TECHNOLOGY ASSESSMENT NOTE / NOTE SUR L'EVALUATION DE LA TECHNOLOGIE

COMPUTER ASSISTED TRANSLATION:
PRELIMINARY MARKET AND INDUSTRIAL ASSESSMENT

Prepared By:

Bev Hillman

Technology Policy and Planning Branch

Department of Communications

June, 1986.

P
308
H56
1986

technology,
policy &
planning.

politique &
planification
de la technologie.

Queen
P
308
H56
1986

Industry Canada
Library Queen

JUIL JUL 24 1998

Industrie Canada
Bibliothèque Queen

2. COMPUTER ASSISTED TRANSLATION:

PRELIMINARY MARKET AND INDUSTRIAL ASSESSMENT

COMMUNICATIONS CANADA

NOV 18 1991

LIBRARY — BIBLIOTHÈQUE

Prepared By:

/. Bev Hillman

Technology Policy and Planning Branch

Department of Communications

June, 1986.

DD 10916861
DW 10945599

P
308
H56
1986

TABLE OF CONTENTS

PAGE

1- INTRODUCTION 1

2- TRANSLATION MARKET

2.1 MARKET SIZE ESTIMATES 2
2.2 LATENT DEMAND 3

3- MACHINE TRANSLATION MARKET

3.1 DESCRIPTIONS 3
3.2 SIZE ESTIMATES 5
3.3 MARKET NICHES 5

4- USER NEEDS 7

5- PRODUCTIVITY ISSUES 11

6- SYSTEM EVALUATIONS

6.1 THE WEIDNER SYSTEM 13
6.2 THE SYSTRAN SYSTEM 16
6.3 THE TAUM-AVIATION SYSTEM 18
6.4 THE METEO SYSTEM 18
6.5 CONCLUSIONS 19

7- MARKET TRENDS/STRATEGIES 19

8- MARKET SUPPLY

8.1 INDUSTRIAL PLAYERS 21
8.2 TRANSLATION AGENCIES 23

9- SUMMARY AND CONCLUSIONS

9.1 NEED FOR TRANSLATION 25
9.2 TRANSLATION MARKET 26
9.3 MACHINE TRANSLATION 26
9.4 OPPORTUNITY 27

10- DEFINITIONS 28

11- REFERENCES 29

1- INTRODUCTION

The language barrier impedes the free movement of knowledge between peoples. The limitation of linguistic skills and the high cost of translation are the basic causes of poor permeation of information from one language into another. Although language does not constitute an impermeable barrier, it is certainly an obstacle which is frequently difficult, and occasionally impossible, to surpass.

The need for language translation results from the necessity to communicate between speakers of various languages. Van Slype (Reference 11.6) noted that this need is on the increase because of the following factors:

- improvements in transport;
- improved telecommunications;
- concentration of economic activity around major multinational groups;
- the emergence of new countries in the world of literary, scientific and technical publications, and the increasing determination of governments in these countries to guarantee cultural development in their own languages;
- in bi- or multilingual countries and in the majority of international organizations, official acts are required to be published in various languages and meetings held in several languages;
- pressure from the international market is forcing a number of English-speaking businesses to translate their technical and commercial documents into the language of their clients. In the same way, scientists and experts are asking for more and more translations in the face of the growing mass of literature in languages with which they are unfamiliar;
- the acceleration of the decision-making process, more up-to-date information is needed on domestic and distant environments;
- the expansion of international meetings which have themselves led to a major expansion in the demand for translation: agendas, working documents, verbatim reports, minutes, etc.

Within this environment of growing interdependence of world economies and the desire for maintaining cultural independance, the need for language translation continues to grow. Faced with the upward spiraling cost of human translation and ever-increasing workloads, centralized translation agencies from Japan to Germany have been pursuing long-range development plans. In Canada, the pressure to find computer aids to translation is growing increasingly intense in a context where budgetary limitations make it necessary to find some way to increase the individual translator's productivity.

Research and industry support activities are currently underway within DOC to develop a world class computer assisted translation capability in Canada. The purpose of this report is to review and consolidate available information on the machine translation markets and industries.

2- THE TRANSLATION MARKET

2.1 MARKET SIZE ESTIMATES

The study "Better Translation For Better Communication" (Reference 11.6) aims to determine for the most important languages and sectors:

- the nature of the translation market, in both qualitative and quantitative terms.
- trends in translation demand
- the share of the market which revised and unrevised machine translations will have.

It is clear that the volume of the professional translation market is very difficult to establish. The study commissioned by the Commission of the European Communities, not only analyses the previously published data, but also draws conclusions from a questionnaire-based survey and from personal contacts in 1979-1980. The statistical data must, however be treated with caution and regarded rather as orders of magnitude.

The world written translation market is believed to be within a range of from 80 to 240 million pages of 250 words, i.e. 20 to 60,000 million translated words per annum. Taking an average price of 8 European Units of Account (ECU) per 100 words, one arrives at an annual world turnover of between 1.6 and 4.8 billion ECU (in September 1981, 1 ECU is worth roughly 1 US $). The world written translation market is expanding by 9 to 10% per annum.

In Canada, three quarters of a billion words per year seems to be a conservative estimate for the whole of the current Canadian market. In 1981-1982, the workload for the Federal Translation Bureau was 276 million words and the demand is still increasing. Given that the average cost of translation in Canada is about 20 cents per word, the Canadian market represents no less than 150 million dollars annually in translation costs and this is almost certainly a conservative estimate.(Reference 11.2)

The number of professional translators working full time, or expressed in full-time equivalents, is believed to be between 57,000 and 170,000, to which must be added a little more than half that number of secretaries and typists. This yields a total staff of between 90,000 and 260,000 persons involved in producing written translations throughout the world. In Canada it is estimated that there are some 2500 professional translators and perhaps an equal number of part-time or occasional translators. This estimate is based on the current employment of some 1200 translators by the federal government and estimates that a roughly equal number are employed elsewhere.

2.2 LATENT DEMAND

In addition to the translation market actually served, there exists a latent demand, currently unsatisfied for various reasons (in particular: availability of translators, cost, time). It is important to try to evaluate this market, now that a new technology (machine translation) may make it possible to meet this demand, at least partially.

The answers obtained in the twelve country survey (Reference 11.6) suggest that the unsatisfied needs would account for 30% of the number of actually translated pages.

Since the cost of and the time required for human translation have been increasing rather than lessening, it is unlikely that the latent demand will be satisfied by human translation. Only through lower costs and/or shorter turnaround times (a possibility with machine translation) could the latent demand be satisfied.

