Translation Journal

Volume 8, No. 3
July 2004

Translators & Computers

Machine Translation and Computer-Assisted Translation:

a New Way of Translating?

by Olivia Craciunescu, Constanza Gerding-Salas, Susan Stringer-O'Keeffe

Olivia Craciunescu was born in 1978 in Piatra-Neamt, a small city in the northeast of Romania. She came to love foreign languages thanks to her mother—a French teacher—and years later she chose to follow the same path.

After graduating from High School, she entered "Alexandru Ioan Cuza" University of Iasi, French - English Department. As a student, she had the opportunity to collaborate with different institutions as a translator and interpreter, thus becoming aware of the importance of translation in international exchanges.

During the last two years at University, she went deeper into the theoretical and practical study of translation, taking courses on translation strategies and techniques, terminological research, interlinguistic translation, simultaneous and consecutive interpreting.

After graduation, she won an Erasmus scholarship in Poitiers, France where she spent a year and obtained a DESS degree (Diplôme d'Etudes Supérieures Spécialisées) in "Ingénierie des Médias pour l'Education". It was on this occasion that she became interested in computer programs and tried to relate them to translation.

Upon her return to Romania, she began a second Master Degree in Translation Studies and Interpreting and decided at the same time to start her own business, opening a translation office. At present, she works as a sworn translator and conference interpreter from French and English into Romanian.

Olivia Craciunescu can be reached at:
olivia_craciunescu@yahoo.com
or mediatrad@caramail.com.

Constanza Gerding-Salas graduated as an English teacher

Abstract

This paper begins with a brief analysis of the importance of translation technology in different spheres of modern life, followed by a concise history of machine and computer-assisted translation. It then describes the technology available to translators in this first decade of the twenty-first century and examines the negative and positive aspects of machine translation and of the main tools used in computer-assisted translation: electronic dictionaries, glossaries, terminology databases, concordances, on-line bilingual texts and translation memories. Finally the paper considers the impact of these new technologies on the professional translator, concluding that s/he will need to acquire new skills in order to remain efficient and competitive in the field.

The Need for Translation Technology

Advances in information technology (IT) have combined with modern communication requirements to foster translation automation. The history of the relationship between technology and translation goes back to the beginnings of the Cold War, as in the 1950s competition between the United States and the Soviet Union was so intensive at every level that thousands of documents were translated from Russian to English and vice versa. However, such high demand revealed the inefficiency of the translation process, above all in specialized areas of knowledge, increasing interest in the idea of a translation machine. Although the Cold War has now ended, and despite the importance of globalization, which tends to break down cultural, economic and linguistic barriers, translation has not become obsolete, because of the desire on the part of nations to retain their independence and cultural identity, especially as expressed through their own language. This phenomenon can clearly be seen within the European Union, where translation remains a crucial activity.

> the Internet with its universal access to information and instant communication between users has created a physical and geographical freedom for translators that was inconceivable in the past.

IT has produced a screen culture that tends to replace the print culture, with printed documents being dispensed with and information being acceessed and relayed directly through computers (e-mail, databases and other stored information). These computer documents are instantly available and can be opened and processed with far greater flexibility than printed matter, with the result that the status of information itself has changed, becoming either temporary or permanent according to need. Over the last two decades we have witnessed the enormous growth of information technology with the accompanying advantages of speed, visual impact, ease of use, convenience, and cost-effectiveness. At the same time, with the development of the global market, industry and commerce function more than ever on an international scale, with increasing freedom and flexibility in terms of exchange of products and services. The nature and function of translation is inevitably affected by these changes. There is the need for countries to cooperate in many spheres, such as ecological (Greenpeace), economic (free trade agreements) humanitarian (Doctors without Borders) and educational (exchange programs), etc. Despite the importance of English, there is the commonly-held belief that people have the right to use their own language, yet the diversity of languages should not be an obstacle to mutual understanding. Solutions to linguistic problems must be found in order to allow information to circulate freely and to facilitate bilateral and multilateral relationships.

Thus different aspects of modern life have led to the need for more efficient methods of translation. At the present time the demand for translations is not satisfied because there are not enough human translators, or because individuals and organizations do not recognize translation as a complex activity requiring a high level of skill, and are therefore not prepared to pay what it

from Universidad Austral de Chile, Valdivia in 1977. She worked in a college that trained executive bilingual secretaries in Santiago, Chile's capital for ten years. There she first became acquainted with translation, mainly in the field of business writing.

She lived in Stockholm, Sweden, between 1989 and 1992. There she became a linguistic mediator for some fellow countrymen, translating/interpreting legal matters, bank documents, official forms, etc.

Back in Chile, she worked at Universidad de Los Lagos, Osorno, between 1995 and 2002, where she taught contrastive grammar, translation studies, applied grammar and English as a foreign language for the English Teaching and Translation Program.

