and difficulty of the source language, and the importance of the texts to be translated (which come from a wide variety of publications) justifies the cost of entering them especially for translation, either by optical scanning or by hand. Usually, however, the cost of MT is more difficult to defend when the texts are not already in magnetic form.

Until recently, MT systems tended to be dependent on the computers for which they had originally been designed. For example, a system installed on an IBM mainframe needed extensive adaptation before it could run on a VAX. This situation is beginning to change, however, especially with adaptation to microcomputers, for which the C programming language has been highly popular. This language is easily ported between different types of mainframes, different micros, and between micros and mainframes. It should be kept in mind, however, that easy portability of the code does not necessarily mean that microcomputer MT systems produce results comparable to those of their mainframe predecessors. The case of the micro is examined in section 5.

Although computers--both mainframes and micros--are by now widely available throughout the world, for many translation services it is still a struggle to obtain enough equipment so that word processing is available to all professional staff. Without this capability, which permits postediting to be done directly on-screen, MT will not realize the savings that are possible.

Finally, even when the fanciest and most modern equipment is available, the acceptance of MT will still depend in large measure on the good will of the users--both the translators and the ultimate consumers of the translations.

1.7 Cost

The cost of machine translation is the factor that has the greatest impact on the success of a project. Cost becomes significant when considered in relation to either the traditional cost of translation or to new benefits that may be generated. It will usually take into account most or all of the following investments: (1) initial and periodic fees to the vendor for use of the translation program, dictionaries, and related software, (2) ongoing maintenance of the software and hardware, (3) overhead for use of the computer and other installed capacity, (4) manpower for dictionary development, (5) cost of inputting the source text by hand or with the aid of OCR, and (6) human intervention. The initial investment and maintenance are usually weighed against the estimated volume of texts to be translated. Formulas exist for calculating the mimimum number of words that justify the installation of MT, depending on the schedule of charges. Computer costs will vary from site to site, but in general the rates are trending downward. What always impacts most heavily on the budget is the expense of human intervention in the process, especially pre-, post-, and/or interactive editing.

Even though editing is the biggest budget item, the total process is usually still cost-effective compared with traditional translation. Pre-editing, as pointed out earlier, is considered worthwhile in the case of one source language into many target languages, and the savings will increase in proportion to the number of the latter. With postediting, the situation varies depending on the purpose of the translation, the type of text, and the experience and ability of the translator/posteditor. With a restricted input language, postediting is minimal--since the human intervention has already taken place upstream. On the other hand, the postediting of general texts is heavier and can be costly--depending on the degree of refinement desired. Still, even with careful postediting, savings of 30% and 40% can be attained (Vasconcellos 1984, 1988a, Magnusson-Murray 1985, Lévy 1988).

Of course, human intervention is only one of the elements to be taken into account; a full cost analysis would have to include all the investments mentioned in this section. And cost will only be meaningful if it is measured in terms of the needs being met. End users may be willing to sacrifice a degree of quality in exchange for new benefits such as faster turnaround and a machine-readable product. But if the product is not serviceable, savings in cost will be irrelevant.

1.8 Putting It All Together

The success of MT will depend on the right combination of all the factors mentioned in this section. The installations that work effectively are the ones that combine the best conditions for each factor in the most efficient way possible.

2. Over the Years

It is helpful to look at MT in the context of its historical development. The dream of translating by machine goes back a long time, and its realization has unfolded closely in tandem with advances in computer technology. At the same time, the impetus for development has responded to the political, economic, and social pressures that existed during the different periods in MT history.[2]

2.1 Prelude

It may come as a surprise to some people that the concept of MT was already being contemplated, formulated, and refined more than half a century ago. On July 22, 1933, Georges Artsrouni, an engineer of Armenian descent living in France, obtained a patent for a machine that substituted words or groups of words for their equivalents in other languages. The words of the input text, coded as punches on a paper tape, were matched up to list of words (the "source dictionary") that were similarly perforated on another band. A model of the system, which was intended to serve other purposes in addition to translation, was actually built and demonstrated, and the French railway and telegraph services might have started using it if they had not been distracted by the coming of World War II.

In September of that same year, Petr Petrovich Trojanskij, working independently in the USSR, also obtained a patent for the design of a machine that translated from one language into several others simultaneously. He spelled out a detailed process of machine translation the principles of which remain valid until today. He identified the three phases of the process: analysis, transfer, and synthesis. The transfer, he said, should be an intermediate universal language expressed in logical representation. There were two dictionaries: a list of words in the source language and another of equivalents in the target language. In the first, associated with each word there was also a string of linguistic codes ("marks of logical analysis") which referred to a set of grammatical rules. Trojanskij considered that it was necessary to have human intervention at both ends of the process: in the preparatory phase, to identify the basic forms of the words and their syntactic functions, and at the output end, to supply the inflexions and review the text that the machine produced. In 1941 he demonstrated an electromechanical model of his system, and seven years later he proposed an electromagnetic machine that was quite similar to the Mark I, the first operational computer.

2.2 The Birth of MT

It can be seen, then, that the idea of machine translation long preceded the invention of computers that were able to turn it into a reality. And indeed, the official birth of MT came only a year after the ENIAC, the world's first fully electronic computer, made its debut. The MT initiative is attributed largely to Warren Weaver, a vice president of the Rockefeller Foundation who at the time was involved in the sponsorship of research on computers. Weaver was convinced that the techniques of cryptoanalysis could be used to encode and decode the meaning of natural language, and that this would be the key to translating by computer (Weaver 1947). He believed that languages were based on universal concepts that could be represented in a logical language (Weaver 1949). On March 6, 1947, Weaver had a meeting in New York with Andrew Booth, a British researcher, and the two scientists exchanged ideas on the feasibility of MT. Booth, recalling that conversation much later (1985), said that he, Booth, had already talked about the subject with Alan Turing some years before, and Turing himself was to mention the possible application of computers to translation in 1948.

Later in 1947 Booth and his colleague Richard Richens, working in England, were to begin the first MT research using an electronic computer. These investigators conceived the idea of reducing the size of the dictionary by using "split" instead of "full" forms. The dictionary lookup program contained rules for morphological analysis which, in conjunction with tables of inflexional endings, made it possible to recognize all normally inflected forms in the input text. Fully inflected forms were only included in the dictionary in exceptional cases. Booth and Richens also introduced the concept of microglossaries, which were to provide special overriding translations for different subject areas and types of discourse. And finally, they proposed solutions for dealing with the words in the input text that were not found in the source dictionary (gap analysis). All these concepts are routinely incorporated in today's state-of-the-art MT systems.

2.3 The Challenge of Real Texts

Other pioneers joined the effort. The results of the first attempts were rather primitive, and strategies were quickly developed to deal with the linguistic problems that the machine could not yet solve. Erwin Reifler, at the University of Washington, was the first in the United States. In 1950 he advanced the notions of pre- and postediting à la Trojanskij, considering that both these steps were unavoidable.

The selection of the right gloss from a series of alternatives was obviously one of the big problems to be solved in MT. To provide the user with as many clues as possible, "slashed entries," were used, with the output text showing all possible alternative glosses. For example:

> Infection/corruption (by/with/as) nodular (by/with/as) bacteria comes/advances/treads especially/peculiarly (it)(is)light/easy(ly) at/by/with/from (of) plants, (of) weakened/loosened (to/for)(by/with/as) nitrogen/nitrous (by/with/as) starvation, and/even/too (is)considerable/significant(ly) (is/are)more-difficult(ly) happens/comes-from at/by/with/from (of) plants, (is)energetic(ally) (of)growing on/in/at/to/for/by/with (of) rich nitrogen/nitrous soils (Micklesen 1958).

This approach was endorsed especially by Anthony Oettinger, who headed up a project at Harvard University. One of his examples (Giuliano 1961) was:

> NEW/MODERN/NOVEL METHOD/WAY MEASUREMENT/METERING/SOUNDING/DIMENSION SPEED/VELOCITY/RATE/RATIO LIGHT/LUMINOSITY

which could be postedited to read as follows:

> NEW METHOD FOR MEASURING THE SPEED OF LIGHT

Initially it was believed that these strategies would become less necessary as research gradually yielded up the missing computational solutions. Soon, however, the complexity of natural language became apparent. Weaver's cryptoanalytic approach was recognized as being too simplistic. Investigators began to see that they had an enormous task ahead of them, especially in the study of syntax and discourse structure.

Before long projects were under way in various centers in the United States, Canada, England, and the USSR. At MIT, the first MT full-time researcher was appointed in 1951: Yehoshua Bar-Hillel, who was replaced in 1953 by Victor Yngve. Other institutions that joined the effort early on were the Rand Corporation (1950), Georgetown University (1952), and Harvard University (1954). It is important to remember that since the beginning the development of MT responded to the pressures that generated funds for research support. It was the climate of the Cold War, aggravated by the launching of Sputnik, that provided the impetus in the United States for systems that would translate from Russian into English. Thus, the projects at Georgetown and Harvard were for Russian-English, and in 1956 the U.S. Air Force called on Reifler to switch his concentration to that language pair as well. The investment was great, and its size alone permitted development of the two essential elements of machine translation: large dictionaries with full linguistic coding, and grammars expressed in forms manipulable by the computer. The grammars could only be tested by being run against massive quantities of real text, and for this it was necessary to develop dictionaries that contained all or at least most of the words that were in the input texts.

