June 1972, Ottawa, Ontario, Canada.

> "Nothing is impossible; there are ways that lead to everything, and if we had sufficient will we should always have sufficient means. - It is often merely for an excuse that we say things are impossible."
>
> La Rochefoucault.

MACHINE TRANSLATION IN CANADA

BY ANDRÉ R. GOUIN

INTRODUCTION

Machine Translation (MT) in Canada finds itself at a significant and very important stage of its development. Thus a re-examination of the situation appears to be a necessary exercise in order to assess the progress made and to allow that a decision to proceed be taken on the basis of considered facts in front of the decision maker. A two-day seminar on Machine Translation was held at the National Research Council, Ottawa, Ontario Canada on 27-28 April 72, where authoritative representations were made with the clearly evident purpose of exhibiting as many of these facts as possible.

AIM

The aim of this paper is to discuss some of the views expressed during the conference and to add some remarks and conclusions which, it was felt, could not be made at the seminar proper since this writer was in attendance only as an interested observer.

PURPOSE

The purpose of injecting this into the discussion at the present time is to provide the opportunity to consider, what it is believed new facets of the subject which were not, or did not appear to be discussed exhaustively during the seminar and thus to provide additional facts to the decision makers on which to base their decision:

to discontinue support of research, in this field;

to continue basically in the same fashion as previously; or

to introduce new approaches.

OUTLINE

The subject will be examined from three angles. That is, the views, remarks, observations and recommendations as presented during the seminar will be summarized and discussed. This portion may be difficult reading for those readers who did not attend the seminar since they will not have had the benefit of hearing,

first hand, the remarks under discussion here. However, this should be compensated by the more complete discussion - which will follow this part. This latter discussion will take two attitudes: one will be theoretical and the other practical. Following this will be presented some conclusions which will attempt to formulate recommendations of what should be a reasonable attitude to adopt with regard to continued research related to MT.

THE SEMINAR

The seminar was very successful and invigorating for those participating and for those in attendance as interested observers. Views and ideas were advanced vigorously by the conferees coming from and representing a wide spectrum of interested agencies. It seems appropriate to discuss some views at this time with the full benefits of the completed seminar. Since this must be done completely from recall it is imperative that the reader appreciate that any views or attitudes apparently assigned to certain conferees may not necessarily be theirs and that if, and when, this happens it is only an expression of what this writer felt the attitude of the conferees appeared to be. This feeling may be erroneous in a few cases and the responsibility for this misinterpretation is accepted and apologies are offered in advance. Every care will be taken, naturally, to indicate clearly where what is expressed in fact reflects the impression made on this writer.

Professor Garvin advanced clearly personal views on the field of machine translation in general. Because of Professor Garvin's previous valuable contributions to the field of MT, from a linguistics point of view, these are highly respected and rightly so. One point, which appeared to this writer, that he seemed to wish to make abundantly clear was that the whole subject of MT tended to divide the researchers in the field in two camps; those who claim that MT is impossible and those, on the other hand, who contend that it is possible; with the degree of separation between the two being related to the length of time members of each group have been exposed or have dealt with MT. Another point which he advanced and which is also reflected in his article in the "Feasibility Study" (1), is that the expression "High Quality" in "high quality translation" is undefinable in absolute terms. Thus "High Quality" must be judged on a relativistic scale whereby a product may be of high quality compared to one thing and at the same time be of low quality when compared to something else. As a direct consequence of this, it appeared to this writer that Professor Garvin was wishing to place on guard those who have to judge the results of MT against making an absolute judgment but rather to retain their objective cool and keep in mind the relative degree of quality required by the eventual user of the product. More will be advanced later on this important area of the subject. He further proposed what he felt appeared to be the two dominant approaches in the field today.

Those were what he chose to call the bipartite and the tripartite designs. Summarily these could be said to reflect the attitude of the different researchers in their approach to MT as follows:

BIPARTITE: The bipartite design is where, to use Professor Garvin's words

(1), "the machine translation program will consist of only two portions: a dictionary, and a translating algorithm" where this latter algorithm consists of a parsing algorithm and the set of grammar rules of the source language.

<u>TRIPARTITE</u>: The tripartite design is one where the parsing algorithm is purposely kept distinct from the set of grammar rules of the source language. These rules are set up in table form and are called as required by the parsing algorithm.

He concludes on this subject by suggesting that operationally he prefers the bipartite design because "this is the only way to insure that a given rule will be called only after all the conditions that are necessary for its operation have been recognized by other rules of the program, and that such a recognition has been effected in the correct order." (1).

Professor Garvin made numerous thought provoking interventions during the subsequent duration of the seminar and his paper in "the Study" should be of interest to most of those active in this field.