3- MACHINE TRANSLATION MARKET

3.1 DESCRIPTIONS

The machine translation and machine aided translation market may comprise different kinds of product: (Reference 11.10)

i) Rough machine translation without post-editing

Such translation may be most valuable to people wishing to get a rapid acquaintance with a journal article, a conference paper, or a working group paper, written in a language they do not understand, in the (numerous) cases where they cannot afford the cost and/or the time delay for a normal human translation process. This potential market could represent 30% of the actual market for formal, written, human translation.

The conditions for the realisation of this market are:

- cost and time delay significantly reduced (e.g. one order of magnitude lower than human translation)
- average text intelligibility of MT above 75% (versus 98 to 99% for source text)
- easy access to a MT system.

ii) Machine-Aided Translation

This product with post-editing will reach the same quality as human translation. It permits the carrying out of the same volume of work with a reduced staff of translators, converted into post-editors (or a higher volume of translation with the same staff), at about the same cost as or a little less than human translation; but with considerable reduction in time delay.

Post-edited MT is likely to interest private translation offices and freelance translators, for delivery date reasons and also because, in the medium-term, it is likely to be improved and to cost relatively less than human translation.

The conditions for machine-aided translation development are:

- acceptability to the human translator, which varies from translator to translator,
- use of a word processing subsystem,
- easy access to a MT system.

iii) Human-aided Machine Translation

This approach drops the requirement that MT be fully automatic. While the machine has the initiative in the translation process, the human will be requested to assist the machine with problems that are beyond its capabilities. Human intervention can take place either before (pre-editing), during, or after (post-editing) machine processing.

iv) Translation of controlled text

This type of system retains automation at the expense of specializing the system to certain well-defined types of texts. There are two ways of restricting the inputs:

- by artificially constraining what the author is allowed to write, may be used by organizations producing their own texts, and thus able to impose from the start, writing rules that would allow the post-editing efforts to be minimized (e.g. instruction manuals, maintenance manuals)
- by designing translation systems specifically for some highly restricted natural domain. (eg. certain technical manuals, medical reports, financial market reports, etc. which use only a part of the whole language's syntax and vocabulary)

The conditions for the development are:

- control of the source texts,
- access to a MT system,
- design of the controlled language subsystem and of the MT subsystem such that the correction ratio is brought to a low level (i.e. 5 to 10%), permitting a reduction in time delay and in costs.

3.2 SIZE ESTIMATES

Machine translation is currently about a $10 million annual business in the U.S. and growing at about 30% per year, according to Harvey P. Newquist III, editor of the AI Trends Newsletter. He believes the growth could be greater if the technology were better understood by U.S. companies. (Reference 11.9)

The conclusion in Reference 11.6 was that: it does not appear unrealistic to expect that, in the medium-term (5 years), MT could occupy a share of the market representing 5 to 15% of that of human translation, taking into account the increase in the general translation market (50% in 5 years).

	1979 Estimated		1984 Estimated (MT excl)		1984 Estimated (MT incl)	
	Quantity (in million 250-word pages)	Value (million US $)*	Quantity	Value	Quantity	Value
Actual market						
--Human Translation	80-240	1600-4800	120-360	2400-7200	117-333	2340-6660
--Post-edited MT	0	0	0	0	3-27	60-540
-- Raw MT	0	0	0	0	3-27	9-84
TOTAL	80-240	1600-4800	120-360	2400-7200	123-387	2409-7284
Latent Market	24-72	token	36-108	token	33-81	token

* 1 U.S. $ was worth roughly 1 European Unit of Account in 1981.

3.3 MARKET NICHES

Are there particular subject fields better suited than others to computer-assisted translation?

a) In 1983, Dr. Kittredge conducted a study to identify texts whose linguistic regularities make possible the automation of some important part of the translation process. (Reference 11.3) It would appear highly desirable to automate in cases where translations must be done quickly, or involve a high annual volume, involve a large amount of terminological lookup, are repetitious or tiring for human translators, etc. However, experience has shown that what is desirable in machine translation is not always possible. It is usually the linguistic properties of the texts which will make or break the success of any attempt to automate translation.

The study examined ten different types of text that were being translated manually at the Translation Bureau. Text varieties which were either low in complexity, or high in homogeneity, or high in annual volume were put in a set of possible candidates for automation. Linguistic investigations were carried out on the vocabulary, grammar and style of each text variety, leading to an initial ordering of the text- types with respect to their complexity for automatic analysis and translation. Adding such practical considerations as the volume of texts to be translated and the requirements of output quality led to the identification of three areas in which automation could be recommended. The areas identified were:

- agricultural market reports,

- weather synopses,

- competition notices and other staffing documents.

Other text varieties studied were found to be too complex or heterogeneous to warrant automation of their translations. This does not preclude the use of computer aids such as terminology banks, text editors.

In the course of this study, no text area was found which presented the same fortunate combination of high annual volume and low complexity as do weather bulletins. The most promising current candidates for automation are either significantly more complex, or combine low complexity with low annual volume in the present circumstances.

b) At the level of demand the emphasis is moving from the traditional sectors such as literature, law, administration, commerce, to new fields and types of text like scientific and technical review articles, patents, standards, minutes of meetings, congress and visit reports, accident reports, installation and maintenance handbooks for large industrial equipment,...Unlike the traditional consumers of translations, who were not very concerned about economic profitability, the new requesters are considerably more sensitive to such factors as cost, quality and deadlines.

The survey (349 respondents in twelve countries) in the study "Better Translation For Better Communication" (Reference 11.6) shows the following distribution of translation volume among subject areas:

- commercial	35.4%
- industrial	21.0%
- scientific	20.0%
- legislative	9.3%
- press,current activities	3.5%
- audio-visual	2.1%
- educational	1.5%
- literary	0.3%
- miscellaneous	6.9%
Total	100.0%

c) Studies show that in each discipline, 2 or 3 languages provide approximately 75% of the world literature: for example, in geography, 80% to 90% of the publications are in English or Russian. In mathematics, the principal languages are English, Russian and French. In information science, more than 75% of the publications are in English, German or Russian. As regards scientific and technical publications, only five languages play any significant role: English, with between 40% and 80% of articles according to the subject matter with the remaining divided between Russian, German, French, and Japanese.

A recent study (Jagodzinski and Latour- 1980) confirms that:

- in the "advanced" modern disciplines, such as genetics and biochemistry, authors from non-English-speaking countries publish primarily in English
- while in more traditional disciplines, such as medicine and biomedical engineering, authors publish more readily in their mother tongue.(Reference 11.6)

d) Is there a certain pattern in the decisions to translate scientific publications?

