

MACHINE TRANSLATION SUMMIT

September 17-19, 1987 HAKONE Prince Hotel, Japan

Concluding remarks by Loll Rolling

Ladies and Gentlemen,

In his keynote speech, John Hutchins presented a perfect survey of MT today. During the conference, it was completed by a range of factual informations and by a number of declarations of opinion, both of which have enriched our vision. I do not want to come back on all these, but I wish to emphasize a number of topics that were only superficially discussed and which, in my opinion, deserve some more consideration and comments:

1. MT funding
2. MT design
3. Market studies
4. Text and script
5. Acceptance
6. Evaluation criteria
7. The new profession
8. Future development
9. Role of MT summits.

But before going into the details I want to clarify something. I do not necessarily disagree with Martin Kay but I would suggest a different definition of our objectives.

The objective of M.T., in my opinion, is to produce translations, for someone who needs them and who is willing to pay, either in order to receive a quicker, a cheaper or a better translation than human translation can supply. You will probably agree that MT is quicker and cheaper, but maybe not better, and therefore I would recall what was said in some of the previous papers about higher consistency of terminology in MT and about a number of typical errors in human translation. I personally have seen a human translation which conveyed the wrong message simply because the translator had forgotten the word "not".

Today people do invest in their translation and they do pay to obtain text translation by machine; if they do so, they must have studied the market and compared the cost, the quality and the timeliness of the translations they can get from the machine. Therefore we should not say that machine translation serves no useful purpose and that further basic research is indispensable before one can think of using it. I can say that at the Commission we have shown that for post-edited Systran translation of high translation quality, the cost is about 50% of human translation, the productivity about 300% of a human translation and the speed about 500% of that of human translation.

MT FUNDING

Most of the funding up to now has been made by government agencies and intergovernmental organizations such as the U.S. Air Force and the Pan American Health Organization.

The second type of sponsors are universities. You have the famous examples of the Georgetown University and 12 other universities in the U.S. which had their funding stopped after ALPAC. Other examples of university funding are Grenoble, Saarbrücken, Kyoto, Provo and Montreal.

But government agencies and universities funded these machine translation systems for their own use, not for somebody else; they wanted to make economies of scale in their own translation departments and expected linguistic research to yield fallout for didactic purposes.

Another type of funding is supplied by the hardware industry; this is what is happening nowadays in Japan but it has been going on since 1954, when IBM started its first MT initiative.

Siemens and Philips are also in the business today. Thus 90% of the funding comes from companies and institutions which have mainly developed MT systems for their own use.

A very small exception to this is the development of machine translation systems by entrepreneurs, that is by individuals who tried to develop systems in order to sell licenses or provide translations to individuals or companies.

Three examples of this are Peter Toma who developed the Systran system, John Smart who initiated the Smart system and Bernard Scott who started the Logos system on their own.

For entrepreneurs, machine translation involves high risk. I know a number of entrepreneurs who started out and lost a lot of money. It is also a low profit enterprise, as investments can only be recovered by large volumes of translation. The market studies are sometimes very misleading. I am sure that Mr. Scott made a very good market study but he forgot that the Vietnamese war would have an end and that the Shah of Iran maybe would not be there forever. So he had to adapt his system to a number of other languages. Someone who develops machine translation must not forget that sales and promotion are expensive and that customer service is even more costly.

MT DESIGN

Economic efficiency in machine translation is more important than scientific theory or orthodoxy. I also think that practical experience does more for the improvement of machine translation than basic research and here I have seen my opinions shared by Mr. Fuchi. I also think, like Joann Ryan, that sound computing is more important than linguistic perfection. At the Commission and also at other places where machine translation developments are undertaken, we can say that we need an average of 9 linguists for 1 programmer. This has led to many false interpretations: we do not think linguistics is more important than computing; on the contrary, linguistic perfection leads to economic disaster.

I will not go into the differences of opinion on the subject of pivot languages, interfaces or interlingua, but I think that basic research is difficult to manage. The objective of machine translation is definitely not to replace a translator. If it can obtain the same result more quickly, it should do so and not try to imitate the workings of the human brain.