Professor Vauquois gave a short review of the situation of MT as it developed in the recent past and as it is found to exist at the present time in France. He pointed out that a recent move of their system from the IBM 7044 to one of the third generation computer introduced a horrendous problem of program conversion. Nevertheless the continued cooperation between Grenoble and Montreal is expected to overcome these problems in time. Professor Vauquois then introduced the linguistic approach to MT which they favored. This, as understood by this writer, proposes to study the translation problem in an attempt to discover the linguistic level at which a transfer from the source language to the target language can best be effected. Among these different levels are found the morphological, the syntactical and the semantical levels. He indicated that it was realized early that a transfer at the morphological level was clearly insufficient and hardly useful. This led their research to study the higher syntactical level. Here the transfer seemed to provide more useful results but at the same time pointed to the necessity to address the yet higher level of semantics. Professor Vauquois seemed to imply, to this writer at least, that this problem will require a tremendous and lengthy effort before we are in a position to state that we have a semblance of solution at this higher level. This was a useful way of presenting the translation process as it is helpful in conceptualizing the stages through which a human translator would be required to pass to achieve a so-called high quality translation. More on this later.

Mr. R. Kittredge introduced the method used by the University of Montréal group. The method is the so-called Q-system. Unfortunately the presentation did not make completely clear what exactly the Q-system consisted of. Nevertheless, the presentation coupled with a number of private conversations with Mr. Kittredge and others of the Montreal group provided sufficient information to allow the following impressions to form. It appears that the Q-system is a higher level programming language, somewhat like a set of Macro programs or commands, which provides a

powerful tool to the linguists in dealing with the computer from their point of view. The system is applied on a CDC 6600 and is developed in the COMPASS programming language. The linguistic approach, as described by Mr. Kittredge, seems to favor a high degree of formalization of the tree structures in the source language, English, in this case, as an attempt to describe the so-called deep structure of the input sentence to the fullest extent possible before the transfer to the target language takes place. The approach is certainly an interesting avenue to

follow which could provide useful results applicable in Professor Vauquois' links between the syntactical and semantical levels or in Professor Garvin's bipartite method by allowing a more founded basis for rule inclusion in the translating algorithm.

Mr. B. Harris then introduced another approach adopted by the Montreal group. This, he called the English-French transcoding system. The significance or advantages which could be gained from this system were not detailed, but these seemed to be based on the assumption that it is easier to recognize syntactical errors in a native language than it is to discern whether the morphemes used are the correct ones to convey the meaning of the source language. If this assumption is admitted, then it would seem a good first step to transcode English morphemes into French ones and to provide an output which gives French morphemes within an English syntax framework. The suggestion was made during the presentation that the method was not much, if at all, different than a word-for-word translation and little was offered to contradict this suggestion. Nevertheless, the approach was novel to this writer and seemed to offer some interesting possibilities in a different application which will be suggested later. Mr. Harris continued with a presentation of the Montreal group's transformulator approach for text with a highly formatted contents such as specification sheets, contracts etc. Here the standard clauses of paragraph length or only one word long are assigned formula numbers. The corresponding target language formula is retrieved on the basis of the source language reduction, with only the odd parameters, which are known to change, requiring to be indicated. The system was demonstrated showing promise for this type of specialized application.

Mrs. K. Booth presented some of the effects on a MT system when a new text is introduced as a test. The steps required to be taken to bring a system to the degree of versatility needed to handle new text was exposed during this presentation. The description, clearly reminded this writer of Yngve's, and others, cyclic method, although not stated as such. Convinced of the benefits offered by this method it was gratifying to see it applied. For those who may not be conversant with the cyclic process and as a recall to others, basically the method consists of the following steps:

    a. a dictionary is coded for a limited text;

    b. a parsing program is developed for the text;

    c. the text is introduced to test the parsing and dictionary;

d. additional rules are developed to improve on the text;

e. these new rules are tested to ensure they exhibit the desired effects on the same original text;

f. a new text is introduced;

g. the dictionary is updated to include the "not-found" words;

h. new rules are developed to handle the special cases found in the new text; and

j. the procedure is repeated continuously.

The proponents of this approach contend that the method is convergent, that is, as more and more text is introduced less and less additions to the dictionary are required and also the necessity to introduce new rules is also decreased. It is clear that this approach necessitates that the introduction or change of a rule must not require changes in the other rules otherwise it would blow up out of proportion. Some interesting statistics were presented by Mrs. Booth to show that when a new text is introduced into a system developed on statistics built up on another type of text, the results may or may not be acceptable. An interesting possibility emerges from this suggestion which will be detailed later.

Other members of the Saskatchewan group presented detailed approach to a number of problems encountered by them in their system. These need not be entered here as they reflected more on concerted efforts of a more specific nature such as:

- Algorithms for distinguishing the reduced passive and past tenses.

- Improvements in the idioms and dictionary routines.

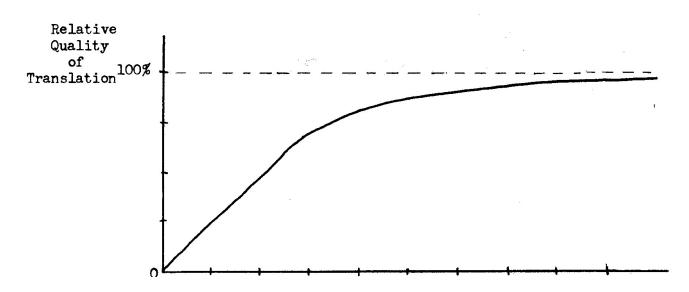- Constructing a micro-glossary for a technical paper.