On the basis of a sample of 133 translations of foreign scientific and technical translations into English, and of 133 foreign articles which were not translated, dating from 1970, Anderson (1978) made a study of the formal characteristics of these documents, and attempted to determine the likelihood of their being translated. He found that the translated articles differed significantly from the untranslated articles as regards: the subject matter, the original language, and the number of references. If these factors were found to be constant in time it would be possible to calculate the likelihood of translation. In this way it would be possible to choose those articles most likely to require translation as soon as they appear in their original languages, and to send them for translation immediately. The average time lapse between the publication of an article in a foreign language and its translation at the request of the first requester (at present 1.9 years) for translations into English could in Anderson's view thus be considerably reduced.(Reference 11.6)

4- USER NEEDS

4.1 A study entitled "Functional Specifications For A Translator's Work Station" (Reference 11.4) was recently completed by Socioscope Inc. This study prepared a detailed analysis of the translation tasks as seen from the perspective of the individual translator. The results of the detailed analysis of translation tasks were presented to a panel of technical experts (including specialists in translation, software, hardware, and telecommunications). The various areas where computerization could be introduced or extended to facilitate the work of the individual translator were then determined by this panel. The objective is to provide a workstation which is MAHT (Machine Assisted Human Translation), keeping in mind the future possibility of some further development of HAMT (Human Assisted Machine Translation), and with the even more remote possibility of MT (Machine Translation).

4.2 The functional areas identified below are areas which the panel of technical experts identified as potentially benefitting from computerized support.

i) Access to databases

- access a terminology bank employing a key word search system with Boolean operators (up to 30% of the translator's time is spent looking up terms)
- automatically consult a dictionary or thesaursus (or a terminology bank) about a word with a cursor in front of it
- maintain a personal "electronic cardex" to file words, terms, and expressions the translator has found useful or has annotated

ii) Text Manipulation and Text Editing

- automatically receive and view the source document and the target document on a split screen display
- scroll rapidly through source and target documents independently
- employ independent cursors to indicate the translator's progress within the source and target documents
- highlight terms, expressions, and sentences employing a cursor and a dedicated key or icon (i.e. electronic equivalent of a yellow pencil)
- erase words, phrases, and paragraphs, and recall these erased areas on command
- identify spelling and minor grammatical errors automatically, without interrupting word entry

iii) Administrative Tasks

- automatically count the number of words within the source and target documents
- display the contents of the electronic "in basket" in a tabular format indicating: the document's time of arrival, the time when delivery is expected, the word count of the document/s, the name of the client, and the priority of the document

- display, in tabular format, the descriptions listed above for all documents translated in the previous three months (for administrative purposes and to integrate previous translations into current ones)

iv) Keyboard Simplification and Special Characters

- employ a mouse and a menu of icons as an alternative to dedicated keys, in order to simplify the keyboard
- employ a virtual keyboard displayed on the screen and controlled with a mouse, so as to be able to develop and manipulate special type fonts (i.e. mathematical symbols, accents, scientific notations)
- enter, receive, store and transmit images (graphics, diagrams, pictures) contained in documents received for translation

v) Electronic Transmission and Storage of Text

- receive (or transmit) the source document, the target document, or both, from (to) other stations such as those belonging to a client, to other translators within the office, to the revisor, the translation system manager, or to an external translator.

  There is a clear need for a means for integrating translators into broadly based electronic networks.

- transmit stored source and target documents to a local or remote printer

  In order to transfer documents between computers of different vendors, standards will evolve. Many new standards are being developed at the international level as part of the escalating drive towards Open Systems Interconnection (OSI). OSI standards are built in a layered fashion. Protocols for document transfer are at the highest layer, called the application layer. The most important standards for document transfer will be:

- Office Document Architecture/Office Document Interchange Format (ODA/ODIF)
- Teletex
- Message Handling Systems
- NAPLPS (Videotex), to a lesser extent

vi) Voice

- transmit a translation recorded on dictaphone to a remote location in digital form so as to permit its storage and subsequent typing. It is worthy of note that the Message Handling System services are intended to incorporate voice content within a message and will provide an appropriate, standard protocol.

vii) Draft Machine Assisted Translation

- evoke a machine assisted translation routine to provide a rough and ready translation aid for the professional translator (for example in certain specialized translation areas or as a rough first draft)

viii) Shared Visual Space

- employ a section of the screen as an electronic blackboard so as to enable two or more translators to cooperate on the translation of the text

ix) Security

- encrypt the translated transmission, should this be required

Other features which must be considered as well are the physical environment, ergonomic features and psychological factors. Many current systems have serious human factor shortcomings and close attention must be paid to the real requirements of the end user. Many of the criteria already exist in the human factors literature with respect to the principles, guidelines, rules and experimental data for designing ergonomically appropriate computer work stations. The human factors literature also contains criteria for designing "user friendly" software.

On the other hand, only a very detailed and specifically focussed task analysis (as part of the development of technical specifications for the device), examining the tasks to be performed, their structure, their relation to each other, and the working modes of the translators, will allow designers to define those requirements that are specific to the translation process.

A comprehensive description of the translator population in terms of age, training, computer knowledge, etc., is necessary to define areas such as training requirements, user documentation.

4.3 Given the current state of NLP technology, many claim that the best way to help solve translation problems is not to develop MT systems but rather to provide the translator with tools that would help him do his work more efficiently, while leaving the basic initiative in the translation process to the translator. A specialized translator's workstation would integrate in a coherent environment several tools, including: a powerful split-screen text processor that can align source language and target language texts, produce concordances, access on-line dictionaries and multilingualthesauri, phrasal lexicons, and terminological data banks. In addition, they might provide the translator with text critiquing facilities, as well as rapid on-line access to public databases of previously translated materials. Very few of these features are now available on commercial systems. Nonetheless a few American companies (eg. Weidner, Alps) are currently introducing workstation products with a few of these features as part of their machine translation systems. These products run on relatively cheap microcomputers and might soon be affordable for individual translators.(Reference 11.2)

5- PRODUCTIVITY ISSUES

5.1 In the search for productivity gains, a number of users have been induced to experiment with commercial systems which, though based on a relatively weak linguistic approach, promise gains on the level of dictionary lookup, text editing, composition for printing, document data bank maintenance, etc. The economics claimed for such systems are apparently due mostly or entirely to these non-translation functions, and to the fact that the systems are used for texts which are frequently updated, where it is quite economical to merely edit the already stored translation of an earlier version. When it comes to translation per se such systems depend on a heavy amount of post-editing.