At present, she works at the Department of Foreign Languages, Universidad de Concepción, Concepción, where she teaches translatology, English-Spanish contrastive analysis, translation workshops and linguistics for translators. She is also member of TermUdeC, a terminology research group at Universidad de Concepción. Currently she is writing her Ph.D. thesis in Education.

Ms. Gerding-Salas can be reached at cgerding@udec.cl.

Susan Stringer-O´Keeffe has undergraduate degrees in French and History (Melbourne, Australia), Applied Linguistics (Besançon, France) and German (Paris-Nanterre), an M.A. in French (London) and a Ph.D. in French and German (Colorado) as well as completed graduate courses in Latin American literature in Spanish (Michigan State). She is the author of *The Senegalese Novel by Women* (Peter Lang, 1998) and has published articles on literature and translation. She taught languages for nine years at the University of Michigan-Flint and has been a literature and translation professor at universities in Chile since 1998. She has worked on a number of translation projects with government institutions.

Ms. Stringer can be reached at sstringer@udec.cl.

Front Page

Select one of the previous 28 issues.

Select an issue:

is worth. In other words, translation is sometimes avoided because it is considered to be too expensive. In part, human translation is expensive because the productivity of a human being is essentially limited. Statistics vary, but in general to produce a good translation of a difficult text a translator cannot process more than 4-6 pages or 2,000 words per day. The economic necessity of finding a cheaper solution to international exchange has resulted in continuing technological progress in terms of translation tools designed to respond to the translator's need for immediately-available information and non-sequential access to extensive databases.

This paper aims at examining the new technologies (machine translation, electronic dictionaries, terminology databases, bilingual texts, grammatical concordances, and translation memories) in order to determine whether they change the relationship between the translator and the texts, and if so, then in what way. We will try to answer the following questions:

- Which computer tools are genuinely useful to translators?
- Do the new technologies threaten the livelihood of the translator?
- Does automation imply the disappearance of translation as we know it?

A Short History of Machine Translation

It was not until the twentieth century that the idea of creating automatic dictionaries appeared as a solution to the problem of linguistic barriers. In the 1930s two researchers worked independently towards the same goal: the Franco-Armenian George Artsrouni and the Russian Petr Smirnov-Troyanskii. The latter was the more important of the two because he developed the idea that three stages are necessary for a system of automatic translation: first an editor who knows the source language analyzes the words and converts them into base forms according to their syntactic functions; then a machine organizes the base forms into equivalent sequences in the target language; finally, this rough version is corrected by a second editor, familiar with the target language. Despite the significance of Troyanskii's work, it remained generally unknown until the late 1950s.

The invention of the computer led very quickly to attempts to use it for the translation of natural languages. A letter from Warren Weaver to the computer specialist Norbert Wiener in March 1947 is considered to mark the beginning of this process. Two years later, in July 1949, Weaver publicized his ideas on the applications of the computer to translation and shortly afterwards a number of universities in the United States initiated research into the field of machine translation. In 1954 the first feasibility trial was carried out as a joint project between IBM and the University of Georgetown. Although very limited in scope, the demonstration was considered a success, leading to the financing of other projects, both in the US and the rest of the world. The first versions of machine translation programs were based on detailed bilingual dictionaries that offered a number of equivalent words in the target language for each word listed in the source language, as well as a series of rules on word order. The complexity of the task made it necessary for developers to continue improving the programs because of the need for a more systematic syntactical focus. Projects were based on advances in linguistics, especially on the development of transformational generative grammar models that appeared to offer new possibilities for machine translation.

However, initial optimism soon disappeared. Researchers began to think that the semantic barriers were insurmountable and no longer saw a solution on the near horizon to the problem of machine translation. IBM and the University of Washington produced an operating system called Mark II, but the results were disappointing. By 1964 the US government was becoming so concerned about the inefficiency of machine translation programs that it created the ALPAC (Automatic Language Processing Advisory Committee) to evaluate them. In 1966 this committee produced a highly critical report that claimed that machine translation was slow, inefficient and twice as expensive as human translation, concluding that it was not worth investing money in research in this field. Nevertheless, the report stressed the need to encourage the development of tools to assist the translation process, such as computer dictionaries, databases etc. Although criticized for its lack of objectivity and vision, the ALPAC report led to a freeze on research into machine translation in the US for more than a decade. However, research continued in Canada, France and Germany and two machine translation systems came into being several years later: Systran, used by the European Union Commission and Taum-météo, created by the University of Montreal to translate weather forecasts from French to English.