These were very interesting indeed and quite inviting of comments, but, best reserved to a communication outside of this paper.

Mr. D. Savage then followed with a review of the technical progress to date as related to MT. He aptly pointed out that advancement in computer power and programming languages had surely been accelerated due, at least in part, to pressures exercised by the MT research community. Because of these advances, the problem of insufficient capacity of computers to accept the vast amount of data required for the data base and programs of MT, could more or less be considered overcome. For instance he used the input problem as one typical example which still needed to be eradicated. He mentioned that this was a problem plaguing all computing centres and that industry was seriously addressing the situation with all the marketing advantages clearly in view that would ensue to the first proponent of a valid solution. Thus it is for the user of computing facilities to clamour for the eradication of any technological misgivings by stating clearly what the problem is

and by requesting an immediate solution. Mr. Savage then addressed the subject of compatibility of software or program development between the two groups. He emphasized that the incompatibility which existed between the Montreal and Saskatchewan groups was not the responsibility of either group as this was imposed on them by the machine provided by the University, but he reiterated that efforts should be devoted to resolve the situation. The problem of incompatibility should not be considered, as serious as it was made out to be and the reasons will be apparent later when the theoretical and practical aspects of the MT activity are discussed. He reiterated the attitude prevalent throughout the conference that the computer hardware and software hurdles previously facing MT had, for all practical purposes been eliminated and that the most urgent requirement for improvement could now be satisfied only by linguists. It is clear, Mr. Savage added, that some problems of a technological nature still remain but that these are no longer of such constraining seriousness as to prevent further advances from the linguistic sphere of MT.
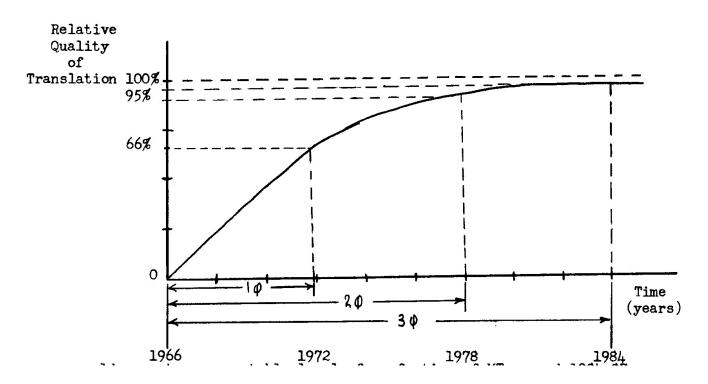
Mr. Clément then produced some statistics reflecting on the degree of effort required to bring machine produced translation up to the standard of that produced by a human. The comparative table presented indicated that MT was still a very long way from becoming a useful translating method. Serious reservations were expressed by some conferees on some of the criteria used to evaluate the relative merits of MT as compared to human translation. Nevertheless, the comparison certainly has valid merits but should be enlarged to provide a more objective scale on which to judge the level of achievement of MT. This subject will be discussed more thoroughly later when concrete proposals will be made to improve on this evaluation method.

Mr. Mayer indicated that the level of achievement of MT as demonstrated was still unacceptable to assist the Bureau in their translation tasks. The requirement of the Bureau, he defined as being 10% of the total translation produced needed to be excellent, 50% had to be very good and the remaining 40% required to be good. The seminar concluded with a general discussion. The major thought which appeared to be expressed during this closing period was a subtle plea, by those interested in the activity, that research support should be continued. The chairman then presented a chart on the board to the conferees as shown here :

Time (years)

The "years" scale was purposely undefined and similarly the definition of "perfect translation" left open. He, then invited comments as to where on the curve one might find MT as it exists today. Dean Booth introduced a few engineering techniques by selecting a point on the ordinate (translation axis) as representing a level of approximately 66 2/3% of achievement. By projecting this point to the abscissa (time scale) and stating that the time taken to achieve this, was 6 to 7 years, we had what is called in engineering terms one time constant. Based on the usual definition of time constants and, if this theory was applicable in this case, then we could expect to achieve a 93-95% level in three time constants. This would indicate, as reflected on this modified graph, that we could expect an acceptable level of perfection of MT around 1984-87.

This was obviously proposed by Dean Booth with tongue in check but nevertheless does reflect on a possibility which has been found to be true in countless other circumstances, even in cases where there were no remote possible link.

Professor Vinay reflected that : the goal of MT was to translate from a language L1 into a language L2; that the problem was energetically addressed by linguists and computer technologists; and, that therefore it was rather curious and most unfortunate that the translators or the translation theoreticians could not be found to be actively participating in this enterprise of particular significance to them. He emphasized on the absolute necessity for these two groups to influence the activity.