A first gain in productivity of about 10% will be obtainable by recourse to multilingual terminology data banks accessible on-line, text processing software, internal reference data banks representing the collective memory of a group of translators, in a word by organizing the intellectual work of the translator around a number of handy data-processing tools which can help him to work faster and to ensure a more constant quality.(Reference 11.6)

5.2 The next addition to the commercial scene may be a second-generation system called METAL, very similar to AVIATION, being developed at the University of Texas under the sponsorship of the German electronics giant, Siemens, A.G. Large scale in-house tests have recently indicated that the times and costs for machine translation plus post-editing can compare favourably with the human sequence, at least for the German-English technical documentation tested. (Reference 11.3)

With further reaching computerization and the introduction of computer-assisted and automatic translation systems an additional gain in productivity of about 50% is foreseeable.(Reference 11.6)

5.3 In sum, three areas of development are bound to increase translators' productivity. The first consists in the marketing of rough machine translations, in the form in which they leave the computer, without any human intervention. The rough translation will be returned to the applicant in less than 24 hours for a total cost of about 25% of a human translation.Since this market does not exist at the present time, it is impossible to quantify it.

Post-edited computer translation, interactive or otherwise, will develop at a cost currently comparable with or slightly higher than that of human translation, but with delivery times which are twice as short. This clear improvement in the factor "time" seems likely to encourage the emergence of a part of the latent market, even at constant cost. A realistic - and cautious - forecast is that within 5 years post-edited machine translation will occupy 5 to 15% of the market, and will then develop much more quickly.

Lastly, a large market will continue to be covered by translators who, while not having access to the large automated systems, will gradually equip themselves with smaller-scale data-processing aids which will increase productivity by permitting faster production of high-quality translations in highly specialized fields.

5.4 The assessment of the quality of a translation, like that of the original text, is a highly subjective matter. It is based on a number of criteria, the nature and relative weight of which vary according to the point of view adopted. The following breakdown may be made:

as regards the criteria of quality:

- intrinsic qualities, (in theory) independent of the reader:
- terminological accuracy,
- grammatical accuracy,
- orthograghical accuracy.
- extrinsic qualities, dependent on the "text-reader":
- intelligibility,
- fidelity to the meaning, aims, and nuances of the original text.

as regards the point of view:

- the point of view of the user, for example the lawyer will be particularly concerned with fidelity
- the point of view of the head of the translation department, who will stress the intrinsic qualities.

In general, the assessment of a translation is based on criteria of quality, but also on criteria of an economic nature: mainly the cost, duration and turnaround time of the translation.

5.5 The survey carried out with translators from twelve countries (Reference 11.6) also related to the criteria for assessing translations in various fields. The list of these criteria is given below, together with an indication of their relative weights in the 400 answers examined, and an indication of the fields for which the weighting of the criterion is especially important.

Criteria	Relative weighting	Areas where criteria are significant (weight)
Intelligibility	34	education(40), industrial(39), scientific(39)
Fidelity	32	legal and administrative(47), scientific(42)
Speed	12	news(32), commercial(16)
Style	11	literary(40), audio-visual(25), education(18)
Presentation	6	audio-visual(20), education(11)
Cost	5	industrial(7), audio-visual(7)

6- SYSTEM EVALUATIONS

6.1 THE WEIDNER SYSTEM

An evaluation of the Weidner computer-assisted translation system was conducted by the Bureau of Management Consulting, DSS for the Department of the Secretary of State. (Reference 11.5)

The purpose of the trial was to determine:

1) whether the Weidner Microcat system improved translators' productivity, that is, whether it reduces translation time for specific texts;

2) the relative cost of computer-assisted translation as compared with human translation; and

3) the impact of computer-assisted translation on translators and their duties.

The findings of the trial with respect to these objectives indicated:

1) The evaluation methodology consisted in comparing parallel translations of the same texts, first produced by means of human translation and then, usually a week later, by means of computer assisted translation. Human translation was performed at an average rate of 10.77 words per minute and computer-assisted translation at a rate varying between 7.54 and 7.09 words per minute, depending on which hypothesis was used for the calculations. This comparison was made for 24 texts totalling 15,760 words. This indicates that human translation using a word processor is between 30% and 34% faster than computer-assisted translation of the same texts, at least within the context of this trial.

The texts translated with the assistance of the Weidner system and revised by the translators were of deliverable quality.

2) A cost analysis was conducted to define the break-even point computer-assisted translation must reach to be an efficient tool for Translation Bureau purposes. The break-even point is defined by determining the number of additional words the system must help translate, over and above the number of words produced using human translation, for the Bureau to recover its costs.

For a system with one workstation for one translator, the number of additional words to be translated over three years is calculated as follows:

$$\text{System cost per day (SCD)} = \frac{\text{cost of workstation}}{\text{\# of days/year * \# of years}} = \frac{39{,}200}{\text{235 days * 3 years}} = \$55.6$$

Equivalent in words translated (EWT) = SCD / Cost per word = \$55.6 /\$0.28
= 199 words

The cost per word of \$0.28 refers to the current approximate cost established by the Translation Bureau and includes all administrative costs.

The percentage of efficiency the system must reach for the Bureau to recover its costs can be estimated. If it is assumed that the average translator in the Bureau produces 1,250 words per day by means of human translation the Weidner system's break-even point, calculated over three years, would be 1,449 words per day. (i.e. 1,250 + 199) If expressed as a percentage, the computer-assisted translation workstation will break even when it increases productivity by 15.9%. (i.e. 199 additional words / 1,250 words * 100)

Since texts translated using the Microcat system must be keyed in, 14.4% must be added to this percentage for the time spent by an expert WP operator entering texts, which is equivalent to 7.7% of a translator's salary. Therefore, for computer-assisted translation to be cost effective, its production level must be 23.6% higher than that of human translation.

A shared-time installation with four workstations requires one IBM PC-XT (microcomputer) on which the Weidner program runs, and ordinary PCs at the other three workstations for wordprocessing and revision. In a four workstation configuration over three years of operation, the average cost for one workstation is $16,675, the number of additional words per workstation per day to break-even is 84, and the percentage of production increase that computer assisted translation must reach for the Bureau to recover its costs is 14.4%.

3) Analysis of the data suggested that the work of translators will be radically changed when computer assisted translation is perfected. The translation operation as it now exists will remain, but there will also be positions for full-time revisors of computer-assisted translation output.

The participating translators were asked which were the system's main advantages and disadvantages. On the positive side was the vocabulary function. The Weidner system comes with a core dictionary of general vocabulary, in addition to which users may create their own subdictionaries. When the system is operated, the user sets the lookup sequence in which the translation program is to consult the various subdictionaries. On the negative side, the translators felt that the system is not fully perfected and did not produce texts of satisfactory quality. However, these were the reactions of the translators and not a scientific evaluation of the system.

Methodology

Six translators participated in the trial, two each from National Defence, Agriculture and Statistics Canada. The trial was conducted over four months, from May to the beginning of September 1984. The trial attempted to provide translators with optimun conditions: texts selected from restricted subject fields, adequate technical support, prior terminology research and compilation of dictionaries, and a quiet workplace.