Important advances occurred during the 1980s. The administrative and commercial needs of multilingual communities stimulated the demand for translation, leading to the development in countries such as France, Germany, Canada and Japan of new machine translation systems such as Logos (from German to French and vice versa) and the internal system created by the Pan-American Health Organization (from Spanish to English and vice versa), as well as a number of systems produced by Japanese computer companies. Research also revived in the 1980s because large-scale access to personal computers and word-processing programs produced a market for less expensive machine translation systems. Companies such as ALPS, Weidner, Globalink (North America and Europe), Sharp, NEC, Mitsubishi, Sanyo (Japan) needed these programs. Some of the most important projects were GETA-Ariane (Grenoble), SUSY (Saarbrücken), MU (Kyoto), and Eurotra (the European Union)

The beginning of the 1990s saw vital developments in machine translation with a radical change in strategy from translation based on grammatical rules to that based on bodies of texts and examples (for example, the Reverso Program). Language was no longer perceived as a static entity governed by fixed rules, but as a dynamic corpus that changes according to use and users, evolving through time and adapting to social and cultural realities. To this day machine

translation continues to progress. Large companies are now using it more, which also increases software sales to the general public. This situation has led to the creation of on-line machine translation services such as Altavista, which offer rapid email services, web pages, etc. in the desired language, as well as to the availability of multilingual dictionaries, encyclopaedias, and free, direct-access terminology databases.

The Translation Market

The development of machine translation is based on supply and demand. On the one hand, there is new technology available, and on the other, political, social and economic need for change. Yet, despite the advances, machine translation still represents only a tiny percentage of the market. At the beginning of the 1990s the translation market was as follows (Loffler-Laurian, 1996):

	HUMAN TRANSLATION	MACHINE TRANSLATION
Europe & the United States	300 million pages	2.5 million pages
Japan	150 million pages	3.5 million pages

It can be seen that only 6 million pages were translated through machine translation, compared with 450 million through human translation, i.e. MT represented only 1.3% of the total. Market analysts predict that this percentage will not change radically by 2007. They say that machine translation will remain only about 1% of an over US $10 billion translation marketplace (Oren, 2004). The languages for which there was most translation demand in 1991 were:

	English	Japanese	French	German	Russian	Spanish	Others
As source lang.	48%	32%	8%	5%	2%	---	5%
As target lang.	45%	24%	12%	---	5%	10%	4%

As expected, English dominates the market. The importance of Japanese reflects the role of Japan in technology and foreign trade, which accounted for two-thirds of translation volume at the end of the 1990s:

Technology	Foreign Trade	Science	Teaching	Literature	Journals	Business Adminis-tration
40%	25%	10%	10%	5%	5%	5%

At this stage it is important to make a distinction between two terms that are closely related and that tend to confuse non-specialists: machine translation (MT) and computer-assisted translation (CAT). These two technologies are the consequence of different approaches. They do not produce the same results, and are used in distinct contexts. MT aims at assembling all the information necessary for translation in one program so that a text can be translated without human intervention. It exploits the computer's capacity to calculate in order to analyze the structure of a statement or sentence in the source language, break it down into easily translatable elements and then create a statement with the same structure in the target language. It uses huge plurilingual dictionaries, as well as corpora of texts that have already been translated. As mentioned, in the 1980s MT held great promises, but it has been steadily losing ground to computer-assisted translation because the latter responds more realistically to actual needs.

CAT uses a number of tools to help the translator work accurately and quickly, the most important of which are terminology databases and translation memories. In effect, the computer offers a new way of approaching text processing of both the source and target text. Working with

a digital document gives us non-sequential access to information so that we can use it according to our needs. It becomes easy to analyze the sentences of the source text, to verify the context in which a word or a text is used, or to create an inventory of terms, for example. Likewise, any part of the target text can be modified at any moment and parallel versions can be produced for comparison and evaluation. All these aspects have profound implications for translation, especially in terms of assessing the results, since the translator can work in a more relaxed way because of the greater freedom to make changes at any time while the work is in progress.

It is important to stress that automatic translation systems are not yet capable of producing an immediately useable text, as languages are highly dependant on context and on the different denotations and connotations of words and word combinations. It is not always possible to provide full context within the text itself, so that machine translation is limited to concrete situations and is considered to be primarily a means of saving time, rather than a replacement for human activity. It requires post-editing in order to yield a quality target text.

Cognitive Processes

To understand the essential principles underlying machine translation it is necessary to understand the functioning of the human brain. The first stage in human translation is complete comprehension of the source language text. This comprehension operates on several levels:

- Semantic level: understanding words out of context, as in a dictionary.
- Syntactic level: understanding words in a sentence.
- Pragmatic level: understanding words in situations and context.

Furthermore, there are at least five types of knowledge used in the translation process:

- Knowledge of the source language, which allows us to understand the original text.
- Knowledge of the target language, which makes it possible to produce a coherent text in that language.
- Knowledge of equivalents between the source and target languages.
- Knowledge of the subject field as well as general knowledge, both of which aid comprehension of the source language text.
- Knowledge of socio-cultural aspects, that is, of the customs and conventions of the source and target cultures.