The seminar was efficaciously conducted and covered a wide spectrum of the MT field. It was most gratifying to have the opportunity to attend since a much better insight of the MT problem was gained by attentively listening to the authoritative views expressed there. The seminar might have dealt more with longer range objectives than it did but this might have required an unacceptable lengthening of the proceedings. Setting the sights farther in the future might have revealed the unfairness of comparing the results obtained from MT, which has only had about six or seven years of limited endeavour, to those achieved by a human translator who has anywhere from 16 to 20 years of extensive and specialized scolarity, besides the additional experience acquired by him at work and outside. This last observation, it is hoped, will help to set the stage for the next two sections which we will now enter into. In these it is proposed to advance some more personal thoughts on the subject of MT by looking at the MT activity from the theoretical point of view and then to follow this with practical considerations.

THEORETICAL CONSIDERATIONS

First of all, it should be quite clear that before a theory of MT can be elaborated, proposed and tested, a theory of human translation (HT) will have to be established. This will surely not be an easy task and will require a concerted effort from the translation theoreticians. It is proposed to show that MT may prove a useful tool to advance this theory of human translation.

Not understanding clearly what the theory of HT is all about, the activity of MT must therefore search elsewhere for a raison d'être at least for the present. There are two possibilities. One is, that if such a system achieved a certain degree of quality it could be immediately applied to alleviate the quantitative burden faced by human translators by taking over the translation of material for which it is qualitatively capable of performing. We will have more to say on this possibility later when we deal with the practicality of MT. The other possibility is that, since we do not have a theory of translation, then MT could serve as an indispensable model to help us understand the process and thus help us build or promote such a theory of translation. Therefore, from the theoretical point of view, MT must be regarded as an attempt to build a model of HT.

The next question might be, or a first question might have been, is a theory of translation a necessary and worthwhile human activity? It is suggested here, as is often done elsewhere, that if the contention that natural languages are vehicles of thought, then their study might reveal some interesting clues on the thought process itself. If this suggestion is admitted and since translation deals more or less simultaneously with two such vehicles, any advancement of a theory of translation could in many ways provide additional clues on the thought process. Another next question might be: how can MT help the translation theoreticians? Theoretically it is possible to surmise that if MT were to develop more or less along the lines that HT follow then this parallelism should afford more possibilities of explaining certain of the rules reflected from one into the other. That is to say, that while we are modelling the HT process in MT we can expect some wild results, especially when we do not

understand clearly what it is we are attempting to model. On the other hand, if we develop the MT model along intuitive lines of how HT is suspected to work, then we may obtain some valid results some of the time (these at least would add to what we now have).

It will be noticed that there is an imposing element of chance in what is advanced here, but it should be one of the prime goals of the translation theoreticians to reduce this element of chance by providing a firmer theoretical base on which to develop the model. Until such time as these theories are forthcoming we have no choice but to approach the activity from the empiricist point of view. The final linguistic solution is not available yet and may likely not ever be, therefore the theory of translation can only evolve empirically. The present activity of HT offers little possibility of providing the required vast amount of observations of the translation process from which theories can be developed. The reason for this is that the human translator is forcibly interested in producing translation, improving the quality of his product by taking every opportunity to augment his experience, and not in describing how he does it. A possible reason why so few translators have tended to become translation theoreticians must be that they realize the magnitude of the task based on their own experience at attempting to describe, in their own mind, the process they actively pursue. This becomes much more evident when they realize that, even if they should propose a theory or a partial theory, they would be faced with the almost insurmountable task of assembling the sufficiently wide supporting data base in actual application, knowing full well that few of their test subjects (other translators) would have the time or inclination to devote to this distracting activity (from their own normal and pressing duties). Faced with this dilemma most translation theoreticians can hope to see their theories disputed or corroborated in only a few, if any, cases. This situation is obviously not very encouraging to those few who devote much effort at arriving at their hypothesis and conclusions.

How can MT help to overcome this intolerable situation? One possibility which seems to emerge, rather obviously, is that once a minimal MT system is evolved, then any translation theories which pertain to translation improvement methods could certainly be tested on the MT system within its limits. If the theory fails then only two conclusions are possible: either the MT model is not a true reflection of the HT process in that particular case or the theory is not valid. One or the other conclusion leads to improvement in the field and as such must be entertained as a valid intellectual activity. On the other hand, the ability of the machine to process and correlate vast amounts of data cannot be argued. This asset can be used to suggest to the theoreticians innumerable avenues worthwhile of their investigation. For example, the machine can easily provide history on how a particular linguistic structure has been handled in a vast number of varying but similar cases. The theoretician can reject the results which are obviously invalid; he can amend those which are quasi-valid, and; he can attempt to explain or correlate those which are valid. Without the machine this task can be appreciated as being almost unachievable within a reasonable time span. Thus, in conclusion, since progress in MT can only emanate from linguistics, since a theory of' translation can only be achieved within a reasonable time period with the help of the machine, (in particular

as used in MT), and since discoveries in linguistics must be applied, proved or rejected somehow on a translation process, whether human or machine, and that the human translator appears disinclined to applying theories which would only retard his pressing activity, it would appear that the MT activity is bound to remain alive whether we like it or not. We had better like it if we are honest about wishing to advance the theory of translation. Why is it that after so many years of effort in MT do we not have more encouraging results today? Let us attempt to answer this question in the next section on practical considerations.