The translators participated in the trial on a part-time basis only. For most translators the learning time to thoroughly master the Weidner system was until mid-July, i.e. approximately 60% of total trial time. It should be noted that TAO Canada, the company that distributes the Weidner system in Canada, states that translators need four months' experience with the system to reach their maximun production level.

It should be noted that the study discusses a number of evaluation models applicable to this type of trial in Appendix C. The models were:

1. Comparison of daily computer-assisted translation production with data on translators' production in the sections,
2. Timed comparison of human translation in the sections and computer-assisted translation during the trial,
3. Comparison of production times of computer-assisted and human translations of different but comparable texts (in terms of content, level of difficulty, length, format and other factors)
4. Comparison of production times of computer-assisted and human translations of the same texts.
5. Comparison of several translators' production using computer-assisted translation in the same subject field. Plus monitor translators' productivity in subject fields other than their specialities.

The study findings require continual updating to remain valid due to the fact that the relative costs of hardware change considerably over time.

The significance of the trial's conclusions is limited by the small number of participants (six translators) and by the experimental context. The conclusions are therefore valid only in relation to this trial.

The Weidner system is marketed in the form of a turnkey system. For a standard language pair, the 100 system, comprising a Digital Equipment minicomputer, 4 screen-keyboard terminals, a printer, text-processing software, the translation software and a current dictionary is sold at US $240,000 (i.e. 180,000 European Units of Account). The 200 system, with 20 terminals, is sold at US $400,000. The cost of maintenance, i.e. of the supply of an improved version of the translation software, every six months, is 1% of the purchase cost.

An assessment carried out by I. Pigott on the Weidner Spanish-English system on a limited sample of texts showed that the post-editor time spent was a third of that taken by a human translator to translate the same text.

6.2 THE SYSTRAN SYSTEM

The Systran system used by the Commission of the European Communities has undergone several quality evaluations. (Reference 11.6) The results can be formulated as follows:

- unedited machine translation is acceptable by the end-user, under certain circumstances, when the intelligibility exceeds a threshold of from 70 to 75% (this is achieved in the case of English-French and of English-Italian)
- raw machine translation can, after post-editing, be less expensive in terms of translating manpower, than the chain: human translation+ revision, for a finished product of equivalent quality. However, for this to be the case the post-editing rate must be below a threshold from 30 to 40% (this is already the case for English-French)
- the performance of Systran is dependent, to a large degree, on the size of its vocabulary: 40 to 50 thousand words and expressions would appear to constitute a minimum threshold
- GM Canada and Xerox, both users of Systran, have increased their productivity 3-4 times and 2-10 times, respectively. This gain in productivity is closely linked to the fact that Systran is used in restricted linguistic domains and with the support of very large dictionaries (over 100,000 words). (Reference 11.4)

A cost price study of Systran was carried out by Chaumier and Van Slype in 1977 at the request of the CEC. Following are the results for the direct cost of translation (personnel + equipment):

Activity	Human translation		Machine translation	
	in official institution	in private bureau	in official institution	in private bureau
	(Cost in Units of Account* per 100 words, 1976 values)			
Total without revision or post-editing	11.375	5	3	2
Total inclusive of revision or post-editing	16	5	11.375	4.5

---

* 1 Unit of Account was worth roughly 1 U.S. $ in 1981.

The capital expenditure necessary for the development of the software and creation of the dictionaries, by language pair and subject area, was evaluated at 350,000 Units of Account. On these bases the profitability threshold (i.e. the number of words which must be translated in order to cover the investment cost and to obtain a consequently lower operating cost for MT rather than for HT) can be calculated on several assumptions:

- with or without "dedicated" acquisition of the source text
- with or without revision or post-editing
- by official organization or private offices.

The threshold ranges from 3.3 million words (the most favorable case: official service, no correction, text available on machine-readable medium) to 70 million words (the most unfavorable case: private office, with correction, source text to be inputted).

The Systran Institute (1980) of Munich offers the Systran software under licence and presents the following calculations:

- total fixed costs (licensing costs and amortization of startup expenses): 92,500 UA
- cost of computer + licensing fee according to volume (not including data acquisition): 0.2 UA/100 words
- post-editing cost (at a rate of 5,400 words/day for a translator working 200 working days per annum and costing, in salary and overheads 40,000 UA/year): 3.7 UA/100 words
- total cost of MT (not including acquisition): 3.9 UA/100 words
- cost of HT (at a rate of 1,800 words/day for a translator working 200 working days per annum and costing, in salary and overheads 40,000 UA/year): 11.1 UA/100 words

The breakeven point is thus:

$$\frac{92{,}500 * 100}{(11.1 - 3.9) \quad * 1.000} = 1.3 \text{ million words for translation per year}$$

Another analysis of the Systran system is available in the article "Economic Aspects Of Machine Translation", Reference 11.10 in which the break-even point and pay back period for machine assisted translation systems are discussed. In order to compute the profitability, one should establish the number of words necessary to be translated to write off the investment costs. A rough formula would be:

$$N = \frac{I}{CH - CM}$$

where N = the minimum number of words to be translated to write off the investment
I = investment
CH = cost of human translation, per word
CM = cost of machine translation, per word

Using this formula, the break-even point for Systran would be:

$$\frac{\pounds\ 187{,}000}{\pounds .0854 - \pounds .0607} = 7{,}570{,}000 \text{ words}$$

Since The Commission's activity is less than N words, the investment will become profitable only after several years, and one should take into account the interest rate on the capital invested. The formula now becomes:

$$I = \sum_{t=1}^{x} \frac{n\,(CH - CM)}{(1 + i)^{t}\ x}$$

Supposing n = 1,900,000 words/year and i = 10%, then:

$$187{,}000 = \sum_{t=1}^{x} \frac{1{,}900{,}000\,(.0854 - .0607)}{(1 + i)^{t}} = \sum_{t=1}^{x} \frac{46.930}{1.1^{t}}$$

If t = 5,
187,000 = 46,930 * 3.7907 = 177,898
If t = 6.
187,000 = 46,930 * 4.3552 = 204, 390

Thus, the pay back period is between five and six years.

6.3 THE TAUM-AVIATION SYSTEM

Gervais (1980) puts the direct cost of translation (personnel + equipment) at 12.7 for human translation and 16.0 for machine translation.(price in Units of Account by 100 words, 1980 values).This assessment gave a negative result, mainly as a result of two elements:

- the excessive cost of processing of the MT on the computer;
- prohibitive cost of the coding of the entries in the dictionaries.(Reference 11.6)

6.4 THE METEO SYSTEM

Meteo is a machine translation system that has been in successful operation since May 1977. Out of the 18,000 words that the Meteorological Sub-section has to translate every day, about 11,000 are submitted to the Meteo system. It translates about eighty percent without human intervention. Meteo translates between three, and three and a half million words a year. The existence of the system has allowed the French translation service for the weather bulletins to be extended across Canada without any increase in staff.