Given the complexity of the phenomena that underlie the work of a human translator, it would be absurd to claim that a machine could produce a target text of the same quality as that of a human being. However, it is clear that even a human translator is seldom capable of producing a polished translation at first attempt. In reality the translation process comprises two stages: first, the production of a rough text or preliminary version in the target language, in which most of the translation problems are solved but which is far from being perfect; and second, the revision stage, varying from merely re-reading the text while making minor adjustments to the implementation of radical changes. It could therefore be said that MT aims at performing the first stage of this process in an automatic way, so that the human translator can then proceed directly to the second, carrying out the meticulous and demanding task of revision. The problem is that the translator now faces a text that has not been translated by a human brain but by a machine, which changes the required approach because the errors are different. It becomes necessary to harmonize the machine version with human thought processes, judgements and experiences. Machine translation is thus both an aid and a trap for translators: an aid because it completes the first stage of translation; a trap because it is not always easy for the translator to keep the necessary critical distance from a text that, at least in a rudimentary way, is already translated, so that mistakes may go undetected. In no sense should a translation produced automatically be considered final, even if it appears on the surface to be coherent and correct.

Machine Translation Strategies

Machine translation is an autonomous operating system with strategies and approaches that can be classified as follows:

- the direct strategy
- the transfer strategy
- the pivot language strategy

The direct strategy, the first to be used in machine translation systems, involves a minimum of linguistic theory. This approach is based on a predefined source language-target language binomial in which each word of the source language syntagm is directly linked to a corresponding unit in the target language with a unidirectional correlation, for example from English to Spanish but not the other way round. The best-known representative of this approach is the system created by the University of Georgetown, tested for the first time in 1964 on translations from Russian to English. The Georgetown system, like all existing systems, is based on a direct approach with a strong lexical component. The mechanisms for morphological analysis are highly developed and the dictionaries extremely complex, but the processes of syntactical analysis and disambiguation are limited, so that texts need a second stage of translation by human translators. The following is an example that follows the direct translation model:

Source language text					
La	jeune	fille	a acheté	deux	livres

Breakdown in source language					
La	jeune	fille	acheter	deux	livre
Lexical Transfer					
The	young	girl	buy	two	book
Adaptation in target language					
The	young	girl	bought	two	books

There are a number of systems that function on the same principle: for example SPANAM, used for Spanish-English translation since 1980, and SYSTRAN, developed in the United States for military purposes to translate Russian into English. After modification designed to improve its functioning, SYSTRAN was adopted by the European Community in 1976. At present it can be used to translate the following European languages:

- Source languages: English, French, German, Spanish, Italian, Portuguese, and Greek.
- Target languages: English, French, German, Spanish, Italian, Portuguese, Greek, Dutch, Finnish, and Swedish.

In addition, programs are being created for other European languages, such as Hungarian, Polish and Serbo-Croatian.

Apart from being used by the European Commission, SYSTRAN is also used by NATO and by Aérospatiale, the French aeronautic company, which has played an active part in the development of the system by contributing its own terminology bank for French-English and English-French translation and by financing the specialized area related to aviation. Outside Europe, SYSTRAN is used by The United States Air Force because of its interest in Russian-English translation, by the XEROX Corporation, which adopted machine translation at the end of the 1970s and which is the private company that has contributed the most to the expansion of machine translation, and General Motors, which through a license from Peter Toma is allowed to develop and sell the applications of the system on its own account. It should be noted that in general the companies that develop direct machine translation systems do not claim that they are designed to produce good final translations, but rather to facilitate the translator's work in terms of efficiency and performance (Lab, p.24).

The transfer strategy focuses on the concept of "level of representation" and involves three stages. The analysis stage describes the source document linguistically and uses a source language dictionary. The transfer stage transforms the results of the analysis stage and establishes the linguistic and structural equivalents between the two languages. It uses a bilingual dictionary from source language to target language. The generation stage produces a document in the target language on the basis of the linguistic data of the source language by means of a target language dictionary.

The transfer strategy, developed by GETA (Groupe d'Etude pour la Traduction Automatique / Machine Translation Study Group) in Grenoble, France, led by B. Vauquois, has stimulated other research projects. Some, such as the Canadian TAUM-MÉTÉO and the American METAL, are already functioning. Others are still at the experimental stage, for example, SUSY in Germany and EUROTRA, which is a joint European project. TAUM, an acronym for Traduction Automatique de l'Université de Montréal (University of Montreal Machine Translation) was created by the Canadian Government in 1965. It has been functioning to translate weather forecasts from English to French since 1977 and from French to English since 1989. One of the oldest effective systems in existence, TAUM-MÉTÉO carries out both a syntactic and a semantic analysis and is 80% effective because weather forecasts are linguistically restricted and clearly defined. It works with only 1,500 lexical entries, many of which are proper nouns. In short, it carries out limited repetitive tasks, translating texts that are highly specific, with a limited vocabulary (although it uses an exhaustive dictionary) and stereotyped syntax, and there is perfect correspondence from structure to structure.