PRACTICAL CONSIDERATIONS

Research, in Canada, has been addressing MT for six or seven years. Any one who would consider this a long time for the little results reaped cannot appreciate the length of training a human translator has to undergo to achieve worthwhile results himself. It is nevertheless good to review the work done and the results obtained and this was done during the seminar to a certain degree. It is hoped to show here where some aspects of the work done by the two groups in Canada were not fully appreciated and where certain alternatives were not exhaustively explored. The objective will be to suggest how, when and where these alternatives might be pursued.

From the seminar clearly emerged three points of view from which MT was surveyed. Two of these as practised by the two Universities (of Montreal and of Saskatoon), were from a theoretical vantage point while at the same time straddled with the iniquitous responsibility of showing immediate practical results from their efforts. On the other hand the third point of view concerns itself with the practical aspects of MT: that is, the Secretariat of State is clearly interested in how the results from the other two can be applied immediately to relieve the pressures imposed on their limited resources. It is contended here that the present approaches and attitudes will never achieve useful and practical results, either theoretical or applied, so long as the two separate objectives remain intermingled as they are now. In other words, pure research should not be constrained by or dependent on any practical results it obtains and applied research should not be conducted on systems primarily oriented toward theoretical work. This, attempts to summarize what experience has shown to be true in innumerable cases and it is not the intent here to convince the reader of the truth value of the pure versus applied research need of separation while retaining some reciprocal influence.

Admitting this proposition as true we can then look at what exists in the field of MT in Canada and determine how the two objectives can best be achieved. The two objectives need be repeated:

From Pure Research - The prime objective of pure research in MT is to develop linguistic theories which when tested and proven can be applied to the development or advancement of a practical system.

From Applied Research - The prime objective of applied research in MT is to devise ways to test, prove and incorporate the correct linguistic theories as emanating from

pure research to improve an MT operational system and consequently to improve its practical usefulness.

Members involved in one or the other activity must always keep in mind the ultimate goal of MT to provide a useful and true model of the HT process so that advances made in one can be transposed to the other and vice versa.

With these points in mind we can ask and answer the question: What does exist now in Canada and what must be created to satisfy these two objectives? In the pure research we can say that the efforts at the Universities of Montreal and of Saskatoon fulfil this requirement well, if only to a very small degree. They would do much better if they were relieved of the annual reaffirmation of their need-to-be. It seems that this has been the unfortunate result of expecting applicable results from this type of research. How can the two groups be relieved of this impractical constraint was suggested above in the implementation of a separate operational development system, separate and distinct from the system on which each group depends for its research, It may be suggested that this would only introduce a third element of incompatibility into the activity. This may be true but not necessarily since the incompatibilities which exist now in the two systems are of techniques and not of theory. In other words it does not matter in what programming language a linguistic algorithm is written or on what machine it runs so long as the algorithm provides the desired results.

It is true that programming language compatibility offers significant advantages in transferring from one machine to another. This compatibility is achievable by the use of higher level programming languages. For example, in the SYSTRAN system used at Dayton, Ohio, a set of MACRO commands has been developed which allow the linguists to write their own programs with relatively no regard to the machine on which the programs will run. This is a type of language which could be propagated throughout the different systems in Canada thereby eliminating the programming language incompatibility and at the same time, because of its machine independence, eliminate or significantly reduce the machine incompatibility as well. Linguists can learn this language and become efficient in its use in less than a week. The language however introduces a degree of inefficiency in use of the machine but, at the same time, allows so much more efficiency to the human user that its adoption is or should be indisputable (after all, what are machines for). Therefore a powerful, easy to learn higher level language is available to eliminate both the programming language and machine incompatibilities mentioned earlier. Under the present conditions then, the programming language incompatibility between the two research centres in Canada can be resolved as suggested above.

The next step is to remove the burden of practical proof from the two pure research groups. This can only be done by initiating an operational development MT system separate from the pure research activities. Before this can be contemplated however, it must be ascertained that a real need exists for the type of translation (low quality but quick turnaround) which would be produced initially by a MT system. To discover this, one must look at the 40% portion of the total translation requiring a "good" level

as brought forth during the seminar. Some of that portion could certainly be satisfied by a less than good output from the machine. There might also be some advantages gained in processing a good part of the rest of that 40% with the assignment of post-editing to bring the machine output up to the required standard of goodness. At first it would be realized that this is not the most economical way of obtaining translation of this particular portion of the total but it is a necessary requirement of any developmental enterprise.

In addition to this small portion of the Secretariat of State's total assigned translation task there is a definite need in many governmental departments, for a large amount of translation which is just not done at the present time or is done by bilingual personnel untrained for this complex duty. The level of goodness obtained from MT would certainly surpass that which is not translated and it would be very near the quality, if not better, than that of the untrained. This is the area where it is believed the initial need which could substantiate the initiation of an operational MT system could be found. Also, as the volume of text requiring translation augments the point where the human resources to produce these translations is surpassed is soon approaching, if it has not arrived already. It is clear that what will be sacrificed when this happens will be the type of translations which now requires a level of good and which will then not even achieve any level. Much more could be said to substantiate the initiation of an operational development MT program. However, in order to keep this paper within a reasonable length let us assume that the need has been shown and that the task now is to decide how to satisfy it as well as to try and indicate how such a system could eliminate some of the discrepancies indicated above in relation to the pure research activity.