The first demonstration of MT on an electronic computer was in 1954, the result of an experiment undertaken by Georgetown University, in Washington, D.C, in conjunction with IBM. This demonstration was followed two years later by a large grant to Georgetown from the U.S. Government to undertake the translation of Russian into English. The Georgetown project, under the direction of Léon Dostert, had an eminently empirical approach. The research was based on a large corpus of scientific articles--genuine texts with existing criterion translations already done by human translators--instead of isolated sentences, many of them conjured up by the investigators themselves. This corpus dictated the coding of the dictionary, which grew to more than 49,000 base forms (Zarechnak, cited in Vasconcellos 1988b). The syntactic analysis was based on the notion of transformations, new at the time, of Zellig Harris, mentor of Noam Chomsky. The system was modular: the output of each module served as input to the next, which facilited the introduction of improvements and the division of labor among different members of the team.[3]

Georgetown's approach contrasted with those pursued at MIT and the University of Texas, which were regarded as "theoretical." Research at Texas, begun in 1958, was for the translation of German into English and had financing from the U.S. Army. It was carried out at the Linguistics Research Center under the direction of Winfred Lehmann. The project's aim was to make maximum use of a monolithic grammar that undertook to solve all the problems that could be anticipated. Whenever possible, the solutions were to be <u>context-free</u>--in other words, dependence on specific context was avoided.

As time passed, the two approaches, empirical and theoretical, gradually took inspiration each from the other, until today the distinction is no longer meaningful. Both systems ultimately became operational. Georgetown's GAT was used to translate texts in the field of atomic energy at Euratom in Ispra, Italy, from 1963 until 1970 and at the Oak Ridge National Laboratory, Tennessee, from 1964 until recently (Henisz-Dostert 1979:159, Zarechnak p.c. 1988). The Texas system, in a more recent version underwritten by Siemens AG, is being marketed by this firm under the name of METAL (see 3.1.8 below).

2.4 ALPAC

By the end of the first half of the 1960s MT was in full production--the Georgetown system (GAT) since 1963, and IBM's Mark I system at the U.S. Air Force in 1964, followed in 1965 by the Mark II. They were both translating from Russian into English for information purposes. Although the quality had not yet reached the FAHQT ideal, the great majority of end-users found the product to be adequate for their needs (see, for example, Henisz-Dostert 1979 on GAT).

Still, for several reasons the process was expensive. First and perhaps most important, the input texts had to be keypunched character by character. Labor was so expensive in the United States that Georgetown arranged to have the texts sent to Germany, where trained operators keyed them in and the resulting decks of punched cards were airlifted back to the States for processing. Another cost was the revision of output. The translator had to indicate the corrections in the machine output by hand and then pass on the revised text either for handling by someone who was trained in data entry, or for recopying from scratch. Moreover, computer time, counted in seconds, was very expensive, and computer operators and programmers commanded high salaries. Programmers were an elite minority, and they wrote in an arcane and difficult language, Assembly. And on top of all this, of course, was the cost of ongoing linguistic development.

With the production of translations already under way, the agencies of the U.S. Government that were financing the various research projects began to look at whether or not to continue their investments. They wanted to know: Is the current output adequate? Is it worthwhile to spend more money and get better results? How much additional effort needs to be made? What should future research be focusing on? To answer these questions, in 1964 the U.S. National Academy of Sciences and the National Research Council appointed the Automatic Language Processing Advisory Committee, a panel of eight scientific linguists, to assess the government's needs for translation and review the status of MT as of that date. There are those who say that the composition of the team already prefigured its conclusions (Zarechnak 1979:52-53; Hutchins 1986:165). The majority of the members were strongly committed to theoretical linguistics, and there was no one on the Committee who was greatly experienced in translation. Few translators were consulted in the course of the study.

The Committee looked at the status of translation in the U.S. Government, with particular focus on Russian into English, and reached the conclusion that the supply of translators was far in excess of the demand. Moreover, it went on to venture the opinion that a considerable proportion of the texts being translated were not really needed. Because of this perceived situation, the panel concluded that unless a high-quality MT product could be generated without human intervention, the process could not be justified. As long as the output required editing, MT was too expensive to be worthwhile. The Committee failed to ponder whether the machine translations already being produced were meeting the needs of their target end-users.

The ALPAC study design called for comparison of three raw machine translations (GAT, Mark I, and Mark II) against three human translations of the same text. Today it is generally agreed that the study was defective in several regards: the presentation of the texts, the criteria used for evaluating the translations, and the choice of participants (Vasconcellos 1988c). While it is true that the Committee might be forgiven for its ignorance, in view of the then state of the art of testing and of discourse and translation theory, there is little excuse for its failure to reflect on the possible benefits of future research along the practical lines that had been traced by MT up to that time.

The conclusions contained in the final report (ALPAC 1966) were fatal for the MT projects already under way in the United States. The Committee recommended that all ongoing MT research be abandoned in favor of: (1) further studies in theoretical linguistics, and (2) greater investments in the development of electronic tools for human translation such as word processing and term banks.

2.5 The Post-ALPAC Period

Although the ALPAC report cast a pall on MT development, especially in the United States, for more than a decade, fortunately its effect was not decisive. In the period that followed, political and economic pressures brought about shifts in the translation picture, while at the same time progress in computer technology, in the form of vastly improved storage and processing capabilities, began to permit computational solutions to problems which up to then had seemed insurmountable.

In international politics, the dissipation of the Cold War opened the way for translations with friendlier purposes. In Canada, the legislation that granted linguistic equality to French meant that enormous quantities of texts, both official and commercial, would have to be translated. In Europe, the Commission of the European Communities, with six official languages, addressed the challenge of producing translations in 30 combinations--a total that jumped to 72 with the incorporation of Greece, Spain, and Portugal.

At the same time, fabulous developments were taking place in computer technology. It should be remembered that in the 1960s computers were still quite limited in terms of space and power. Computer time was expensive and programmers were few and highly sought-after. Moreover, in the case of MT, in order to submit a text to the computer it had to be keypunched first. All this was to change, however, with the advent of miniaturization, which permitted not only compact storage of the large dictionaries that MT requires but also rapid access to the respective records--all this taking place with previously undreamed-of efficiency and speed. Programmers began to proliferate, especially with the introduction of higher-order programming languages that simplified and demystified the task. As linguists learned how to write their own programs, a new profession emerged: that of the computational linguist. For machine translation in particular, perhaps the most important advance was the development of word processing. This technology greatly facilitated the input of texts for MT, since it created a pre-existing corpus of machine-readable material, and it also permitted, for the first time, direct postediting on-screen. No longer was it necessary to have the posteditor's corrections entered by a second person or to recopy the entire text. At the same time, developments in optical character recognition (OCR) provided automatic assistance for the capture of texts that were not already in magnetic form. Thus the most tedious steps in the process--input and postediting--suddenly became easy and cost-effective.

And finally, all these advances in computer and related technologies provided yet another pressure that turned out to be a boost for MT: the economic pressure for the new industries to sell their products in foreign markets. Technical manuals had to be provided in the local language wherever the machines were being sold, and it was natural that MT should be enlisted to promote its own host.

3. The Systems

In the absence of public sector support for practical research after 1966, the private sector took over. For the first time, MT began to be developed from the start as a commercial venture. The installation of SYSTRAN at the U.S. Air Force in 1969 marked the beginning of a new era in the history of MT. The other pioneer on the commercial scene was Logos, a company whose activities began in 1969, although they did not culminate in a commercial product until 1983. In the meantime Weidner (later WCC), founded in 1977, installed its first system in 1979. In 1980 ALP Systems (now ALPNET) entered the picture with easily manipulable electronic dictionaries and interactive MT. These are the four firms which have dominated the market during the post-ALPAC years. There have also been other commercial initiatives, and some public-sector ones as well. The following pages offer a "grand tour" of MT systems around the world. They are presented in the order in which they became operational.

3.1 Commercial Systems

3.1.1 SYSTRAN

SYSTRAN, developed by Peter Toma, is without doubt the MT system that has been most tested, most used, and most widely implemented throughout the world. Of all the general systems, it is the one that "tries-anything" the most and the hardest. Toma, a Hungarian-American working in California in the mid-1950s, first invented a small multilingual system, the precursor of SYSTRAN, which received wide publicity in 1957. Shortly afterwards he joined the team at Georgetown University, where he worked from 1958 to 1960. After he left Georgetown he was associated for three years with the firm Computer Concepts, where he developed two other systems, AutoTran and TechnoTran (Toma p.c. 1988). Returning finally to his independent efforts, he announced SYSTRAN, translating from Russian into English, in 1964. An improved version of this software was installed at the U.S. Air Force in 1969 (Toma 1976) and has been running there full steam ever since (Bostad 1982, 1987, Gachot p.c. 1988).