MARKET STUDIES

Market studies are very important for people who are developing machine translation, but they concentrate on the presently existing human translation market. Institutions and companies use human translators for translating their correspondence, their reports and their documentation. This market is rather limited. We made a translation market study in 1981 which is no longer valid, but it is useful as a model for further market studies. Still it is important not to concentrate only on the existing market because it is conditioned by the limited number of translators available. Now that machine translation is developing, we have noticed that there is a latent market and that new markets are now coming up because of the relatively low cost and the high speed of machine translation. This market covers especially databases, not the full translation of databases but material that is retrieved out of databases. And M.T. will be extensively used now by companies in the export industry for trade literature, and also for cover-to-cover translation of journals. The evolution in the next 5 years will be rather explosive. This latent market will be covered by machine translation.

And there is a third part to this market, which will come into being in the next few years because of the progressive integration of M.T. into communication networks which are just starting to be used internationally. We already have Teletext, Videotext and electronic document delivery, but there is a lot to be done to integrate machine translation into communication channels; the first example was the Minitel implementation by Gachot in France.

Market studies do not need to be subdivided by text types and subject fields; if you have developed a machine translation system that covers certain text types and subject fields, the market study should of course cover them. But on the basis of my own experience, I would say that of all the text types existing in the Commission a system like Systran can cover up to 85%; only 15% are not suitable for translation by Systran.

Coverage of subject fields is even higher, around 95%. But individual subject fields can have a very large vocabulary and that is cumbersome to introduce into M.T. systems. Also the coverage by language couples is important. We now have a coverage of about 35% of the texts translated for the Commission and will soon reach about 55% due to the current development of new language pairs.

TEXT AND SCRIPT

There are processing problems which deserve mentioning here. In Europe we have an extensive problem of character coding. When an English translator buys a text processor and connects it to a machine translation computer, he enters an English text and he will get a French text back. He finds then that the word processor he bought from a British company cannot handle French accents, not to speak of German umlauts and the Spanish tilde. Our findings are that of the text processors and mini-computers available in Europe now, the great majority are incompatible as far as character codes are concerned, and something has to be done about this urgently. Now, at the end of the conference, we cannot expect any solution to the problem, but I am confident that some commercial computer companies will solve it. The problem of character conversion also exists in one and the same language; the Kana-Kanji conversion is a good example and the Chinese language has comparable problems. My question is whether all JEIDA member companies are using compatible schemes for Kana-Kanji conversion.

The OCR problem has practically been solved in Europe by introducing learning systems so that whatever is read by the machine is shown on the screen. The human operator can make corrections, the machine learns from these, and the errors will not occur any more. I am confident that the Japanese will also very soon produce optical character readers for their script, together with appropriate learning systems.

My experience is, and I would like to have your reactions, that in a relatively advanced system, 90% of the errors found in the translated text are due to errors in input, spelling errors, punctuation errors, grammatical errors and what I call formatting excentricities. To this problem there are a number of fundamentally different approaches. I would just list them and give you my preference :

1. Guiding and training of authors so that they will not produce errors of syntax. This is not very effective, since it does not eliminate errors added by the typist.
2. Pre-editing, before a text is submitted to the machine, is relatively expensive and does not always produce positive results.
3. Spelling checks and grammar checks can improve the text to a certain level, but the editing is not complete.
4. M.T. programs and dictionaries can be adapted in order to cope with the errors which are most likely to occur.
5. Post-editing, in my opinion, is presently the best way to cope with these problems.
6. Another way to solve problems is interactive input. After the OCR you can run programs to identify unknown words which can then be corrected semi-automatically.

ACCEPTANCE AND PROMOTION

You know that quite a number of people in different positions have negative attitudes to M.T. in general. I am not referring to linguists: they feel that their science is threatened. I would rather speak of the translators who feel that they might lose their jobs; I think it is rather easy to convince them that, since the market is expanding, those who have a literary training can remain in their jobs and those who wish to adapt to new working conditions can very well find employment. Translators who have negative attitudes can be persuaded that it is a noble way to enhance their profession to help improve machine translation systems. Negative attitudes can also be found among the authorities who have to decide about the use of machine translation systems in their organization. Very often the translators are willing, but the authorities are not, because it would disturb the operation of a department which has been working well for many years. Surprisingly, sometimes the head of the computer center doesn't like the idea of devoting computing time to machine translation. Post-editors sometimes feel their linguistic skills are not used as they might; the solution is obviously to take subject specialists with good knowledge of the target language to do the post-editing.

For laymen who have never heard about machine translation, we can make demonstrations and hand out samples to convince them. The most important people to convince are the customers, the end-users of machine translation. What they tend to criticize is the complexity of access to machine translation and there, of course, the promoters of M.T should make sure that the service unit provides the best possible human interface.