An estimate of the computing costs of the system established them at about $0.03 (or $0.04 at the most) Canadian dollars per word. It is difficult to evaluate the cost more exactly. Should one, for example, use the commercial rate for fifteen CPU minutes a day of the CDC 7600 computer - which would have gone unused in any case? Meteo, with its limited vocabulary and sentence structures, is considerably cheaper than the other systems.(Reference 11.10)

6.5 CONCLUSIONS

In conclusion, in 1981:

- the cost of raw MT works out at 1.25 UA/100 words, not including the acquisition cost or amortization of the capital expenditure
- the post-edited MT costs much the same as revised HT, or slightly less, around 8 UA/100 words

If only the direct cost is considered, MT appears economically viable:

- without post-editing, to provide users with a product which is:
    - -- two to four times less expensive than HT
    - -- in principle, produced much faster than HT
    - -- but of a quality which is much poorer than the corresponding unrevised HT.

It can be suitable to permit scrutiny, by the end-users, of non-official and short-term documents, and of scientific and technical texts:

- with post-editing, to provide users with a product:
    - -- either costing the same as unrevised HT, but of a better quality and a turnround time which can be slightly faster
    - -- or costing less by (nearly 30%) than revised HT, with an equivalent quality and a much shorter turnround time.

It can be suitable to provide official and longer-term documents, on condition however that the employers succeed in motivating some translators sufficiently to be willing to do post-editing work.

It seems that if MT is to take off, the capital outlay required to achieve the various technological improvements needed will need to be lower.

7- MARKET TRENDS/STRATEGIES

7.1 Of the three linguistic levels required for accurate translation--syntactic analysis, information sharing between sentences, and general world knowledge-- only the first is fairly well understood and present in current translation products. Determining a sentence's meaning based on its relationship with other sentences is not so advanced, although it represents a hot area of AI research and progress can be expected in this field. Creating a computer with a broad base of world knowledge, however, is a problem of profound difficulty. Giving a machine such ability is far beyond the current state of the art in computer science and will not occur in the foreseeable future. (Reference 11.9)

7.2 A careful look at current development projects (eg. METAL, Projet National, EUROTRA) * reveals that there is currently a strong tendency to build ambitious wide-scope systems on the basis of a relatively conservative technology. The projects will turn out second-generation MT systems which do not seem to embody much in the way of dramatic technological improvements over the systems developed in the seventies.(Reference 11.1, Appendices)

Instead of trying to build general purpose MT systems, Cognos Inc. (Reference 11.1) believe that the best approach would be to use what is now known about the relative complexity of sublanguages that would appear promising, even if more difficult, for example, than weather forecasts. There seems to be no reason why Canada could not become a leader in MT research by applying, adapting and improving the latest techniques in syntax, semantics, pragmatics and knowledge representation. In addition, some original research on translation theory and its relations to AI/NLP could be pursued. The aim would be to acquire, after a few years of research, a capability to develop MT systems that can clearly overcome some of the limitations of current systems.

The Cognos study recommends that immediate consideration be given to mounting a machine translation program with the objective of developing new generation, intelligent machine translation systems. Although capability and interest in several MT, NLP, and AI fields exists in Canada, it is generally fragmented and separated geographically at various universities and institutions across the country. It is emphasized that such a MT program must be:

(a) diverse in its approach,

(b) varied in its goals (long-term goals must be kept in mind as well as sub-optimal solutions to satisfy critical immediate needs),

(c) cast as part of the more general goal of developing AI and NLP research programs, as well as basic research programs into knowledge representation, computational linguistics, computational reasoning, and the human-computer interface.

---

* a) METAL (sponsored by Siemens Corp. at the University of Texas) is a second generation system which has been developed for German-to-English translation of telecommunications equipment manuals.

b) Projet National (formerly ESOPE) is a pilot project oriented towards technology transfer between university and industry. Prototype systems would be created for two or three pairs of languages: English to French (and possibly Spanish), and French to English.

c) EUROTRA is a large-scale project that is intended to be capable of translating several types of texts among all 7 languages of the EEC (i.e. between 42 language pairs).

7.3 Four of the sublanguages included in the study by Dr. Kittredge (Reference 11.3) proved to be sufficiently simple that one can already foresee the introduction of a completely new approach to bilingualism in these texts. Certain texts such as informative stock market reports, agricultural market reports, weather reports, and other "statistical" reports all contain information which can be reduced to tabular form. In most of these cases, it is now possible to begin generating the natural language reports directly from the data. Both English and French texts will be synthesized from the same data. When reports can be generated in one language, it is trivial to generate them in a second, third, etc. language provided that text and sentence grammars have been worked out in those languages.

One of the best ways to prepare for introducing this new technology to Canadian translation is to prepare detailed grammars for all the sublanguages in question for experimental construction of bi-lingual text-generation systems. Within three to five years it should be possible to synthesize certain texts directly from data bases concerning market activity.

8- MARKET SUPPLY

8.1 INDUSTRIAL PLAYERS

The products of three companies -- Automated Language Processing Systems (ALPS-Provo, Ut.), Weidner Communications (Northbrook,Ill.) and Logos Computer Systems (Wellesley, Mass.)-- represent the state of the art in commercial translation software. (Reference 11.9) Each company approaches the translation task in a different way, but they are similar in that they make no claims of performing totally automatic translation.

All of the available machines translate by reading each sentence in isolation. The software packages are also similar in that each takes the user through a dictionary-building phase. When presented with a new document to translate, the machines scan it and produce a list of words they don't recognize. The user can then enter a definition along with other information about each word for the machine's future reference.

Beyond these product similarities, some differences emerge. ALPS' product, for example, is designed to work in an interactive manner, with the translator guiding the software through the text, one sentence at a time. The products from Weidner and Logos operate in batch, with the software making a first pass at the translation and a human translator editing the output.

The following table provides a comparison of selected features of the Alps, Weidner, Logos and Systran systems.