The next SYSTRAN combinations to be demonstrated were Chinese-English and German-English, in 1972. Soon afterwards a request came from NASA for English-Russian MT to support the Apollo-Soyuz mission, and accordingly the company's linguistic team, under the coordination of Joann Ryan, switched over to the analysis of English as a source language and worked on this project during 1974-1975. The next product to be developed, English-French, was purchased by the Commission of the European Communities (CEC) in 1976 (Pigott 1988) and later by other clients. One of the first commercial clients, Xerox Corporation, requested English-German, -Spanish, -Italian, and -Portuguese to assist in the publication of technical manuals. These combinations were installed in 1978. By the end of the 1970s SYSTRAN was operational in numerous language combinations and had clients in the United States, Canada, and Europe. Also, an ambitious project had been launched for the development of English-Japanese and Japanese-English. Today at the CEC alone there are 12 combinations in different stages of implementation or development: English-German, -Spanish, -French, -Dutch, -Italian, and -Portuguese; French-German, -English, -Italian, and -Dutch, and German-French and -English.

In 1986, after 30 years in the field of machine translation, Dr. Toma decided to pass SYSTRAN on to new owners in order to be able to devote his full efforts to the cause of world peace. The rights sold for $7 million. Today all the combinations except Japanese/English are in the hands of the Gachot family, which recently brought the different companies together under the single name of Systran Translation Systems, Inc. The Japanese combinations remain with Tokyo-based Iona Corporation, which has developed them both to the operational level.

SYSTRAN has always run on IBM mainframes. On-line service is offered by telephone: the client can send a text by modem from any part of the world, to be translated on SYSTRAN's mainframes either in California or in Paris, and sent back to the requester by wire. In addition, in 1988 a microcomputer product was announced, but it was not yet being marketed (Gachot p.c. 1988).

3.1.2 LOGOS

After SYSTRAN, the oldest commercial MT company is Logos, established by Bernard Scott in 1969. Logos' first project was a system for the translation of aircraft training manuals from English into Vietnamese. This combination represents a major challenge because the two languages divide up the Universe very differently. To deal with these differences, Logos devised a clever set of semantic tables which were to stand them in good stead for the analysis of other languages as well in future projects. The system was used in Vietnam from 1971 to 1973, when the U.S. presence came to an end.

Logos undertook a series of other projects in the years that followed. The largest of these was an English-Farsi system to be used in Iran. Once again, however, Destiny was to interfere and it was never installed.

Turning to languages of less specialized interest, the firm then began to develop products for which there was a broader market. In 1982 a general system was unveiled for translation from German into English, and this was followed shortly afterwards by the reverse direction. The success of the German-English was reported by Lawson (1984), who found increases in productivity at four sites visited soon after the system was first installed. An example of German-English LOGOS output is shown in Figure 2. Today Logos offers English-French, English-Spanish and German-French. They run on an IBM mainframe and the Wang VS-100 minicomputer, and the system has also been ported to UNIX.

One of the outstanding features of the LOGOS system is the interactive software for updating the dictionaries. It comprises two utilities, ALEX, the Automatic LEXicographer, and his girlfriend SEMANTHA (Wheeler 1988). SEMANTHA was originally a tool used by in-house developers to write linguistic rules based on syntactic and semantic preferences, and it is now available to users as well (Wheeler 1988). Recently Logos announced the birth of FILIUS--offspring of ALEX and SEMANTHA. This is a split-screen word processor that runs on a PC and assists in the postediting of LOGOS translations (Cave 1988).

3.1.3 WCC

The Weidner Corporation (renamed Worldwide Communications Corporation--WCC--in 1987) was founded in 1977 by two brothers of the same name, and its products were introduced on the market in 1979. The first operational combination was English-French, which was soon followed by English-Spanish and -German and Spanish-English. The first client was the Canadian firm Mitel, which used MT to translate technical manuals from a relatively restricted English into French. In this installation MT was one of the links in an automated publishing chain (Hundt 1982). Another important WCC client, which uses the system for a broad range of text types, is ITT (today ESC), a translation bureau in England that employs a large team of MT posteditors. In this latter setting productivity with MT postediting has been double that of traditional translation since 1982 (Magnusson-Murray 1985:178). At both installations, Mitel and ITT (ESC), postediting rates have been estimated to range between 800 and 1,000 words an hour. The system has also been used by WCC itself at its headquarters in Deerfield, Illinois, for its own translation service.

The firm was purchased by Bravice International of Japan, which has been working on the development of Japanese-English. As of 1988, WCC had 10 operational language combinations: English-German, -Arabic, -Spanish, -French, -Italian, -Norwegian, and -Portuguese, plus German-, Spanish- and French-English (Strozza p.c. 1988).

Updating of the WCC dictionariess is interactive. The coding is less complex than that of SYSTRAN or LOGOS.

Weidner made MT history when it launched MicroCAT, the first system to run on a microcomputer, in 1983. Both MacroCAT (which runs on VAX computers) and MicroCAT are used in various parts of the world.

3.1.4 ALPNET

Tracing its origens to an MT project at Brigham Young University in Provo, Utah, Automated Language Processing Systems (renamed ALPNET in 1988), has been selling products designed to lighten the translator's task since 1981 (Good 1988, Weaver 1988). These include: word processing, selective dictionary retrieval, automatic dictionary retrieval (AutoTerm), automatic access to repeated texts (Repetitions Processing, or repstraction), interactive machine translation (TransActive) and, more recently, two batch MT products (ASK and TransMatic). AutoTerm, which includes both terminology management software and dictionary data sets, is the ALPNET product in widest use.

TransActive, launched on the market in 1982, has six operational combinations: English-German, -Spanish, -French, -Italian and German- and French-English. The user can set switches to control the degree to which the process is interrupted, depending on the nature of the text and the purpose of the translation. In addition to interactive editing, with TransActive there is also the possibility of postediting the final output.

All these products, originally designed for Data General computers, are available today for IBM ATs and clones.

ASK and TransMatic, developed for Spanish- and Portuguese-English, run exclusively on IBM mainframes and are used for information-gathering. ASK translates word for word, presenting all the possible translations for each word in the form of slashed entries (à la Reifler and Oettinger, 2.3 above) in the final output (Beesley 1988). TransMatic, on the other hand, does syntactic analysis and produces a translation in the form of ordinary text. An example of unedited Spanish-English output is shown in Figure 3.

In a recent policy shift, ALPNET has decided to place less emphasis on the sale of software and to focus more on the provision of translation services. It has already put together a network of centers in various parts of the world.

3.1.5 SMART

Smart Communications, Inc., established in New York in 1977 by Australian-born John Smart, offers machine translation, principally on the basis of restricted input, in nine combinations: English-German, -Spanish, -French, -Greek, -Italian, -Portuguese; Spanish-, French- and Italian-English (Smart 1988). Dictionary updating is interactive. If the input text is in

French or English, it can be submitted first for automatic critiquing by the SMART Expert Editor, a product that is also sold separately for use in the preparation of technical manuals. Since 1984 the SMART Translator and the Traducteur SMART have been processing vacancy notices for the Canadian national job bank (Bergeron 1988). The messages are transmitted electronically--more than 2,000,000 a month--with a view to reducing pileups of paper (Smart 1988).

3.1.6 Globalink

The MT products of Globalink, which run on IBM ATs and clones, trace their origins to a chain of ownership that has included Xonics Corporation, Tabor Corporation, Challenge Systems, and Telecommunications Industries, Inc. (TII). This last firm inherited from its predecessors a VAX-based line of MT products and also a young PC version, on which development efforts were then concentrated. Globalink, created in 1988, now offers TWP, a PC-based general translation system from Spanish and French into English and English into Spanish, plus a series of subject-specific microdictionaries, and has clients in the U.S. Government and private industry (Rowe p.c. 1988). The company, which operates out of Fairfax, Virginia, is establishing a worldwide network for distribution of the software. The continuity of this line of products is owed to the perseverance of Bedrich Chaloupka, one of the original members of the GAT dictionary team at Georgetown University.

3.1.7 ATAMIRI

Developed by the Bolivian mathematician Iván Guzmán de Rojas, ATAMIRI uses a central syntactic representation based on the Indian language Aymará. The inventor has frequently cited the special syntactic and logical characteristics of Aymará (Guzmán de Rojas 1985), which he feels make it singularly appropriate as an intermediate structure for machine translation (Guzmán de Rojas 1988). ATAMIRI's Aymará-based design does in fact facilite translation into multiple languages, but one cannot say that it is based on a true interlingua because it does not invoke all the linguistic information that such a system is normally defined as encompassing.

ATAMIRI runs on Wang computers. In the early demonstrations in 1984-85, it still had a very limited vocabulary and was able to translate into several languages simultaneously. However, when the system began to be used in 1985 in practical operations that required large dictionaries, development efforts were focused initially on the English-Spanish combination. Over the next two years English-German, -French, and -Dutch were implemented, and as of late 1988 English-Italian and -Swedish were being implemented as well (Guzmán de Rojas 1988 and p.c. 1988). ATAMIRI is used in Wang International Translation Centers for the production of technical manuals.

3.1.8 METAL

In 1979 the Munich-based international computer firm Siemens AG took over sponsorship of the German-English project that had been carried out at the University of Texas from 1959 to 1974 (Slocum 1988, Bennett & Slocum 1988). In the new project, called METAL (which may stand for "Machine Translation and Analysis of Natural Language" and has other expansions as well), the linguistic development was done on a Symbolics LISP machine, which is very fast and user-friendly for the formulation of grammatical rules. The German-English METAL product was ready for beta-testing in 1986, and by 1988 the system was in full operation at six translation services in Germany and the Netherlands, including the Siemens Corporate Translation Center in Munich (W.S. Bennett p.c. 1988; Language Technology Jan-Feb 1989). It is the first of the "theoretical" systems to reach the market.