The pivot language strategy is based on the idea of creating a representation of the text independent of any particular language. This representation functions as a neutral, universal central axis that is distinct from both the source language and the target language. In theory this method reduces the machine translation process to only two stages: analysis and generation. The analysis of the source text leads to a conceptual representation, the diverse components of which are matched by the generation module to their equivalents in the target language. The research on this strategy is related to artificial intelligence and the representation of knowledge. The systems based on the idea of a pivot language do not aim at direct translation, but rather reformulate the source text from the essential information. At the present time the transfer and pivot language strategies are generating the most research in the field of machine translation. With regard to the pivot language strategy, it is worth mentioning the Dutch DLT (Distributed Language Translation) project which ran from 1985 to 1990 and which used Esperanto as a pivot

language in the translation of 12 European languages.

It should be repeated that unless the systems function within a rigidly defined sphere, as is the case with TAUM-MÉTÉO, machine translation in no way offers a finished product. As Christian Boitet, director of GETA (Grenoble) says in an interview given to the journal *Le français dans le monde* Nº314 in which he summarizes the most important aspects of MT, it allows translators to concentrate on producing a high-quality target text. Perhaps then "machine translation" is not an appropriate term, since the machine only completes the first stage of the process. It would be more accurate to talk of a tool that aids the translation process, rather than an independent translation system.

The following is a relatively recent classification of some MT programs based on the results obtained from a series of tests that focused on errors and intelligibility in the target texts (Poudat, p.51):

Translator	**Address**	**Characteristics**	
Alphaworks®	http://www.alphaworks.ibm.com/aw.nsf/html/mt	Translates English into seven languages; transfer method	3
E-lingo®	http://www.elingo.com/text/index/html	Twenty pairs of languages available; transfer method	2
Reverso®	http://trans.voila.fr	Thirteen pairs of languages available; transfer method	1
Systran®	http://www.systransoft.com	Twelve pairs of languages available; direct transfer method	4
Transcend®	http://www.freetranslation.com/	Eight pairs of languages available; direct transfer method	5

Analysis of Some Errors in Machine-translated Texts

For the purpose of analyzing errors in machine-translated texts, it is revealing to compare such a translation with that done by a human translator. An article from Le Monde Diplomatique has been chosen, as this is a newspaper that is originally written in French but which is then translated into 17 other languages. In this case we will compare the French to English translations produced by Systran, Reverso and a human translator.

SOURCE TEXT: Le Monde Diplomatique, September 2002

Depuis le 11 septembre 2001, l'esprit guerrier qui souffle sur Washington semble avoir balayé ces scrupules. Désormais comme l'a dit le président George W. Bush, "qui n'est pas avec nous est avec les terroristes".

Systran		Reverso		Human translation
Since September 11, 2001, the warlike spirit which blows on Washington seems to have swept these scruples. From now on, like said it the president George W Bush, "which is not with us is with the terrorists". (37 words)		Since September 11, 2001, the warlike spirit which blows on Washington seems to have swept (annihilated) these scruples. Henceforth, as said it the president George W. Bush, "which (who) is not with us is with the terrorists". (35 +2 words)		Since 11 September 2001 the warmongering mood in Washington seems to have swept away such scruples. From that point, as President George Bush put it, "either you are with us or you are with the terrorists." (36 words)

The first point to be made is that MT is a translation method that focuses on the source language, while human translation aims at comprehension of the target language. Machine translations are therefore often inaccurate because they take the words from a dictionary and follow the situational limitations set by the program designer. Various types of errors can be seen in the

above translations.

- **Errors that change the meaning of the lexeme**

1. Words or phrases that are apparently correct but which do not translate the meaning in context:

Original: *l'esprit guerrier*

Systran: the warlike spirit

Reverso: the warlike spirit

HT: the warmongering mood

2. Words without meaning:

Original: *comme l'a dit le président George W. Bush*

Systran: like said it the president George W. Bush

Reverso: as said it the president George W. Bush

HT: as President George Bush put it

Although Reverso's translation is not completely correct, it translates *comme* into "as", which is the correct choice for this context.

- **Errors in usage**

The translation is understandable in that the MT produces the meaning but does not respect usage:

Original: *semble avoir balayé ces scrupules*

Systran: seems to have swept these scruples

Reverso: seems to have swept (annihilated) these scruples

HT: seems to have swept away such scruples

Original: *qui n'est pas avec nous est avec les terroristes*

Systran: which is not with us is with the terrorists

Reverso: which (who) is not with us is with the terrorists

HT: either you are with us or with the terrorists

As already mentioned, human translation concentrates on the target language, preferring to depart from the source language, if necessary, in order to reproduce meaning. For example, the human translator clearly chose "the warmongering mood in Washington" as a better contextual translation of *l'esprit guerrier qui souffle sur Washington* than the more literal versions seen in the machine translations.