A number of ways offer themselves to the initiator of a program such as this. One way is to start from the bottom with next to nothing and to thread the way to a minimal but improvable system. Another way is to make a thorough study of other similar systems, to avoid the errors made there and to benefit from the advances already achieved. Yet another way is to adopt a system already in operation, provided such a system is improvable, and to devote ones energies to improving it. The first alternative is clearly unacceptable if a possibility exists for one of the next two choices to be successful. The second choice, to study existing operational systems in order to avoid the pitfalls and to take advantage of the gains made by each, would seem to be not much less difficult than the first choice since this would mean starting at almost the same point with the addition of lengthy studies of the already implemented systems to discover those pitfalls which are to be avoided. The third choice seems to offer the best chances of success. However, this choice assumes that, in the case of operational MT, such systems in fact exist. The seminar provided the forum where the existence of such operational MT system could be discussed. It seems however that the knowledge of the existence of such systems was limited to the "I have heard of" type. This, in a way, is understandable since those responsible for these few systems are very much more active in improving them than reporting about them. Nevertheless, three systems were mentioned as reportedly being operational or having operational potential. These were: the SYSTRAN system of Russian-to-English at Dayton, Ohio; the Oak Ridge

system which is based on the Georgetown development; and the LOGOS system which translated from English to Vietnamese.

This writer has no knowledge about the Oak Ridge system. The LOGOS system has produced a translation of a 200 page text from English to Vietnamese of reportedly good quality. Little system documentation is available on which to form judgement on its potential to become operational in other pairs of languages although this claim is made by its developer, the LOGOS Development Corporation, Middletown, N.Y. Having done some work on the SYSTRAN system this writer is convinced that at least this one offers the possibilities of satisfying the requirements as stated for the third alternative above. The work done on SYSTRAN by this writer, consisted in modifying it to accept French as the source language instead of Russian. As this work was a Master's thesis project, little time was available to do extensive amendments or modifications to the existing programs., nevertheless the feasibility of modifying the system to accept at least one different language was successfully demonstrated to a level of output which only required linguistic techniques to be applied for improvement. The significance of these results lies in the fact that this small success could even be achieved on such a complex system within very limited time constraints, exceedingly short computer time availability, little linguistic background other than native knowledge of French and learned knowledge of English and engineering background.

While not excluding any other systems from consideration, this writer is convinced that SYSTRAN has all the attributes of the initial basis on which to build an operational English-to-French MT development system. For those interested, references are included at the end of this paper to some documents describing SYSTRAN or making reference to it.

To return now to more specific ideas as promised earlier in the discussion of the presentations made at the seminar. We mentioned then that the English-French transcoding system, as described by Mr. Harris, offered some interesting possibilities. One of these is as follows. If the correctness of the morphemes of a language is more difficult to discern by a native speaker than syntactical errors, then this should be demonstrable. It would seem that a transcoding system, as was proposed, could be made to show that such is the case in non-native writers of one or the other of French or English. In particular, it is thought that experiments of this nature, conducted on non-native text of young people in the process of learning the second language, would offer most useful results. For example, the machine could be used to discover at what level (as described by Prof Vauquois) do different students of a second language effect the transfer from their own language to the second one. It could be used also to indicate at what degree of competence in the second language the transfer at that particular level takes place. Further, such experiments could point to numerous worthwhile avenues of investigation for improving the methods of teaching a second language. The transcoding method is appealing in a way because it seems to compensate for what a non-native finds most difficult at least in the initial stages of learning the other language. That is, while he might not be able to recognize morphemes in his own language as correct or

not, he is quick at recognizing syntactical errors and compensating for them with little or no effort. This suggestion would offer one opportunity for the two Universities to cooperate in assessing the data from their individual areas since one offers a primarily English environment and the other French. These are only a few possibilities in this regard, by no means exhaustive, which could give invaluable results in advancing a theory of translation with the obvious benefits to both HT and MT.

Mrs Booth rightly indicated that surprising results may be obtained when the parsers, developed on the basis of statistics accumulated from one type of text, are applied to a different type of text. This indicated, in some cases, the necessity to develop micro-glossaries. One possibility here is that certain types of text might be so different in basic structure and style as to require a totally different parsing approach. This could be uncovered by applying the statistical method to a large amount of data. Where results are closely related to previous results these could be used to update the latter. On the other hand, if the new results are found to differ drastically from previous ones, these can serve as the basis for creating a different set or sets. Coupled with these different sets can be the development of micro-glossaries.