EVALUATION CRITERIA

In 1978 the Commission financed a study report about evaluation criteria for translation, which is still available from the E.C. There are three main criteria: cost, speed and quality. Cost and speed do not need to be explained. Quality is the critical criterion because of the subjectivity of quality rating. Several criteria have been proposed.

- Intelligibility: that was in the earlier ages when M.T. output was of low quality. The rating depended on users' knowhow and attitudes.

- Accuracy or correctness of terminology: the rating depends on whether the evaluator is a subject specialist or a translator.

- Revision rate: this is, for example, in a 100-word page, the number of words to be revised in order to get a correct translation. In a very primitive M.T. system this may be 30-40%. By bringing up the dictionaries to size you can easily reach the minimum level of 85-90% of correctness which corresponds to 10-15% of words needing revision. This is economically viable.

Another criterion proposed was revision time. If the revisor spends 80% of the time that it would have taken him to translate the document, obviously the translation is not good. If it takes only 20% of the time, then the economic benefits are evident.

But quality criteria are subjective; my suggestion is that an evaluation cannot be done by one single person, but must represent the statistical average of the assessments by several persons.

The differences between evaluations by several individuals shows the value or validity of the evaluation figures. Evaluation results must absolutely be accompanied by an indication of what text types have been evaluated, of what the subject areas in the text were, and about the average length of sentences; long sentences are indeed more difficult to get through M.T. than short ones.

THE NEW PROFESSIONS

A number of new professions have come into being as a result of M.T.

M.T. developers have to have a certain skill of management in addition to their basic knowledge of computing and languages. Remember that it is easier to teach someone who knows the languages to learn a little programming than to teach a programmer the intricacies of language. More numerous are the lexicographers, who introduce lexical data into computer dictionaries. Linguistic programming is often handled by translators who have learned it on the spot.

Pre-editors need to have a good knowledge of the source language and apply the recommendations of the authors' guide, but I do not think that pre-editing will be an important new profession.

A really new and important profession will be that of the post-editor, and as I said before, one should use subject specialists who have not the same attitudes as translators; they do not commit some of the errors that trained literary translators are inclined to make and they do not waste a lot of time on minor stylistic variations.

These literary translators have missed the train. If they are not able and willing to change their attitude, they will be replaced by subject specialists in this new profession.

Another very important new profession is the marketing of M.T. We must not underestimate the need for promotion, training and after-sales service.

FUTURE DEVELOPMENT

A lot was said on future development during this conference. Of course, there will be more research into theoretical linguistics, which will take place whether we like it or not. I do not think it will be very useful in the short term, but it might help to increase the efficiency and accuracy of M.T. in the medium term. I do not share Martin Kay's opinion about ALPAC. It must be remembered that ALPAC I was directed against American universities who used government money for M.T. research. So if there is an Alpac II in Europe, it would not be directed against commercial M.T., but rather against M.T. research in the universities as Alpac I was.

There will be more research into programming languages, and that is very important because I think the present use of Prolog and Lisp is not the best solution. Our experience with Systran has shown that programming in Assembler macro language was much more efficient in terms of computer time than the use of Lisp or Prolog as they are. But I do think that by relying on the requirements of the user one can do very useful things in the future. There will be a lot of improvements to M.T. systems by means of disambiguation routines, as a result of feedback and lexical developments. Also very important, will be the automatic enhancement of M.T. by learning systems. Mr. Boitet underlined the usefulness of establishing multilingual corpuses and parsers to extract equivalent terms and phrases in many languages for this purpose.

ROLE OF M.T. SUMMITS

We all experienced a very useful exchange of information on all aspects of M.T. and this will possibly lead to a joint monitoring of some of these activities. M.T. funding may benefit from the prestige of the M.T. Summit, and also because those who are raising funds now know what is the best way to justify funding and to invest.

Another outcome of the M.T. Summit will be the recognition that it will be very useful and advisable in the future to pool resources, especially lexicons and text corpuses. Another result of M.T. Summits may be the joint promotion of M.T. based on increased market knowledge and universally accepted evaluation criteria. I express the wish that this be the first of a long series of M.T. Summits to be organised in Japan, in America and in Europe. And I wish that in the future we can bring together just as many qualified participants and above all that future conferences will be as cheerful as this one has been.