3.1.9 TOVNA

Launched in 1987, TOVNA, like other commercial MT systems, undertakes to translate general texts. It is sold by a firm of the same name which has its headquarters in Tel-Aviv, Israel, and a branch in London. Hailed as "the learning system," TOVNA is claimed to have an algorithm that is language-independent; the rules of the user's language or sub-language are generated as the system is exercised. Formulation of these rules depends on instructions given by the user, but it is still necessary for the Tovná personnel to assist in tailoring the dictionaries to suit the user's needs. The system does not yet have a track record; as of 1988, it was too early to tell whether the investment required to develop a new language pair or to customize a dictionary is less than with other MT systems. The English-French combination is installed at a translation service in England (Weiss p.c. 1988).

3.1.10 XLT

A proprietary system of the Canadian translation service SOCATRA, Inc., XLT began to translate English into French in November 1987 (Bédard p.c. 1988). It is used internally by SOCATRA for general translation work.

3.1.11 Japanese-language systems

The heavy commercial traffic between Japan and the West has generated a translation market from Japanese into English and vice versa without parallel in the history of language. To give an idea of the economic importance of these combinations, in 1986 the billings for translations done in Japan amounted to more than 1 trillion yen (US$5 billion), or as much as the entire foreign debt of Uruguay for that year. And the demand continues to grow. It is estimated that the Japanese market for MT systems alone will be worth 250 billion yen (US$1.25 billion) by 1990 (Whitelock 1987:147). The MT situation is complicated by the fact that the linguistic difficulties in formalizing these combinations are particularly challenging.

An "MT Summit" was held in September 1987 in Hakone, Japan. On that occasion reports were heard on no less than 14 English/Japanese systems, which, together with others not discussed at Hakone, make for a total of 20. In the following list the asterisk (*) indicates that the system is already in practical operation: *ATLAS-I and -II (Fujitsu), *DUET (Sharp), *HICATS (Hitachi), IBM/Tokyo Research, LAMB (Canon), LUTE (Nippon Telegraph & Telephone), *MELTRAN (Mitsubishi), *MEDIUM/MICRO (Bravice), *Nippon-Data General, PAROLE (Matsushita), *PENSEE (OKI), *PIVOT (NEC), RMT (Ricoh), SMART, SMTR (Resource Sharing), *SWP-7800 (Sanyo), *SYSTRAN (Iona), TAURAS (Toshiba) and VORTEX (Toma p.c. 1988). There is also the national system MU, being developed by Kyoto University, which is mentioned below in Section 3.3.2, and work being done at the Universities of Manchester and Sheffield in England (also 3.3.2).

3.1.12 The Trends

It can be seen, from the progress made since 1966, that often the impulse which led to the implementation of practical MT systems was the need to sell commercial products, especially electronic products, in foreign markets.

The specific nature of this pressure is reflected in the directions that MT development has taken. The commercial systems have tended to be general, rather than specialized, in order to appeal to the largest share of the potential market. In terms of the combinations developed, English has been the source language in most of the operational systems, while as a target language it is already losing importance--confirming the fact that today there is more need, in the United States at least, to disseminate information than to gather it. At the same time, there is an explosive demand for both Japanese-English and English-Japanese, and there is also interest in German as a source language. Combinations that do not include English are beginning to be more important commercially.

Thus, more than anything it has been economic interest, in the form of market expansion in foreign countries, that has established MT as a viable mode of translation. The free market has succeeded in proving what in 1966 the American public sector had thought was impossible: that MT really works.

3.2 Non-Commercial Systems

Even though the majority of systems currently in operation are for commercial use, on the other hand the public sector has not entirely abandoned its efforts in the area of machine translation. Some projects did continue, or were initiated, after 1966 despite the influence of ALPAC. Many of these were projects associated with universities, and outside the United States a lot of them had government support.

3.2.1 AMPAR/NERPA/FRAP

Research in the Soviet Union, begun in 1955, was affected less directly by ALPAC, but nevertheless it went through a cycle similar to that in the United States: optimism followed by a certain disillusionment, a hiatus in the mid-1960s, and after that a return to work with a more pragmatic approach, which led to the development of operational systems (Hutchins 1986). Progress was slowed by relatively limited access to computers.

Since 1974 most of the activity has been concentrated at the Center for the Translation of Scientific and Technical Literature and Documentation in Moscow (Marchuk 1977), where the following systems have been developed: AMPAR, from English to Russian; NERPA, from German to Russian; and FRAP, from French to Russian. These are all general systems whose main purpose is to gather information. Depending on the purpose of the translation, postediting is also used (Marchuk 1984:205).

AMPAR, which is geared to the translation of texts in technical fields, especially informatics (Marchuk 1984:98), became operational in 1979; NERPA in 1981; and FRAP also in 1981 (Hutchins 1986). The first two have an empirical orientation similar to that of the Georgetown project, and the solutions tend to be specific to the respective combinations, while FRAP has a transfer component that is based on theoretical linguistic principles.

All three systems are part of an integrated DP environment; they support an abstracting service and also the compilation of terminology and other lexicographic data.

3.2.2 METEO

Since May 25, 1977, METEO, originally developed by the University of Montreal, has been translating weather forecasts from English into French for the Canadian public 24 hours a day, seven days a week (Thouin 1982).

In 1975 the TAUM group (Traduction Automatique de l'Université de Montréal), financed by the Canadian National Research Council for the preceding 10 years, received a grant from the Department of the Secretary of State to develop a system for the translation of weather forecasts. The result of this initiative was METEO 1, which was installed in 1977 and for the next six years voraciously gobbled up forecasts originally written in a moderately restricted English, turning out serviceable French counterparts at a rate of 11,000 words a day (Isabelle & Bourbeau 1988).

By 1983 METEO 1 had been upgraded and converted to a microcomputer by John Chandioux Consultants, Inc. The new system, METEO 2, handles more complex input (Chandioux 1988) and translates 32,000 words a day (Chandioux p.c. 1988). It has an automatic pre-editing component. There is postediting, but the rate of human intervention is only 3.4%, compared with 20% in the case of METEO 1. In October 1988 translation from French into English was introduced, and at the end of the first month of operation output in the reverse direction was already 8,000 words a day.

METEO can be credited with several positive contributions. Besides effecting savings in the cost of translation, it has paved the way for expansion of the meteorological network in Canada. The human translators, even though their ranks have been reduced, are more content with their work;

whereas before they complained of the monotony, now they are involved in the process and have a hand in improving the system.

3.2.3 GETA/B'VITAL

In France the University of Grenoble has been working on MT since 1961. The Centre d'Etudes pour la Traduction Automatique (CETA) was established at that time and in 1972 became the Groupe d'Etudes pour la Traduction Automatique (GETA), which had support from the National Research Center and the Ministries of Defense, Telecommunications, and Industry (Vauquois & Boitet 1988).

Beginning with translation from Russian into French, the project had a strongly theoretical orientation. Initially the goal was to have an interlingua-type representation of the central component of the system. Intermediate structures were produced in the form of trees. Later, however, the emphasis shifted to having simply a transfer component based on a multi-level mix of syntactic and semantic rules (ibid.).

In 1978 the operational environment ARIANE-78, consisting of software to facilitate system development, was introduced.

Subsequent work led to completion of the CALLIOPE system, launched in 1985 and rebaptized B'VITAL (Bernard Vauquois Informatique et Traîtement Automatique des Langues) in 1987, in memory of the director of GETA who died that year. As of 1988 translations were being produced from French into English in the aerospace field and from English into French for several applications (Boitet p.c. 1988).

3.2.4 TITUS

Begun in France in 1969 as an initiative of the Institut Textile, the TITUS project is geared to providing the textile industry with information from the technical literature. Abstracts are translated only once, from their original language into a simplified form of French (Langage Documentaire Canonique--LDC), from which versions are generated in German, Spanish, French, and English (Ananiadou 1987). LDC also serves as an index for the retrieval of data.

With TITUS I, the user had to rewrite the abstracts manually before they were entered in the system. The second version, TITUS II, introduced LDC. This is a sub-language of French consisting of a list of permitted terms, a fixed inventory of function words, and a limited syntax (Hutchins 1986). In TITUS IV, the drafting process is automated: the LDC restrictions are presented interactively as the user enters the text in ordinary French (Ananiadou 1987).

3.2.5 CULT

Another semi-specialized system is CULT (Chinese University Language Translator), developed by the Chinese University of Hong Kong. The first version was implemented in 1969. In 1975 CULT began to be used to translate the *Acta Mathematica Sinica* from Chinese into English, and from 1976, the *Acta Physica Sinica* (Loh & Kong 1979). Both journals are widely circulated throughout the world. CULT also translates from English into Chinese. The project, under the direction of Shiu-Chang Loh, has had support from the Asia Foundation and the Rockefeller Brothers Fund.