It is found difficult, by many to accept that the study natural languages from a statistical point of view could offer some chance of success. This writer is convinced that the approach has some merits in, for example, helping to eliminate some ambiguities in word functions. It is predicted that in the not too distant future the method might become almost indispensable to any MT system, particularly if suggestions, such as the one made by Dr. Petr Beckmann in his recent article "Computerization of English" in IEEE Spectrum Dec. 1971, are investigated further and shown to be true. There, the link is suggested between English, as a natural language, and the information theory of error detection and correction. Anyone conversant with information theory can appreciate the importance of statistics and can also imagine the impact on the study of languages should the link be further demonstrated and advanced, not to mention the extreme usefulness of already accumulated statistics on languages.

With regards to the development of an objective scale to score MT relative to HT care must be taken to produce a scale devoid of as many subjective elements as possible and of any biases favoring one or the other quantity being measured. The scale, as proposed by Mr. Clément, is repeated here for ease of reference. This was applied to a test corpus of 113 sentences, 2789 words.

Criteria for Evaluation	Revision of MT	Revision of HT
1. Untranslated words	0	0
2. Alternatives	817 (28%) words	0
3. Structural alterations	55 (48%) sentences	26 (23%) sentences
4. Words alterations	78 (2.7%) words	26 (0.9%) words
5. Mistranslations	10	0

5.	Mistranslations	10	0
6.	Idiomatic conformity	5-10%	95%
7.	Time required (for post editing)	17 hrs	4.5 hrs
8.	Working rate or speed (of post editor)	165 words/hr	600 words/hr

The figures in the table reflect on the effort required by a human post-editor to revise the translation of a same text as produced by the MT system of the University of Saskatchewan and that obtained by a human translator TR-2. Looking at each criterion in turn can help divulge some areas of possible improvement. The first, "untranslated words" shows that the machine and the human did equally well. It will be realized that this criterion will be met every time a test corpus is prepared for machine translation since in these cases the text will be compared to the machine dictionary to ensure that all the words are in the dictionary prior to being submitted to MT. Later, when the MT system is sufficiently advanced, the preparatory stage should not be concerned with the completeness of the machine dictionary, it will be assumed complete. The machine will then produce a list of words "not found" which will be a justifiable way of measuring the completeness of the dictionary. As more and more tests are run through, the "untranslated word" quantity should show a diminishing trend, if not, then a serious look at the dictionary build up procedures would be called for.

The second criterion, "Alternatives", reflects on a technique used by MT to leave the choice between alternatives to the human revisor. The same procedure is effected by the human translator, in the other case, who selects the best alternative himself. This is a good indicator of how well the MT system is at making choices and points to the degree of effort required to model the equivalent choices made by the human. The development of that portion of the model, which will allow for a choice between alternatives, could gain much impetus by studying the results obtained from a transcoding study of non-native texts as suggested earlier as well as of statistics and micro-glossaries (sometimes called topical glossaries).

"Structural alterations", the third criterion, indicates, by the two scores obtained, that this is an area which is contentious whatever or whoever does the translation. It seems that the degree of alterations required to make the structures elegant is highly dependant on the revisor's view of what an elegant structure is. Nevertheless, the criterion is valid and even though reservations might be raised in certain cases of a structural alteration, whether of the MT or of the HT product, it must be admitted that MT will require more alterations for some time to come during its development. Therefore the criterion is a good indicator of the relative effort required to bring MT to an acceptable qualitative level. Also, the criterion can help to accumulate data on the

types of structural alterations most often required and so, to indicate where particular effort can be directed to improve the system.

The fourth criterion on "words alterations" as well as the fifth one on "mistranslations" only indicate a disagreement between the human revisor and the human translator selection of a word translation or the human who provided the initial equivalence to the machine dictionary. Therefore, as such, these two criteria contribute nothing to the comparison of MT to HT, per se, since the discrepancy is human. However, it is useful to retain the criteria as relative indicators of the seriousness of the disagreement between MT or HT and the revisor, but the analyser of these results should guard against assigning an equal weighing factor to them in his judgement of the relative merit of MT versus HT.

The sixth criterion on "idiomatic conformity" points to a wide disparity between MT and HT, at least in this case. This can be explained by considering that the MT system in question has really not addressed the idiomatic problem in the concerted way that the human translator of the TR-2 level would have and as such the MT system is apt to close this separation very drastically as soon as it is provided with a model of the HT idiomatic experience. In SYSTRAN, for example, this model is included in an idiomatic table look-up which permits the recognition of idiomatic expressions of up to thirteen words with the quality of the table only dependent on the degree of effort devoted to its build-up by those responsible for this portion of the data base. Therefore, as such, this criterion serves as an indicator of the relative success achieved in developing the correct data base compared to the experience of the human translator.

In the case of the seventh criterion, "time required for post editing" and of the eight on "working rate", it must be recognized that a large portion of the time spent on MT revision by the post-editor was also expanded for HT revision but by the human translator himself. For example the revisor must have devoted a considerable effort in choosing among the 817 alternatives provided by MT while an equivalent effort must have been required by the human translator in choosing among alternatives prior to presenting his product to the revisor. The point to be recognized here then, is that for the sake of equity in arriving at an objective judgement this should be based on an equal basis of comparison. So, it would be necessary to include the time required to translate (which was, for this test, MT: less than 1 min; and HT: greater than 21 hrs) to account for these hidden facts.