OPERATIONAL SYSTEMS

FEATURES	ALPS	WEIDNER	LOGOS	SYSTRAN
SOFTWARE RUNS ON:	IBM PC/AT, Xenix operating system	Digital, VAX/VMS IBM PC/AT	Wang VS and IBM VM/CMS	IBM Mainframe
SUPPORT TRANSLATION FROM:	-English to French, German, Italian, Spanish -French to English	-English to French, German, Italian, Spanish, Portugese -French to English -German to English -Spanish to English -Japanese to English	-English to French, German, -German to English	-English to French, German, Italian, Spanish, Portugese, Japanese & Arabic; -French to English -Russian to English -Japanese to English
COST:	-Each 1-way lang. pair costs $13,000 per workstation -$120,000-$400,000 depending on confirmation	Cost of 2 directions (Eng. to Fr. and Fr. to Eng.) be $50,000; for 4 directions $85,000	Licenses its software -about $56,000 for hardware and software charges according to use, about 25¢ per line translated	Approximately $5,000 per month for software for the first language pair
OTHER FEATURES:	Multilingual word processor and access to on-line dictionaries in both languages	Text editor allows split-screen viewing of source and target texts. MicroCAT, a scaled down version	Semantic Abstraction Language as an intermediate code between source and target languages, specialized dictionaries, post-editing some interactive pre-editing	Specialized dictionaries, post-editing, interfaced with OCR and word processing
USERS:	Agnew Techtran (a translation house in L.A.), Xerox in Rochester) & Control Data Service Bureau in France	About 25 systems sold, Mitel Corp., Cosmopole Communications (a translation house)	Omniplex (translation house), & several German companies	U.S. Air Force, NASA, EEC, GM & Several others.

Comments were made evaluating these systems. All of the systems were felt to have weak linguistic processing. The Alps and Weidner systems provided raw translations that were frequently ungrammatical or incorrect and never more than literal. The Logos system lacked any treatment of compounds which is clearly a problem (especially for German), however, it was still too early to judge this system. The Systran system has very comprehensive dictionaries but cannot achieve more than raw translation with frequent mistranslations or ungramaticabilities.(References 11.9 and 11.1)

An excellent discussion of second-generation prototype systems is provided in Reference 11.1. The system's sponsor, approach taken, goals, coverage, and evaluation are some of the elements discussed.

8.2 TRANSLATION AGENCIES

The Translation Bureau conducted a survey of the official languages translation market in Canada. (Reference 11.8) The goal of the study was to identify , to the extent possible, the official-languages translation market in Canada, including:

- those in the market,
- their location,
- their production capacity,
- their specialties,
- their rates, and,
- the proportion of this market represented by the Translation Bureau.

From various register databases the study identified 1,398 translation agencies including both independent free lances and companies. The 1,398 questionnaires were mailed in December, 1982. A usable response rate of 32.4% (412 questionnaires) was achieved. The percentage of respondents for each geographic area corresponds fairly closely to that area's proportion of the national population. Some of the findings were as follows:

- The average number of full-time translators is 3.3 for the companies and 0.7 for the unincorporated agencies (UAs).
- The average number of part-time translators is 5.3 for the companies and 1.1 for the UAs. If the three companies which each have more than 100 part-time translators are taken out, the average for the companies drops to 2.9.
- For the 193 agencies which stated that they receive work from the Translation Bureau, this source accounts for 67% of their supply.

- 16% of the agencies which stated that they deal with the Bureau obtain almost all of their work from it, and 59% obtain over 75% of their work from the Bureau.

- Private companies were mentioned as a source of supply by 293 agencies, which receive 53% of their work from them. Of those agencies, 61 or 21% work almost exclusively for private firms.

- The Bureau is almost the sole source of supply for only 4% of the companies, but for 33% of the UAs.

- 63% of the respondents felt that the private sector pays more than the public sector, and 27% felt they paid the same

- There were marked differences in the demand forecast depending on the geographic area:

GEOGRAPHIC AREA	INCREASE	DEMAND STABLE	DECREASE
Montreal-Quebec (N= 189)	25%	50%	25%
NCR-Ontario-Toronto (N =163)	26%	66%	8%
West-East (N = 25)	54%	42%	4%

- Respondents were asked if they had considered doing their work on a word processor and providing the text to the client on both diskette and paper. 24% already do it, 40% would consider it, and 36% said no.

- The main reasons given for charging a higher rate were: urgency, difficulty of text, specialized texts, and special physical layout.

- 33% of the respondents offer volume discounts to clients, and another 30% would consider it, only 23% of the respondents were opposed to it.

- "Administration" as a specialty was mentioned a total of 166 times by the 291 respondents to this question.

Using the volumes of specialized work indicated by the agencies and the percentage each specialty represents in their total specialized work volume, the study established the current volume, current annual unused capacity and short-term annual capacity for each specialty by geographic area. The current annual volume for 1982 (calendar year) for all their translation activities was 156 million words --of which 58 million were general and 98 million were specialized. The current annual unused capacity (i.e. the difference between annual volume and annual production capacity) was 113 million words. The annual production capacity an agency could obtain within two years was 300 million words.(See page 33 of the report)

- Comparing the rates charged by the companies to those charged by the unincorporated agencies shows that the former charge nearly 3 cents more for general texts and nearly 4 cents more for specialized material

- For work done for the Bureau, companies charge an average of 1.5 cents more for general texts and 1.7 cents more for specialized texts than UAs.

- The study then estimated the market capacity or more precisely the capacity which the Translation Bureau could use. Some of the factors considered were: the number of respondents, the number of non-respondents in the sample, the national population, the percentage of agencies likely to meet the Bureau's criteria, and the rates charged by agencies compared to those paid by the Bureau. On the basis of these factors and various hypotheses, the Bureau could have access to a total minimum unused capacity of 50 million words and to a maximum unused capacity of 102 million words.

A small survey (thirteen companies) was undertaken asking private companies which have their own translation services when and how they use the services of freelances to meet their translation needs. (Reference 11.8) The percentage of texts contracted out varies tremendously from firm to firm - from 0% for Sun Life to 60% for Bell Canada. However, the reasons for using freelances are much the same everywhere: overloads, tight deadlines, technical work. The rates varied from 15 to 30 cents per word, with most paying between 15 and 20 cents. The rate is occasionally set on the basis of the freelance's reputation, but most often it is determined by the nature and urgency of the document.

9- SUMMARY AND CONCLUSIONS

9.1 NEED FOR TRANSLATION

There has been a significant revival of world-wide interest in machine translation in the past few years, with major new MT research programs in Japan, U.S.A., and the European Economic Community. This research is being pursued in several directions, including both short-term goals that involve severely restricting the translation task or merely providing aids to the translator, and long-term goals of developing machine translation systems based on the latest developments in natural language processing and in artificial intelligence.

In Canada, the pressure to find computer aids is growing increasingly intense in a context where budgetary limitations make it necessary to find some way to increase the individual translator's productivity. The Department of Communications and the Department of the Secretary of State are formulating plans for a cooperative program to develop computer assisted translation systems in Canada. The objectives of the program are as follows:

- to assist the Secretary of State in acquiring the necessary informatics technologies including machine translation systems to more efficiently and effectively fulfill their mandate of providing a translation capability to the federal government, and

- to foster the development of a viable industry capable of supplying translation automation products to meet the requirements of the federal government as well as competing in the domestic and foreign markets for machine translation services and systems.