CULT originally relied on manual pre-editing of the Chinese input, with intervention in approximately 5% of the text--principally to define sentence and phrase boundaries and provide semantic markers for ambiguous characters (Loh et al. 1978). By 1979 interactive software had been developed for this purpose; when the system encounters input that does not match anything in the dictionary, the user is queried and given the possibility of changing the source text interactively (Loh & Kong 1979). With this form of pre-editing, CULT became the first system to pioneer front-end intervention.

3.2.6 SPANAM/ENGSPAN

The Pan American Health Organization, Regional Office of the World Health Organization in the Americas, began to develop MT for internal use in 1977.

The first system, SPANAM, which translates from Spanish into English, in the beginning had certain features that resembled Georgetown's GAT system and SYSTRAN, although its development was totally independent of these projects. Work done since 1979 has taken modern directions. The development of ENGSPAN, which translates from English to Spanish, was partially supported by a grant from the U.S. Agency for International Development (USAID), which permitted the investigation and implementation of an augmented transition network grammar for the analysis component, as well as the development of transfer and synthesis modules drawing on principles from contemporary linguistics (León & Schwartz 1986, Vasconcellos & León 1988). This grant was the first public manifestation of interest in MT on the part of the U.S. Government since the ALPAC decision of 1966. It was motivated by the need to disseminate information on health and agriculture to Third World countries.

The "smarts" of ENGSPAN have been retrofitted onto SPANAM, and the new version of the older system is now in operation.

Both SPANAM and ENGSPAN are general systems with large dictionaries (approximately 63,000 and 55,000 terms, respectively), and they follow a "try-anything" approach. SPANAM became operational in January 1980 and since that date has been in almost daily use, producing an average of 80,000 words a month, most of which are postedited by professional translators. The texts translated, typically documents written by technical specialists in the Organization, are quite varied in terms of subject matter, vocabulary, language style, and discourse type. In addition to public health, the texts are in many other fields, including agriculture, computers, law, management, sanitary engineering, and the physical sciences. ENGSPAN, which has been operational since 1985, translates similar texts, often for dissemination in Spanish-speaking Latin America. It is also installed at USAID, the International Center for Tropical Agriculture (CIAT) in Colombia, and the International Rice Research Institute (IRRI) in the Philippines.

Figures 4 and 5 show raw translations produced by ENGSPAN and the new SPANAM, respectively.

3.3 Systems Still Under Development

The foregoing inventory has attempted to cover all the systems that are currently installed and in regular use. There are a few that were once operational but have since been shelved, and these have been omitted. The next section will briefly mention some that are still under development.

3.3.1 EUROTRA

Even though it has yet to produce concrete results, the EUROTRA project, sponsored by the Commission of the European Communities, is nevertheless a unique phenomenon in the history of MT if one considers the scope of its objective, the investment it has represented, and the jobs it has generated for computational linguists in Europe.

This mammoth project was born of the need to translate the 42 combinations of the CEC's official languages (Danish, Dutch, English, French, German, Greek, and Italian), which more recently became 72 with the admission of Spain and Portugal. It was felt that to develop SYSTRAN pairs for all these combinations would be a totally impractical undertaking (Wagner 1985). The very thought of having to undertake each combination as a separate effort seemed illogical; it would be more efficient to build an interface that would serve as a fulcrum for translating from many to many.

The initiative started to take shape in 1978 with the establishment of ties with the universities of Grenoble, the Saar, Manchester, and Pisa, each of which assumed responsibility for a specific aspect of the research. It was carried forward under the general direction of Sergei Perschke at CEC headquarters in Luxembourg, with coordination being handled by Margaret King, of the University of Geneva's Institute for Semantic and Cognitive Studies, which served as general secretariat.

By 1981 there were 80 researchers working on EUROTRA. In November 1982 the project received a grant from the CEC equivalent to US$12 million for a program to be carried out over a period of five and a half years (Hutchins 1986). Three stages were envisioned: two years for preparation, two years for basic and applied research, and 18 months for stabilization of the linguistic models and evaluation of the results. This last phase ended in 1988, and it is now planned to continue with the practical development of the initiatives already under way and with other parallel projects that have spun off from the major effort.

As Hutchins points out (1986:271), EUROTRA has derived much of its importance so far from the fact that never before has there been an MT project of such magnitude, or one that has brought together so many linguists and related specialists working simultaneously in different countries.

3.3.2 Other University Systems

In addition to the research associated with EUROTRA, there are other projects which have not seen the light of practical implementation but which nevertheless have played an important role in the evolution of MT.[4]

At the University of Kyoto work on MT from English into Japanese began in 1968 and led to the installation of TITRAN, which was used for some time but is now in mothballs (Nagao p.c. 1988). MU, the official project of the Japanese government, was inaugurated in April 1982 under the direction of Professor Makoto Nagao. Although not yet fully operational, as of 1988 MU was already translating in both directions, English-Japanese and Japanese-English, using dictionaries of more than 70,000 items. The system also includes a term bank. It is to be implemented at the Japan Information Center of Science and Technology.

At the University of the Saar the SUSY project, under the direction of Wolfram Wilss, engaged in the development of MT systems for 15 years, from 1972 until 1987, part of this time in collaboration with EUROTRA. Work was undertaken, in chronological order, on the following combinations: Russian-, French-, English-, Esperanto-, Danish- and Dutch-German, and German-English and -French (Freigang 1987). Saarland's ASCOF project made use of COMSKEE, a programming language developed in-house which is much like LISP or PROLOG (Biewer et al. 1988).

Also in Germany, in research begun at the University of Heidelberg in 1973, the SALAT project has developed an interlingua, with initial concentration on translation from German into French. At the University of Stuttgart the SEMSYN system focuses on translation from Japanese into German.

In England, at the University of Manchester Institute of Science and Technology (UMIST), NTRAN is being developed to generate texts in Japanese. It is intended to be used interactively by monolingual speakers of English. A system in the reverse direction, Japanese-English, is being developed at the University of Sheffield.

In Canada, the TAUM group at the University of Montreal followed its work on METEO with an initiative known as TAUM-AVIATION. The goal of this project, begun in 1976, was to translate a series of maintenance manuals for aircraft hydraulic systems. Like METEO, TAUM-AVIATION was for translation from English into French in a specialized subject area. The prototype was demonstrated in 1979, but the next year the Canadian Government decided to stop funding the project because of the high cost of development (Isabelle & Bourbeau 1988:238).

Work in Canada continues to stress the sublanguage approach. Two systems currently under development are CRITTER (Dymetman & Isabelle 1988), for the translation of data on the livestock market, and MARWORDS, which generates marine weather texts simultaneously in English and French using information from a data base (Kittredge et al. 1988). (GET)

In the U.S., at the University of New Mexico, Yorick Wilks heads the Computing Research Laboratory, where MT development proceeds on the basis of his "preference semantics" and other work in artificial intelligence that he has been pursuing since 1970. The XTRA system translates from English into Chinese (Huang 1985).

Carnegie Mellon University in Pittsburgh established the Center for Machine Translation, under the direction of Jaime Carbonell, in 1986. There is a large team, which includes Sergei Nirenburg, formerly of Colgate University, and Masaru Tomita, from the University of Kyoto. Their system is intended to be many-to-many and makes use of a hierarchically organized knowledge base.

3.3.3 Commercial projects

There are also several major commercial projects that have not yet reached the operational stage. The most feverish activity is going on in Japan (3.1.11). In addition, since 1980 Philips Laboratories in the Netherlands have been working on ROSETTA. This system, based on Montague grammar, will be multilingual, with initial focus on English/Dutch and Spanish to be incorporated later. Also in the Netherlands there is DLT (Distributed Language Translation), being developed under the direction of Toon Witkam since 1982 by the firm BSO. DLT uses Esperanto for its central syntactic representation. The initiative has had support from both the CEC and the Dutch government. The prototype, which currently translates from English into French, was demonstrated in December 1987, and the target date for a commercial product is 1993 (*Language Monthly* 1988).

IBM has work going on at a number of sites. There are also other companies with projects under way, but information tends to be sketchy and is often confidential.

3.4 Other Products

All the MT systems described above were initially designed to run on mainframes, minicomputers, or powerful workstations such as the Symbolics LISP machine. With the advent of personal computers, some of the existing systems were downsized, or adapted, to run on the new hardware. Most of them antedate the microcomputer; certainly none of them started out as a PC product. More recently the mass marketing of the PC has spawned several low-end multilingual packages that are sold by mail. The first to appear was Linguistic Products (Dessau 1986), which in its early releases was offered in Spanish-English or English-Spanish and was entirely contained on one floppy diskette. There was a sizable vocabulary of fully inflected forms, with the entries coded only for one of five possible "parts of speech," and a small set of rules for rearranging the text (Dessau p.c. 1986). Another low-end PC product, TOLTRAN, was demonstrated in English-Spanish by its inventor, Bruce Tolin, but as of the end of the year it was not yet being distributed (Tolin p.c. 1988).

There has also been a wave of multilingual packages containing phrase dictionaries for the generation of business letters. Four such products were announced in 1988: Correspondence, from MultiLingua; LinguaWrite, from A Propos; TransWord, from Tron B.V.; and TickTack, from Primrose (Language Technology Jan/Feb 1989).