In conclusion on the development of a set of criteria to objectively compare MT and HT it can be said that the set provided at the seminar was a useful start at building such a table. It fell far short of being sufficiently comprehensive and conclusive to form a judgement on that basis alone. As the set is expanded to include more relevant factors and eventually to serve to indicate how far MT has to reach to achieve the high level of quality translation obtained from an experienced human translator, the comparison should also be made against the product of an inexperienced translator, or even a bilingual person with no training in translation, to establish how improved MT is becoming relative to the low level of translation

obtained from this latter subject. The reasoning here is that, it seems just as valuable to realize how far MT has advanced as it is to realize how far it has still to go. Overall then the need for and value of a set of criteria to compare MT and HT can be accepted as validated.

CONCLUSIONS AND RECOMMENDATIONS

The lack of extensive and positive results obtained from MT to date may be attributed to two basic reasons. One was the let down which occurred when the over optimistic predictions, of the early sixties, failed to materialize. These predictions were unfortunately based on a grossly incomplete appreciation of the magnitude of the linguistic problem by the computer specialists of the day and of the limitations of the computers by the linguists. The "brute-force" approach, as Professor Garvin calls it, failed as should have been expected. This failure propelled the pendulum to swing the full half-cycle in the opposite direction, leading to the over pessimistic group of "perfectionists" who required a total theoretical solution to the linguistic problem before any hope of obtaining practical results could be entertained. The supporting funded research was thus directed to address the theoretical linguistic problem with full vigour. Again this attitude seems to have failed because even though some interesting hypothesis have been proposed, as a result of this concerted research effort, we now find that no instrument exists to test these theories in practice. The prime danger which, in this writer's view, the field in Canada is facing today, appears to be that since we have not really gone through both of the cycles as described above, attitudes may develop such that research is forced to go through these cycles before the more mature middle-of-the-ground position is adopted. This would be a most unfortunate and costly dead-ended avenue to follow. Paraphrasing what Professor Garvin suggested, in "Machine Translation - Fact or Fancy" in Datamation, April 1967, the reasonable middle ground lies between the task-oriented, awe-inspiring "theoretical research couched in dazzling quasi-mathematical symbolism" and "the naïveté of the (original) brute-force approach".

It would seem then, that what needs to be done in Canada would be to continue to support the two Universities in their research, as in the recent past or even on an increased basis. It would be advisable also to relieve the Universities of the pressures of having to produce immediate practical results. They should however be provided with the necessary instrument to test the theories they uncover or advance. This can be achieved by entering into a developmental operational MT project. The way in which this should be approached was suggested in the practical considerations. It should be a simple matter to investigate the current availability of operational systems since these appear to be restricted to only three, the LOGOS, the SYSTRAN and the Oak Ridge systems. If no other systems can be uncovered then, the best of these should be adopted with the definition of "best" being directly related to such concepts as: machine independence, open-endedness, improvability, versatility, universality potential, ease of programming language, etc. The implementation of an operational MT system, on that basis, would seem to offer the best chances: of providing the instrument required by the pure research facilities to test their hypothesis; of advancing a modest beginning, which should be cautious at

first but determined; and, of eventually hoping to have a practical system capable of producing useful results.

Initially, the results obtained should be opposed to those obtained from an equivalently inexperienced human translator (e.g. a child learning a second language or a bilingual person with little or no translation training). The comparison of MT to the translation obtained from an experienced translator is a valid long range activity. However, the comparison should be made against more than one human translator to avoid a varying base or upper reference against which MT is measured. The conclusions drawn from such a comparison should be strictly reserved to exhibiting the degree of improvement of the MT system since the previous assessment. They should not attempt to establish the degree of acceptability, or otherwise, of the product of MT as this is better left to the actual user of the MT output.

Canada, as a bilingual country, with English and French as its two official languages, has much to gain from the results which may be obtained from such an activity but it also has the responsibility and opportunity to contribute to the world store of knowledge by an active and determined participation in the field.

The author of this paper, André R. Gouin, may be reached at :
<mailto:argo@moncanoe.com>.

REFERENCES

(1) Lebmann W.P., and R. Stachowitz, *"Feasibility Study on Fully Automatic High Quality Translation"* University of Texas. RADC-TR-71-295, Vols 1 & II. December 1971. 576 p.

(2) Toma, P., L.A. Koslik, and D. Perwin. *"SYSTRAN Machine Translation".* RADC-TR-70-168. September 1970. 146p. (AD 875290).

(3) Gouin A.R. *"French-to-English Machine Translation System based upon Digital Computer Software Programs (SYSTRAN)".* USAFIT Master's Thesis, March 1970. 278p. (Available as DDC document).

(4) Aldridge L. *"French-English Machine Translation - Some Preliminaries and Lexical Routine Development".* USAFIT Master's Thesis. December 1971. 116D. (Available as DDC document).