9.2 TRANSLATION MARKET

No comprehensive data on translation expenditures is available either for Canada as a whole or internationally. The reason for this is that, except for the small number of larger translation users (governments and international agencies), the market for translation is highly fragmented.

The world written translation market is believed to be within a range of from 80 to 240 million pages of 250 words, i.e. 20 to 60,000 million translated words per annum. A rough estimate of the volume of translation in Canada is 750 million words per year. At an estimated rate of 20 cents per word, the estimated cost for this translation is about $150 million per year.

It is also believed that there is a substantial latent demand, currently unsatisfied for reasons such as cost, time, availability of translators. This potential market could represent 30% of the actual market for formal, written, human translation.

9.3 MACHINE TRANSLATION

The potential for machine translation is still unclear and more information is needed on the user's requirements and whether significant increases in efficiency and effectiveness of translation will be realized with machine translation.

Some developments in machine translation or machine assisted translation indicate at least the potential for productivity gains. The first consists in the marketing of rough machine translations, in the form in which they leave the computer, without any human intervention. The rough translation will be returned to the applicant in less than 24 hours for a total cost of about 25% of a human translation. Such a translation may be valuable to people wishing to get a rapid acquaintance with an article, written in a language they do not understand, in the (numerous) cases where they cannot afford the cost and/or time delay for a human translation.

Post-edited computer translation, interactive or otherwise, will develop at a cost currently comparable with or slightly higher than that of human translation, but with a clear improvement in the "time" factor. One forecast was that within five years post-edited machine translation will occupy 5 to 15% of the market, and will then develop much more quickly.

Lastly, a large market will continue to be covered by translators who, while not having access to the large automated systems, will gradually equip themselves with smaller-scale data-processing aids which will increase productivity by permitting faster production of high quality translations in highly specialized fields.

9.4 OPPORTUNITY

A recent study by Socioscope entitled "Functional Specifications For A Translator's Work Station" identified numerous functional areas as potentially benefitting from computerized support. Very few of these features are now available on commercial systems, although a few companies are beginning to introduce products with a few of these features as part of their machine translation systems.

An opportunity exists for Canada in the machine translation area since:

- There is a demonstrated expertise in Canada in the design of advanced and large scale machine translation systems, as evidenced by the TAUM project. A large amount of general experience with translation and machine aids to translation (such as terminology banks) exists in the Translation Bureau of SoS and in a number of private translation services. In addition, Canadians experience in expert systems and NLP serves to provide depth to this background of expertise.

- Canada's strong national need for translation could serve as a stimulant to the technology. The most conservative estimate of the Canadian translation market gives figures of 750 million words per year. Automation of even 10% of this volume would result in a saving of at least $20 million per year.

- In the world market, MT is a significant niche market in which Canada has a tremendous opportunity.

Two features characterize current research in machine translation. The first is a basic research program directed at a better understanding of the possiblities and limitations of current technologies. It is now widely accepted that a fully automatic high-quality translation of unrestricted text require systems with a broad base of world knowledge and a deep understanding of natural language. The second feature of current research is the attempt to provide useful though limited systems for machine translation or machine aided translation. In other words, it is a search for various suboptimal solutions to the machine translation problem. There is a critical need to maintain a balance between the unquestionably important task of applying available techniques to practical problems, and the equally important effort to extend our basic knowledge and understanding in order to provide a foundation for future generations of MT systems.

10- DEFINITIONS

TRANSLATION

- the "traditional definition" : the process of replacement of a text written in a source language by a text written in a target language, the objective being a maximun equivalence of meaning.

- the "modern definition" : the process of transfer of message expressed in a source language into a message expressed in a target language, with maximization of the equivalence of one or several levels of content of the message: i.e. referential (information for it's own sake, eg. organizational note), expressive (centred on the sender of the message, eg. speech), conative (centred on the recipient, eg. publicity), metalinguistic (centred on the code, eg. dictionary), phatic (centred on the communication, eg. courtesies), poetic (centred on the form).(Refernce 11.6)

Machine Translation

The majority of the MT applications require a human intervention at some stage of the translation process:

- either before inputting the text into the machine: the pre-editing consists in rewriting the source text to simplify it and/or to remove a number of ambiguities

- during the work: ambiguities not removed by the machine are submitted to the translator who resolves them

- or after completion of machine translation: post-editing makes it possible to correct the machine's mistakes and to provide a text of a quality similar to that of a human translator

Typology of translators

- technical translator
- multilingual secretary
- terminologist
- conference interpreter
- interpreter-translator
- conference translator
- press translator, information, advertising
- language teacher
- revisor
- documentalist
- minute-writer
- cinema-TV translator
- unspecialized translator
- editor
- literary translator
- translator or interpreter attached to a court

11- REFERENCES

11.1 "Natural Language Automated Processing and Artificial Intelligence"
Volume 1: State-of-the-Art
Volume 2: Technology Thrust
Volume 3: Possible Social and Economic Impacts
Volume 4: Opportunities for Canada
Volume 5: A Program Plan
Volume 6: Priorities for Canada
By: Cognos Inc. and sub-contractors
For: Secretary of State and Dept. of Communications
Date: February 29, 1984

11.2 "Machine Translation and Natural Language Processing: Opportunities for artificial intelligence in Canada"
By: Cognos Inc.
For: Dept. of Communications
Date: February, 1985

11.3 "Sublanguage-Specific Computer Aids To Translation: A Survey of the Most Promising Application Areas"
By: Dr. Richard Kittredge, Universite de Montreal
For: Secretary of State
Date: March, 1983

11.4 "Functional Specifications for a Translator's Work Station"
By: Socioscope Ltd.
For: Canadian Workplace Automation Research Centre, DOC
Date: March, 1986

11.5 "Trial of the Weidner Computer-Assisted Translation System"
By: Bureau of Management Consulting, DSS
For: Secretary of State
Date: October, 1985

11.6 "Better Translation for Better Communication"
By: G. Van Slype
For: Commission of the European Communities
Date: 1984

11.7 "A Survey Of Machine Translation: Its History, Current Status and Future Prospects"
By: J. Slocum, in Computational Linguistics, Vol. 11, no. 1
Date: 1985

11.8 "Survey of the Official Languages Translation Market in Canada"
By: Bureau of Management Consulting, DSS
For: Translation Bureau, SoS
Date: 1983

11.9 "Machine Translation Poised For Growth"
By: Charles Connell
For: High Technology
Date: June 1986

11.10 "Practical Experience of Machine Translation"
By: V. Lawson (editor)
For: Proceedings of a Conference, London
Date: November, 1981

INDUSTRY CANADA/INDUSTRIE CANADA
98741

HILLMAN, BEV
--Computer assisted translation :
preliminary market and industrial
assessment

P
308
H56
1986

DATE DUE

JUN
JUIN 30 1992