

Parallel session III Machine translation

Summary of discussion

*John Hutchins, Monique L 'Huillier and
Brian McCluskey*

To start off the discussion, Linn Downs was invited to give a shortened version of her paper on 'Selecting a machine translation software vendor', reproduced as the previous chapter. The first question from the audience - 'Which system did you choose in the end?' - was unfortunately one which she preferred to answer in private.

As a supplement to Peter Wheeler's update on machine translation earlier that day, Veronica Lawson, the discussion leader, gave some details received from Smart Communications, Inc. (New York). Smart, better known for its pre-editing software than its Translator program, based its approach to translation on 'cleaning up' the input, using the company's Plain English Program, to reduce ambiguity and post-editing. The subsequent translations were said to be considered fast and functional. Some 3,000 clients were using the Smart Translator daily, the largest group being Canada's Ministry of Employment and Immigration. In 1985 Smart hoped to market its software in Europe on UNIX-based hardware (the 'Translator's Workstation'), microchip technology having now put the power of Systran (or the Georgetown system) on a desktop.

More MT service bureaux were opening. The Commission of the European Communities (CEC) had called for tenders for bureaux using its Systran program. In Luxembourg, the European Centre for Automatic Translation offered public bodies Systran translations of texts presented in machine-readable form, on paper or via a telecommunications link; clients could specify whether they wanted 'raw' MT output or post-edited versions. In France, Gachot S.A., again offering Systran, expected to concentrate mainly on

serving small translation offices. Service bureaux enable organisations to profit from good-quality translations at low cost, with no large investment in computer or word processor equipment, and the most compelling reason for using them was speed. It was pointed out that clients need not contact the service bureaux direct - they can still work through familiar translation agencies; the agencies have recourse to MT versions whenever they think these are suitable.

A major problem with MT is the motivation of translators and end-users to make use of MT systems, i.e. to induce clients to accept a different kind of product, to accept on occasion some lower-level translations, and for translators to learn the unfamiliar skills of post-editing. Translators are trained to aim for high-quality translations, not to correct poor-quality ones. It is often better to hire people specifically as post-editors and not as translators, as is the policy at Burroughs, staff being appointed primarily as post-editors; however, as 25 per cent of all translations are still done manually, they are not sitting at terminals 'cleaning up' MT output all day.

There were now great differences in attitudes of translators to MT. Ralph Hawes of Logos found that attitudes ranged from complete acceptance of MT to determination to beat or fight the system; the differences had nothing to do with quality or age of the translators. He believed four factors made translators good MT editors: translating skills, familiarity with the subject, facility with word processors, and motivation. The chief of these was motivation - the mental attitude of the person concerned towards MT, and the desire to increase his or her own efficiency. Differences in attitude could account for wide ranges in productivity, from 5 pages a day to 25 pages a day.

A speaker from a United Nations agency commented that all translations have to satisfy three criteria: speed, cost and quality. Organisations with, say, 100 translators wanted, in particular, a reasonable level of accuracy at a reasonable level of cost; MT would be judged on how it met these requirements.

For a speaker from the English Translation Division of the CEC, Luxembourg, the great merit of MT was that it offered the prospect of a wide range of quality in translations, where users could decide the level suitable for their purposes. The natural inclination of human translators was to provide as high a quality as possible, the client having no say in what the standard should be. The best-quality translation (as Veronica Lawson pointed out) was not always best for the client. Consequently, in addition to its standard (full) post-editing service, the CEC now offered

'rapid post-editing': cursory editing of only the most glaring errors. Post-editors could improve the low-quality MT output to different degrees, according to the needs of the clients. MT had opened up a previously unrecognised market for low-quality translations; just as there were markets for the whole range of original texts (from popular journalism to academic monographs), there were markets for a whole range of translations. The major problem was the psychological block of translators and agencies, a reluctance to send out something of lower quality than they were capable of. However, if translators found that they were providing a useful service to the customer by offering translations of lower quality by using MT, they would generally be prepared to continue providing the service, regardless of their own views as to the intrinsic quality of the translations produced.

Stephen Kerce of Weidner Communications Corporation (WCC) reported on the Weidner service bureau, which produced one million words of MT-generated finished translations a month. The conviction of WCC was that clients wanted human-quality translations of good standard, and WCC therefore post-edited all texts, providing high-quality translations, and made a profit.

Linn Downs underlined the cost-effectiveness of MT in Burroughs Corporation (Detroit, Michigan), an organisation with a throughput of 21 million words a year. At Burroughs they preferred their writers to keep to a uniform terminology, but there were no restrictions on style.