Finally, there are hand-held devices that offer to translate phrases for the traveler. At least seven such products are on the market (sold by Langenscheidt, Seiko, Sharp, and Texas Instruments), with vocabularies ranging from 600 to 40,000 items. None of them uses linguistic rules.

These other products have a different history and a different usership than the full-fledged MT systems. There is a enormous potential market for any product that runs on a PC and successfully accomplishes its purpose. Once the larger systems are adapted to this environment, the consumer will be faced with a bewildering array of options, and it will be important to be able to sort them out and understand the different objectives for which the systems and packages have been designed.

4. Evaluation of MT

As it can be seen, there is a plethora of systems and projects, and much duplication of language combinations. The list presented here gives an idea of the considerable progress already achieved in this field. It would appear that there is little terrain which has not been explored at one time or another, with greater or lesser success.

One may well wonder: Why so many? Are there significant differences between one system and another? What are these differences? Are they important enough to warrant the continued introduction of new products, many of which seem to duplicate what is already available? And finally, what are the criteria for judging?

In the 1960s, decisions about MT were often based on analysis of the quality of the output translation, coupled with calculations of cost versus the benefits to be gained. Of course cost is still important today, but we also know that the real value of an MT system goes beyond the quality of the output at any given moment: it depends, more than anything else, on the requirements of the specific application being considered.

It is the entire process that is of interest, not just the product in isolation. The relative importance of different aspects of this process will vary depending on the perspective of the person responsible for the decision: the administrator, the translator, or the client. Administrators will want to serve the interests of their institution, preferably at a lower cost than in the past. Translators will look at the work that is left for human beings to do after the machine has done its part--as well as the ease with which they will be able to adapt to the new working mode, the possibilities for increasing their productivity, the prospects for lightening their load in the future, and, of course, the potential impact on their careers. Clients, on the other hand, want to know only whether the quality and form of the product are adequate for their purpose, and whether their deadlines are met; what mode was used to produce the translations is of little importance to them (Klein 1988). Any evaluation should therefore give priority to these factors.

The assessment of raw MT is fraught with pitfalls. To begin with, the criterion of output quality is problematic. Despite decades of scientific study, the evaluation of translations, whether human or machine, continues to be a vague and uncertain exercise (Rose 1987). To some extent, the definition of error will vary depending on the purpose of the translation and the values of the user community. Moreover, output can be misleading because it is difficult to know the circumstances in which the text was produced (Bédard 1988). For these and other reasons, a formal analysis of the raw product is simply not a sufficient criterion for judging an MT system.

More interesting than direct evaluation of the output is the reaction of a person who has already used the system. In a study by Henisz-Dostert (1979), 58 scientists who had been using translations produced by Georgetown's GAT system over the period 1963-1973 were polled for their reactions. The responses showed a high level of satisfaction: 91.4% found the quality good or acceptable; 93% considered that the translations were informative; 87% preferred MT over human translation, and 96% would recommend MT to their colleagues (ibid. p. 208). In this case the product was judged by the end-users in terms of their actual needs.

On the other hand, raw output, if expertly examined, can yield a typology of shortcomings that are indicative of the system's potential, and of the investment that would be required in order to make the necessary repairs (Vasconcellos 1988c). Also, in translation service where there will be no time or budget for postediting, it is indispensable to test a system's performance on randomly selected texts.

The basic conceptualization and structure of the system, including the depth of dictionary coding, will influence the amount of effort needed, in the short term, to produce translations that are viable and, in the long term, to add other language combinations.

The degree to which a system is specialized is also a major consideration, since it may well determine the system's extensibility to other applications (Shann 1987:89, Lehrberger & Bourbeau 1988). Obviously, if the specialized system is already adequate for the purpose desired, it will be more effective than a general system that has yet to be adapted. But such a coincidence of purposes would be rare, and it is more likely that the specialized system will not lend itself easily to other domains.

For the potential user, it is important to know which language combinations have already been developed, the size of the dictionaries, and the ease with which they can be adapted to the subject areas of concern. For general translations, the dictionaries should have at least 20,000 basic entries of general vocabulary with the coding already incorporated, based on a scheme that is adequate for the application envisaged. There should be a means of adding the user's specialized terminology and, in addition, there should be several ways of specifying alternate translations. Finally, dictionary updating should be relativaly easy to learn and to perform, though not at the expense of precision and linguistic power.

It is also important for the prospective user to consider the total environment into which the system is to be fitted. The system's potential for future growth will be derived in no small measure from the material and human resources that can be counted on in the prospective setting. The needed hardware must be available, including workstations in adequate number. The right kind of human resources are important at all points in the process. The vendor should be prepared to offer continuing support: MT is complex, and ongoing cooperation with the vendor is essential. On the user's side, the translators who will be using the system should be ready to make a long-term commitment; they cannot be temporary staff. They should be prepared to become involved in dictionary improvement as well as in postediting (McElhaney & Vasconcellos 1988, Santangelo 1988), and, when possible, they should participate in making suggestions about future development (Ryan 1988, Myer 1988).

In the final analysis, what distinguishes one system from another is the <u>function that it performs</u> (Vasconcellos 1988c). There is no single "right" way to evaluate machine translation, but the exercise will be more valuable to the extent that it takes into account the needs to be met, the purposes of the institution, and the environment in which MT is to be, or is already being, used. Formal evaluation of the output text, when this is done, should be part of a larger investigation that gives priority to the functional factors that will determine the system's future over the long term: its capacity to grow, and the possibility that the users will be able to make a meaningful contribution to this growth.

5. <u>And Now What?</u>

Finally, it's time to ask: What does the future hold for MT? Where is it headed?

It is reasonably safe to hazard a few projections about: language combinations, types of systems, degrees of specialization, hardware and software, shifts in public opinion, and, finally, the purposes of translation and the clients served.

With regard to the language combinations, we have already seen that recently English has tended to be the source language, whereas in the past, especially prior to 1966, it was the target language. Combinations that do not involve English are beginning to appear, and Japanese is becoming an important focus of attention. More Arabic, Chinese, Korean, and other non-Indo-European languages can be expected. It will be increasingly possible to marry systems that were developed in different circumstances (e.g. White 1988). In general, there will be a greater trend toward cooperation between teams in different parts of the world.

Systems based on a theoretical approach are beginning to reach the commercial market, where, alongside those already in use, their merits will be put to the test. There will be new modalities that make use of interactive interrogation, and systems that generate texts in several languages directly from data bases. And there will be attempts, with relatively limited success, at real-time interpretive MT--that is, the generation of an oral translation of spoken language--as has already been demonstrated with preprogrammed texts (Tomita et al. 1988).

Hutchins (1988b) believes that specialized systems will ultimately prevail over the try-anything type, since the results are easier to predict and therefore more reliable. On the other hand, one can expect that economic pressure will contribute to the continuing proliferation of general systems, even though they may produce inferior results, because demand for the specialized types is necessarily limited.

In terms of hardware, MT will be running on increasingly smaller computers, and for this reason it will be more accessible to the general public. Until recently, a number of factors were holding back the adaptation of the serious MT systems--those designed to translate large volumes of text--to microcomputers. To begin with, the texts often had to be input manually, or else documents had to undergo awkward and expensive conversion from one word-processing package to another. This problem has been greatly diminished through the availability of user-friendly TC and conversion software, coupled with improvements in OCR technology and falling prices thereof. Another obstacle was that dictionaries of the size required for general-purpose translation could not be contained on disks as small as 10 and 20 Mb, which had been the norm for a number of years. Now, thanks to impressive increases in hard disk capacity and innovative compression techniques, these dictionaries can be accommodated and the storage problem is no longer an issue. Even when the dictionaries do fit, however, lookup has been slower than on a mainframe or mini by orders of magnitude. Unless machine lookup is considerably faster than the human translator, the appeal of these systems will be limited. This speed is improving markedly with the general availability of new high-speed chips, UNIX-like operating systems, and more efficient database management. All these trends are already facilitating the implementation of serious MT systems on microcomputers.

Other problems, however, remain to be solved. For the vendors, customer support for dictionary updating is labor-intensive and incompatible with mass marketing strategies. This difficulty is being overcome in part by interactive software, which makes updating not only easier to perform but also faster, more efficient, and more accurate. Dictionary updating on networked PCs can quickly get out of control, but new software support for local area networks (LANs) will help to rationalize the situation. Finally, vendors have been reluctant to release large and deeply coded dictionaries which represent years of work, since these data sets can easily be purloined. Smaller dictionaries, on the other hand, make for MT systems that don't translate as well as they could, and hence bad press for MT. These last difficulties remain to be worked out. In general, however, the hurdles are falling away one by one, and it is not out of line to predict that all the major MT systems will be available on some form of microcomputer before the end of the first half of the 1990s.

Public opinion is shifting in favor of MT. Its availability on microcomputers will contribute to an upsurge in its use, since for the first time it will be within the reach of small translation services and independent translators. In general, it can be expected that, thanks to MT, the volume of translations in the world will increase. There will be continuing interest in texts for commercial and administrative purposes, and there will be much more real-time, or near-real-time, translation for information purposes. The public will come to accept a final product that is less polished. And with all these developments MT will reach out to a much larger public than the small translator/linguist community that is currently involved.