Because MT aims primarily at comprehension and not at the production of a perfect target text, it is important to follow two basic rules in order to make the best use of programs. First, we need to recognize that certain types of texts, such as poetry, for example, are not suitable for MT. Second, it is essential to correct the source text, as even one letter can radically change meaning, as in the following example: *We shook hand* translates into "Nous avons secoué la main"; but *We shook hand***s** becomes "Nous nous sommes serré**s** la main". The omission of an **s** in the source text is enough to make the machine translation incomprehensible. It is of additional interest to note that the final *s* of *serrés* is a mistake because the MT program does not take into account the subtleties of French grammar with regard to the agreement of the past participle.

Computer-assisted Translation

In practice, computer-assisted translation is a complex process involving specific tools and technology adaptable to the needs of the translator, who is involved in the whole process and not just in the editing stage. The computer becomes a workstation where the translator has access to a variety of texts, tools and programs: for example, monolingual and bilingual dictionaries, parallel texts, translated texts in a variety of source and target languages, and terminology databases. Each translator can create a personal work environment and transform it according to the needs of the specific task. Thus computer-assisted translation gives the translator on-the-spot flexibility and freedom of movement, together with immediate access to an astonishing range of up-to-date information. The result is an enormous saving of time.

The following are the most important computer tools in the translator's workplace, from the most elementary to the most complex:

Electronic Dictionaries, Glossaries and Terminology Databases

Consulting electronic or digital dictionaries on the computer does not at first appear radically different from using paper dictionaries. However, the advantages soon become clear. It takes far

less time to type in a word on the computer and receive an answer than to look through a paper dictionary; there is immediate access to related data through links; and it is possible to use several dictionaries simultaneously by working with multiple documents.

Electronic dictionaries are available in several forms: as software that can be installed in the computer; as CD-ROMs and, most importantly, through the Internet. The search engine Google, for example, gives us access to a huge variety of monolingual and bilingual dictionaries in many languages, although it is sometimes necessary to become on-line subscribers, as with the Oxford English Dictionary. On-line dictionaries organize material for us from their corpus because they are not simply a collection of words in isolation. For example, we can ask for all words related to one key word, or for all words that come from a particular language. That is to say, they allow immediate cross-access to information.

For help with specific terminology there is a wide range of dictionaries, glossaries and databases on the Internet. Le Nouveau Grand Dictionnaire Terminologique developed in Quebec, Canada contains 3 million terms in French and English belonging to 200 fields. Another important resource is EURODICAUTOM, a multilingual terminology database created by the European Union in 1973 that covers a variety of specialized areas, both scientific and non-scientific (the list begins: Agriculture, Arts, Automation...). In addition, there are web sites that offer information on terminology that is helpful to translators. One such site is that of the TERMISTI research center attached to the Higher Institute for Translators and Interpreters (ISTI) in Brussels (http://www.termisti.refer.org) which provides information on the following:

- Dictionaries available on Internet such as those mentioned.
- Terminology networks such as RIFAL (Réseau international francophone d'aménagement linguistique), RITERM (Red Iberoamericana de Terminología)
- European terminology projects such as Human Language Technologies, Information Society Technologies.
- Translation Schools
- Forums and diffusion/discussion lists
- Conferences
- Journals such as the International Journal of Lexicography, La banque des mots, L'actualité terminologique, Méta, Terminogramme, Terminologies nouvelles, Terminology, Terminometro, Translation Journal, Apuntes.

Concordances

Computer concordances do not replace tools such as dictionaries and glossaries, but provide an additional method of handling texts for translation. They are word-processing programs that produce a list of all the occurrences of a string of letters within a defined corpus with the objective of establishing patterns that are otherwise not clear. These letters may form part of a word, such as a prefix or suffix for example, or a complete word, or a group of words. Specific functions of concordances include giving statistical data about the number of words or propositions, classifying words etc. in terms of frequency or alphabetical order and, most importantly perhaps, identifying the exact context in which the words occur. Information can be accumulated and stored as more texts are translated, producing a database available for consultation at any time in a non-sequential way.

Concordances are particularly valuable for translating specialized texts with fixed vocabulary and expressions that have a clearly defined meaning. They ensure terminological consistency, providing the translator with more control over the text, irrespective of length and complexity. However, they are not so helpful to literary translators, who are constantly faced with problems relating to the polysemic and metaphorical use of language. Nevertheless, some literary translators use concordances as they clearly have a potential role in all kinds of translation.