This scenario is not for tomorrow or the day after. Many of these things will take time. Meanwhile, MT will be generating more demand for translation, just as the copying machine generates a demand for copies. It is not unreasonable to envision a future translation market some three times larger than it is at present, with a considerable share of the work still being produced in the traditional way and with translators also postediting machine output and contributing to the further enhancement of MT systems.

ACKNOWLEDGMENTS

The author is grateful to Plínio de Almeida, Veronica Lawson, and Marjorie León for their review of the text and, in each case, their valuable suggestions. In addition, many thanks are due to all the colleagues who kindly supplied information that has been included herein: Claude Bédard, Scott Bennett, Christian Boitet, John Chandioux, Ralph Dessau, Denis Gachot, Iván Guzmán de Rojas, John Hutchins, Makoto Nagao, Ian Pigott, Alan Portela, Tim Rowe, Joann Ryan, John Smart, Ralph Strozza, Bruce Tolin, Peter Toma, Elizabeth Wagner, Alan Weaver, Alex Weiss, Peter Wheeler, and Michael Zarechnak.

NOTES

[1]Explanations of the MT process are found in Macdonald 1979, Hutchins 1986 and Lehrberger & Bourbeau 1988, and there are descriptions of several systems in King 1987 and Slocum 1988.

[2]Information on the history of MT comes from Zarechnak (1979:47-57), Hutchins (1986), and the author's personal experience.

[3]For a detailed description of the project, see the final report (Macdonald 163).

[4]The information in this section is based in large part on Hutchins (1986).

[5]Some of the predictions in this section were inspired by, or coincide with, Hutchins (1988b).

REFERENCES

Automatic Language Processing Advisory Committee. 1966. Language and Machines: Computers in Translation and Linguistics; A Report by the Automatic Language Processing Advisory Committee (ALPAC). Washington, D.C.: National Academy of Sciences, Division of Behavioral Sciences. National Research Council Publication 1416.

Ananiadou, Sophie. 1987. "A Brief Survey of Some Current Operational Systems." In: Machine Translation Today: The State of the Art, ed. Margaret King. pp. 171-191.

Bédard, Claude. 1988. "You Trust Your Mother, But YOU Cut the Cards." Language Technology 7:26-27, May-June.

Beesley, Kenneth R. 1988. "Language Identifier: A Computer Program for Automatic Natural-Language Identification of On-Line Text." In: Proceedings of the 29th Annual Conference of the American Translators Association, ed. Deanna L. Hammond. pp. 47-54.

Ben-Ari, Danit, Daniel M. Berry & Mori Rimon. 1988. "Translation Ambiguity Rephrased." Proceedings, II International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (Pittsburgh, 12-14 June 1988). Pittsburgh: Carnegie Mellon University, Center for Machine Translation.

Bennett, Winfield S., & Jonathan Slocum. 1988. "The LRC Machine Translation System." In: Machine Translation Systems, ed. Jonathan Slocum. pp. 111-140.

Bergeron, Yvon. 1988. "MT in On-Line Database Retrieval: A Case Study from Canada." Presentation, 29th Annual Conference of the American Translators Association (Seattle, 12-16 October 1988).

Biewer, Axel, et al. 1988. "ASCOF: A Modular Multilevel System for French-German Translation." In: Machine Translation Systems, ed. Jonathan Slocum. pp. 49-84.

Booth, Andrew D. 1985. Letter to W.J. Hutchins dated 16 May 1985, cited in Hutchins 1986.

Bostad, Dale. 1982. "Quality Control Procedures in Modification of the Air Force Russian-English MT System." In: Practical Experience of Machine Translation, ed. Veronica Lawson. pp. 129- 133.

Bostad, Dale A. 1987. "Machine Translation: The USAF Experience." In: Proceedings of the 28th Annual Conference of the American Translators Association, ed. Karl Kummer. pp. 435-443.

Brislin, Richard W., ed. 1976. Translation: Applications and Research. New York: Gardner Press (John Wiley & Sons).

Cave, Jon R. 1988. "New Developments in LOGOS^{tm} Machine Translation." In: Proceedings of the 29th Annual Conference of the American Translators Association, ed. Karl Kummer. pp. 55-9.

Chandioux, Jean, & M.F. Guérard. 1981. "Météo: un système à l'épreuve du temps." Meta 26(1):17-22.

Chandioux, John. 1988. "Météo^{tm}: An Operational Machine Translation System." Presentation, RIAO 88 (Massachusetts Institute of Technology, March 1988).

Dessau, Ralph, & George Mallard. 1986. "Automated Language Translation for Personal Computers." In: Proceedings of the 27th Annual Conference of the American Translators Association, ed. Karl Kummer. pp. 125-128.

Dymetman, Marc, & Pierre Isabelle. 1988. "Reversible Logic Grammars for Machine Translation." Proceedings II International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (Pittsburgh, 12-14 June 1988). Pittsburgh: Carnegie Mellon University, Center for Machine Translation.

Freigang, Karl-Heinz. 1987. 'Research on Machine Translation at the University of Saarbrücken'. In: A Profession on the Move, ed. Catriona Picken. pp. 83-94.

Giuliano, V.E. 1961. "A Formula Finder for the Automatic Synthesis of Translation Algorithms." Mechanical Translation 6:11-24. Cited in Hutchins 1986.

Good, Robert L. 1988. "Automated Lookup: AutoTerm of ALP Systems." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 87-91.

Guzmán de Rojas, Iván. 1985. "Logical and Linguistic Problems of Social Communication with the Aymara People." Ottawa: International Development Research Centre. IDRC Manuscript Reports 66e.

Guzmán de Rojas, Iván. 1988. "ATAMIRI: Interlingual MT Using the Aymara Language." In: New Directions in Machine Translation, ed. D. Maxwell et al. pp. 1-7.

Hammond, Deanna L., ed. 1988. Proceedings of the 29th Annual Conference of the American Translators Association (Seattle, 12-16 October 1988). Medford, N.J.: Learned Information, Inc.

Henisz-Dostert, Bozena. 1979. Users' Evaluation of Machine Translation. Part III of Machine Translation The Hague, Paris, New York: Mouton. Trends in Linguistics, Studies and Monographs 11. pp. 147-244.

Huang, Xiuming. 1985. "Machine Translation in the Semantic Definite Clause Grammar Formalism." In: Proceedings of the Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages, ed. Sergei Nirenburg. pp. 135-144.

Hundt, Michael G. 1982. "Working with the Weidner Machine-Aided Translation System." In: Practical Experience of Machine Translation, ed. Veronica Lawson. pp. 129- 133.

Hutchins, W. John. 1986. Machine Translation: Past, Present, Future. New York: Halsted Press (Wiley & Sons).

Hutchins, W. John. 1988a. "Future Perspectives in Translation Technologies." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 223-240.

Hutchins, W. John. 1988b. "Historical Perspectives." Presentation, II International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (Pittsburgh, 12-14 June 1988).

Isabelle, Pierre, & Laurent Bourbeau. 1988. In: Machine Translation Systems, ed. Jonathan Slocum. pp. 237-263.

King, Margaret, ed. 1987. Machine Translation Today: The State of the Art. Proceedings of the Third Lugano Tutorial (Lugano, 2-7 April 1984). Edinburgh: Edinburgh University Press. Information Technology Series, 2.

Kittredge, Richard, et al. 1988. "Multi-Lingual Text Generation and the Meaning-Text Model." Proceedings, II International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (Pittsburgh, 12-14 June 1988). Pittsburgh: Carnegie Mellon University, Center for Machine Translation.

Klein, Fred. 1988. "Factors in the Evaluation of MT: A Pragmatic Approach." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 198-202.

Kummer, Karl, ed. 1986. Proceedings of the 27th Annual Conference of the American Translators Association (Cleveland, 16-19 October 1986). Medford, N.J.: Learned Information, Inc.

Kummer, Karl, ed. 1987. Proceedings of the 28th Annual Conference of the American Translators Association (Albuquerque, 8-11 October 1987). Medford, N.J.: Learned Information, Inc.

Language Monthly. 1988. "BSO Launch Their DLT Prototype." 52:4-5.

Language Technology. 1989. "Metal Translation Heavies." 11:9, Jan-Feb.

Language Technology. 1989. "Take a Letter, Mac." 11:10, Jan-Feb.

Lawson, Veronica. 1982. "Machine Translation and People." In her: Practical Experience of Machine Translation. pp. 3-9.

Lawson, Veronica, ed. 1982. Practical Experience of Machine Translation. Amsterdam: North-Holland.

Lawson, Veronica. 1984. "Users of Machine Translation System Report Increased Output." Language Monthly 11:6-10.

Lawson, Veronica. 1985. Tools for the Trade: Translating and the Computer 5. London: Aslib.

Lawson, Veronica. 1988. "A Translator's Map of Machine Translation." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 106-115.

Lehrberger, John, & Laurent Bourbeau. 1988. Machine Translation: Linguistic Characteristics of MT Systems and General Methodology of Evaluation. Amsterdam/Philadelphia: John Benjamins.

León, Marjorie, & Lee Ann Schwartz. 1986. "Integrated Development of English-Spanish Machine Translation: From Pilot to Full Operational Capability." Technical Report of Grant DPE-5543-G-SS-3048-00 from the U.S. Agency for International Development. Washington, D.C.: Pan American Health Organization.

Lévy, Moïse. 1988. Presentation, 29th Annual Conference of the American Translators Association (Seattle, 12-16 October 1988).

Lippmann, Edward O. 1986. "Artificial Intelligence in Linguistics: Text and Style Critiquing Computer Program." In: Proceedings of the 27th Annual Conference of the American Translators Association, ed. Karl Kummer. pp. 157-159.

Loh, S.C., & L. Kong. 1979. "An Interactive On-Line Machine Translation System (Chinese into English)." In: Translating and the Computer, ed. Barbara M. Snell. pp. 135-148.

Loh, S.C., L. Kong, and H.-S. Hung. 1978. "Machine Translation of Chinese Mathematical Articles." ALLC Bulletin 6(2):111-120.

Macdonald, R.R. 1963. "Georgetown University Machine Translation Research Project: General Report, 1952-1963." Washington, D.C.: Georgetown University. Occasional Papers on Machine Translation, 30.

Macdonald, R. Ross. 1979. "The Problem of Machine Translation." Part II of Machine Translation The Hague, Paris, New York: Mouton. Trends in Linguistics, Studies and Monographs 11. pp. 147-244.

McElhaney, Terrence, & Vasconcellos, Muriel. 1988. "The Translator and the Postediting Experience." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 140-148.

Magnusson-Murray, Ulla. 1985. "Operational Experience of the Machine Translation Service." In: Tools for the Trade, ed. Veronica Lawson. pp. 171-180.

Marchuk, Yu. N. 1977. "Machine Translation in the All-Union Translation Centre." In: Overcoming the Language Barrier, proceedings, III European Congress on Information Systems and Networks (Luxembourg, 3-6 May 1977), sponsored by the Commission of the European Communities. Munich: Verlag Dokumentation. pp. 623-624.

Marchuk, Yu. N. 1984. "Machine Translation in the USSR." Paper sent to the Conference on Methodology and Techniques of Machine Translation (Cranfield, 13-15 February 1984), pp 195-209. Subsequently published in International Forum on Information and Documentation 9(2)3-8.

Maxwell, Dan, Klaus Schubert & A.P.M. Witkam (eds.). 1988. New Directions in Machine Translation. Dordrecht/Providence: Foris Publishers.

Myer, Ingrid. 1988. "The Role and Design of Computer Studies in the Research-Oriented Translation Program." Paper presented at the Conference on Teaching Computers and the Humanities Courses (Oberlin, Ohio, 16-18 June 1988).

Nirenburg, Sergei. 1985. Proceedings of the Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (Colgate, New York, 14-16 August). Colgate University.

Picken, Catriona, ed. 1987. Translating and the Computer 8: A Profession on the Move. London: Aslib.

Pigott, Ian M. 1988. "MT in Large Organizations: Systran at the Commission of the European Communities." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 159-166.

Rose, Marilyn Gaddis, ed. 1987. Translation Excellence: Assessment, Achievement, Maintenance. American Translators Association Scholarly Monograph I. Binghamton (N.Y.): University Center at Binghamton (SUNY).

Ruffino, Richard, & Peter DeMauro. 1986. "The Impact of Controlled English on Machine Translation." In: World Systran Conference. Special issue of Terminologie et Traduction. Luxembourg: Commission des Communautés européennes.

Russo, Maria J. 1988. "Integration of MT in Workstation Publishing." Presentation, 29th Annual Conference of the American Translators Association (Seattle, 12-16 October 1988).

Ryan, Joann. 1988. "The Role of the Translator in Making an MT System Work: Perspective of the Developer." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 127-132.

Santangelo, Susana. 1988. "Making an MT System Work: Perspective of the Translator." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 133-139.

Shaefer, Leonard A. 1988. "MT and the Independent Translator." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 124-126.

Shann, Patrick. 1987. "Machine Translation: A Problem of Linguistic Engineering or of Cognitive Modelling?" In: Machine Translation Today, ed. Margaret King. pp. 71-90.

Slocum, Jonathan. 1988. "A Survey of Machine Translation." In his: Machine Translation Systems. pp. 1-47.

Slocum, Jonathan, ed. 1988. Machine Translation Systems. Cambridge, New York, etc.: Cambridge University Press.

Smart, John M. 1988. "Getting Smart in Many Languages: MT with an Option of Preprocessing." In: Technology as Translation Strategy, ed. M. Vasconcellos. pp. 124-126.

Snell, Barbara M., ed. 1979. Translating and the Computer. Amsterdam, New York, etc.: North-Holland.

Thouin, Benoît. 1982. "The METEO System." In: Practical Experience of Machine Translation, ed. Veronica Lawson. Amsterdam: North-Holland. pp. 39-44.

Tikhomirov, B.D. 1984. "AMPAR and NERPA Systems of Machine Translation: Some Specific Features of Software and Technology." Paper sent to the Conference on Methodology and Techniques of Machine Translation (Cranfield, 13-15 February 1984), pp 210-217. Subsequently published in International Forum on Information and Documentation 9(2)9-11.

Toma, Peter. 1976. "An Operational Machine Translation System." In: Translation: Applications and Research, ed. Richard W. Brislin. pp. 247-260.

Tomita, Masaru. 1986. "Sentence Disambiguation by Asking." Computers and Translation 1(1):39-52.

Tomita, Masaru, et al. 1988. "The Universal Parser Compiler and Its Application to the Speech Translation System." Proceedings, II International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (Pittsburgh, 12-14 June 1988). Pittsburgh: Carnegie Mellon University, Center for Machine Translation.

Turing, Alan M. 1948. "Intelligent Machinery." In: Machine Intelligence, 5, eds. B. Beltzer & D. Michie. Edinburgh: Edinbrugh University Press, 1969. pp. 3-23. Cited in Hutchins 1986.

Vasconcellos, Muriel. 1984. "Machine Translation at the Pan American Health Organization: A Review of Highlights and Insights." Newsletter of the British Computer Society Natural Language Translation Specialist Group, May 1984.

Vasconcellos, Muriel. 1985. "Management of the Machine Translation Environment: Interaction of Functions at the Pan American Health Organization." In: Tools for the Trade, ed. Veronica Lawson. pp. 115-129.

Vasconcellos, Muriel. 1986. "Functional Considerations in the Postediting of Machine-translated Output: Dealing with V(S)O versus SVO." Computers and Translation 1(1):21-38.

Vasconcellos, Muriel. 1987a. "Postediting On-Screen: Machine Translation from Spanish into English." In: A Profession on the Move, ed. Catriona Picken. pp. 133-146.

Vasconcellos, Muriel. 1987b. "A Comparison of MT Postediting and Traditional Revision." In: Proceedings of the 28th Annual Conference of the American Translators Association, ed. Karl Kummer. pp. 409-416.

Vasconcellos, Muriel. 1988a. "In-House Machine Translation to Meet PAHO Translation Needs: Report of the Controlled Study." Internal document. Washington, D.C.: Pan American Health Organization.

Vasconcellos, Muriel. 1988b. "M. King (ed.) Machine Translation Today: The State of the Art." Computers and Translation 3(2):162-170.

Vasconcellos, Muriel. 1988c. "Factors in the Evaluation of MT: Formal vs. Functional Approaches." In her: Technology as Translation Strategy. pp. 203-213.

Vasconcellos, Muriel, ed. 1988. Technology as Translation Strategy, American Translators Association Scholarly Monograph II. Binghamton (N.Y.): University Center at Binghamton (SUNY).

Vasconcellos, Muriel, & Marjorie León. 1988. "SPANAM and ENGSPAN: Machine Translation at the Pan American Health Organization." In: Machine Translation Systems, ed. Jonathan Slocum. pp. 187-235.

Vauquois, Bernard, & Christian Boitet. 1988. "Automated Translation at Grenoble University." In: Machine Translation Systems, ed. Jonathan Slocum. pp. 85-110.

Wagner, Elizabeth. 1985. "Rapid Post-Editing of Systran." In: Tools for the Trade: Translating and the Computer 5, ed. Veronica Lawson. pp. 199-213.

Walraff, Barbara. 1988. "The Literate Computer." Atlantic Monthly, Jan 1988, 64-71.

Weaver, Alan 1988. "Two Aspects of Interactive Translation." In: Technology the Translation Strategy, ed. M. Vasconcellos. pp. 116-123.

Wheeler, Peter J. 1988. "The Translator and the Dictionary Experience." In: Technology as Translation Strategy, ed. Muriel Vasconcellos. pp. 149-159.

White, John S., et al. 1988. "Application of the Natural Language Interface to a Machine Translation Problem." Proceedings, II International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages (Pittsburgh, 12-14 June 1988). Pittsburgh: Carnegie Mellon University, Center for Machine Translation.

Whitelock, Peter J. 1987. "Japanese Machine Translation in Japan and the Rest of the World." In: A Profession on the Move, ed. Catriona Picken. pp. 147-159.

Zarechnak, Michael. 1979. "History of Machine Translation." Part I of Machine Translation. The Hague, Paris, New York: Mouton. Trends in Linguistics, Studies and Monographs 11. pp. 1-87.