On-line Bilingual Texts

A bilingual corpus normally consists of a source text plus its translation, previously carried out by human translators. This type of document, which is stored electronically, is called a bi-text. It facilitates later translations by supplying ready solutions to fixed expressions, thus automating part of the process. The growth of the translation market has led to increased interest on the part of companies and international organizations in collections of texts or corpora in different languages stored systematically on-line and available for immediate consultation.

Translation Memories

Translation memories represent one of the most important applications of on-line bilingual texts, going back to the beginning of the 1980s with the pioneering TSS system of ALPS, later Alpnet. This was succeeded at the beginning of the 90s by programs such as Translator Manager, Translator's Workbench, Optimizer, Déjà Vu, Trados and Eurolang, among others. In its simplest form, a translation memory is a database in which a translator stores translations for future re-use, either in the same text or other texts. Basically the program records bilingual pairs: a source-language segment (usually a sentence) combined with a target-language segment. If an identical or similar source-language segment comes up later, the translation memory program will find the previously-translated segment and automatically suggest it for the new translation. The translator is free to accept it without change, or edit it to fit the current context, or reject it altogether. Most programs find not only perfect matches but also partially-matching segments. This computer-assisted translation tool is most useful with texts possessing the following

characteristics:

- **Terminological homogeneity**: The meaning of terms does not vary.
- **Phraseological homogeneity**: Ideas or actions are expressed or described with the same words
- **Short, simple sentences**: These increase the probability of repetition and reduce ambiguity.

A translation memory can be used in two ways:

1. In **interactive** mode: The text to be translated is on the computer screen and the translator selects the segments one by one to translate them. After each selection the program searches its memory for identical or similar segments and produces possible translations in a separate window. The translator accepts, modifies or rejects the suggestions.

2. In **automatic** mode: The program automatically processes the whole source-language text and inserts into the target-language text the translations it finds in the memory. This is a more useful mode if there is a lot of repetition because it avoids treating each segment in a separate operation.

A translation memory program is normally made up of the following elements:

a. A translation editor, which protects the target text format.

b. A text segment localizer.

c. A terminological tool for dictionary management.

d. An automatic system of analysis for new texts.

e. A statistical tool that indicates the number of words translated and to be translated, the language, etc.

Thus translation memory programs are based on the accumulation and storing of knowledge that is recycled according to need, automating the use of terminology and access to dictionaries. When translation tasks are repeated, memories save the translator valuable time and even physical effort: for example, keyboard use can be reduced by as much as 70% with some texts. Memories also simplify project management and team translation by ensuring consistency. However, translation memories can only deal with a text simplistically in terms of linguistic segments; they cannot, unlike the human translator, have a vision of the text as a whole with regard to ideas and concepts or overall message. A human translator may choose to rearrange or redistribute the information in the source text because the target language and culture demand a different content relationship to create coherence or facilitate comprehension. Another disadvantage of memories is that training time is essential for efficient use and even then it takes time to build up an extensive database i.e. they are not immediate time-savers straight out of the box. Finally, it should be stressed that translation memory programs are designed to increase the quality and efficiency of the translation process, particularly with regard to specialized texts with non-figurative language and fixed grammatical constructions, but they are not designed to replace the human translator.

Conclusion: The Impact of the New Technologies on Translators

It has long been a subject of discussion whether machine translation and computer-assisted translation could convert translators into mere editors, making them less important than the computer programs. The fear of this happening has led to a certain rejection of the new technologies on the part of translators, not only because of a possible loss of work and professional prestige, but also because of concern about a decline in the quality of production. Some translators totally reject machine translation because they associate it with the point of view that translation is merely one more marketable product based on a calculation of investment versus profits. They define translation as an art that possesses its own aesthetic criteria that have nothing to do with profit and loss, but are rather related to creativity and the power of the imagination. This applies mostly, however, to specific kinds of translation, such as that of literary texts, where polysemy, connotation and style play a crucial role. It is clear that computers could not even begin to replace human translators with such texts. Even with other kinds of texts, our analysis of the roles and capabilities of both MT and CAT shows that neither is efficient and accurate enough to eliminate the necessity for human translators. In fact, so-called machine translation would be more accurately described as computer-assisted translation too. Translators should recognize and learn to exploit the potential of the new technologies to help them to be more rigorous, consistent and productive without feeling threatened.

Some people ask if the new technologies have created a new profession. It could be claimed that the resources available to the translator through information technology imply a change in the relationship between the translator and the text, that is to say, a new way of translating, but this does not mean that the result is a new profession. However, there is clearly the development of new capabilities, which leads us to point out a number of essential aspects of the current situation. Translating with the help of the computer is definitely not the same as working exclusively on paper and with paper products such as conventional dictionaries, because computer tools provide us with a relationship to the text which is much more flexible than a purely lineal reading. Furthermore, the Internet with its universal access to information and instant communication between users has created a physical and geographical freedom for translators that was inconceivable in the past. We share the conviction that translation has not

become a new profession, but the changes are here to stay and will continue to evolve. Translators need to accept the new technologies and learn how to use them to their maximum potential as a means to increased productivity and quality improvement.

References

ARNOLD, Doug, BALKAN, Lorna et al. *Machine Translation: An Introductory Guide.* URL: http://clwww.essex.ac.uk. MTbook/HTML/ (consulted 13 May 2002).

BEDARD, Claude. *Ce qu'il faut savoir sur les mémoires de traduction.* URL: http://www.terminotix.com/Terminotix-fr/info/art_5html (consulted 17 June 2002).

BERTRAND-GASTALDY, Suzanne. *La traduction automatique ou assistée par ordinateur*. URL: http://www.esi.umontreal.ca/~moalv/blt6134/index.html (consulted 6 June 2002).

CHAMPOLION, Yves. *Machine Translation (MT and Future of the Translation Industry*. URL: http://accurapid.com/journal/15mt.htm (consulted 2 May 2002).

DELISLE, Cynthia. *La traduction automatique: espoirs ou déboires?* URL: http://www.cafi.org/colloques/compte-rendu-mard-2000.html (consulted 14 June 2002).

DISTRIBUTED LANGUAGE TRANSLATION (DLT) http://www.fb10.uni-bremen.de/linguistik/khwagner/mt/ppt/DLT.ppt (consulted 19 Oct. 2003).

HUTCHINS, W. John. *Translation. Technology and the Translator.* URL: http://www.bcs.org.uk/siggroup/nalatran/nalamtr7/nalamt76.htm (consulted 9 May 2002).

HUTCHINS, W. John. *Computer-based Translation Systems and Tools.* URL: http://www.eamt.org/archive/hutchins_intro.html (consulted 8 July 2002).

HUTCHINS, W. John. *The State of Machine Translation in Europe and Future Prospects*. URL: http://www.hltcentral.org/page-917.0.shtml (consulted 10 March 2002).

KAY, Martin. *History of machine Translation*. URL: http://www.lsadc.org/Kay.html (consulted 9 May 2002).

LAB, Frédérique. *Linguistique contrastive et traduction automatique. Une étude de cas: les traductions anglaises du présent français*. Defended Université de Paris VII UFRL

LAURENS, Olivier. *La traduction vers le français des expressions résultatives anglaises*. Université Paris 7—Denis Diderot UFR en Linguistique, 1996.

LOFFLER-LAURIAN, Anne-Marie. *La traduction automatique*. Villeneuve d'Ascq: Presse Universitaire du Septentrion, 1996.

OREN, Tim. *Due Diligence* URL: http://www.pacificavc.com/blog/2004/01/27.html (consulted 11 March 2004).

POUDAT, Céline. "La traduction automatique en libre accès sur l'Internet". *Le français dans le monde*, 2001, n°. 314, p. 51-52.

RICO PEREZ, Celia. *New Trends in Machine Translation*, *Meta*, XLII, 4, 1997. URL: http://www.erudit.org/erudit/meta/v42n04 (consulted 9 May 2002).

RICO PEREZ, Celia. *From Novelty to Ubiquity: Computers and Translation at the Close of the Industrial Age.* URL: http://accurapid.com/journal/15mt2.htm. (consulted 9 May, 2002).

RICO PEREZ, Celia. *Universals of Language, Sublanguage and Artificial Languages: May the Machine Compete with the Translator?* URL: http://www.ccl.umist.ac.uk/events/conference/multimedia1.htm (consulted 18 May, 2002).

ROSEN, Mariam et TSAO, Claude. *Demain, la traduction automatique.* URL: http://www.mondediplomatique.fr/mav/Internet/ROSEN/traduction.html (consulted 19 July, 2002).

SAMUELSON-BROWN, G. *A Practical Guide for Translators* (2nd Ed.). London: Multilingual Matters Ltd, 1995.

SEVERINI, Alfiero. *Mise au point d'un système de traduction automatique italien français*. Université Paris 13, 2001.

Traduction—terminotique - Lexicographie.

URL: http://www.owil.org/traduc.htm (consulted 28 July, 2002).

TRUJILLO, A. Translation engines: techniques for machine translation. London: Springer, 1999. URL: http://www.ccl.umist.ac.uk/staff/iat/transeng (consulted 30 Sept.2003).

TRUJILLO, A. Estrategias de traducción automática.

URL: http://www.imim.es/quark/num19/019053.htm (consulted 13 Oct. 2003).

ZIMINA, Maria. *Alignement de textes bilingues par classification ascendante hiérarchique.* URL: http://www.cavi.univ-paris3.fr/lexicometrica/jadt/jadt2000/pdf/77/77.pdf (consulted 3 August

2002).