Ulla Magnusson-Murray reported on the ITT experience as a MT service bureau, using the Weidner system for translations from English into French, Spanish and German, and from those languages into English. Having acted as a translation bureau before purchasing the MT system, ITT already had a nucleus of translators. In the last year ITT had both doubled its output and improved quality, primarily through the integration of the word processing, publishing, typesetting and electronic input systems, which meant that documents could pass from stage to stage with no rekeying. Now the only human intervention is in the post-editing phase. The amount of revision in the Weidner system depended on the quality of source language texts, the terminological difficulties of the subject and the experience of the translator; in general, least revision appeared to be necessary for English-Spanish translation (perhaps 70 per cent was acceptable in 'raw' MT output), but a lot of terminology was needed to reach this level.

In response to a comment about the slow speed of Weidner, as demonstrated at the conference on the IBM PC microcomputer, representatives of ITT and WCC both stressed that their bureaux used the Weidner system on

mainframe computers with much faster results; in both cases the MT system was run overnight, with post-editing being done off-line during the day.

References to the 'acceptability' of translations gave rise to a number of comments: did this not depend on how low or high your standards were? Did '70 per cent acceptability' mean that 30 per cent was gibberish or just of lower quality?

A speaker from the English Translation Division, CEC, Luxembourg, questioned the value of expressing accuracy or acceptability as a percentage, since so much depended on the recipient's knowledge of the source language and target language, and on whether 'intelligibility' alone was sufficient or good quality was essential for the client's purpose.

Merle Tenney (ALPS) made the point that ALPS was the only fully interactive MT system. Interaction had the advantages that additions and changes of vocabulary could be entered into the dictionary whenever the need arose and that particular usages could be specified for specific texts. (For instance, only one decision would be needed to translate 'you' each time as 'tu' rather than 'vous'.) He stressed that comparisons of the speeds of MT and human translation could be misleading if only dictation speeds were included for human translation, and not typing time; in general all translations were faster with machine aids.

It was argued that MT systems should be capable of absorbing the local house vocabulary of the organisation. Above all, MT systems should be accurate, although it was pointed out that it is rare to find two translators in agreement on the question of how accurate a particular translation is.

Questions were asked about the problems of post-editing: were they points of lexis, grammar or style? Of course, it depended on the source text, on the purpose of the translation, and on the MT system. In Veronica Lawson's experience of Systran, the problems were primarily lexical and semantic; there were relatively few mis-parsings and the system was better than others on anaphora, although articles, for example, still presented problems. In the CEC, feedback from post-editors to those in charge of the program had helped greatly in improving quality. Dr Habermann (Kernforschungszentrum Karlsruhe) reported on a favourable experience with Systran: 500 pages translated in half an hour, and no post-editing necessary at all.

Some speakers worried about lower standards and thought that translators ought to fight for higher-quality MT. One of the speakers from the CEC countered that in the new conditions translators would have to adopt lower standards of pride, with no prejudices about quality levels, and with different attitudes to perfectionism. The market forces of MT would determine standards as well as prices.

It was argued by one speaker that even if the source language text is poorly written, clients expect the translation to be well written; translators have to 'correct' texts. (It was mentioned at one point that poorly-written texts were frequently those written by non-native speakers, e.g. Germans writing in French.) Dr Habermann believed that users are more content with low-quality texts than translators are (quoting an example of Systran output which a user found acceptable but a post-editor did not). The natural inclination of translators, Veronica Lawson thought, was to seek the best possible version. Another speaker, however, held that if the source language text had been poorly written then the target language text should be also: GIGO (garbage in, garbage out). In this respect, a speaker from King Saud University saw a drawback with systems like Smart which required texts to be unambiguous or which required pre-editors to establish authors' intentions. The need was for systems which could translate what was given.

The experience of one speaker with MT texts requiring revision was that if texts were intended for general distribution, it was often quicker to do the translation manually.

Thomas Pahl, who as Dr Toma's programmer had worked on Systran from the beginning, believed that the real value of MT was that it allowed texts to be translated for information purposes which could not otherwise be translated at all. MT systems had no problems with quantity; the bottleneck was the input stage. The problems of input were indeed raised by a number of speakers. Manual input was costly, although one speaker thought the problems were exaggerated: there was always a secretary to input text! Another pointed out that the obvious answer was the optical character reader, but unfortunately very few optical character readers would cope with diacritical marks, let alone non-Roman alphabets. A speaker from King Saud University believed that the technology was available; what was lacking was the money to develop the readers needed for MT systems.

A representative of Xerox spoke as a major user of Systran (if Peter Wheeler's figures were accurate, perhaps 10 per cent of total!). One advantage of MT was that it reproduced the original source language text structure unchanged; it did not impose other standards. The quality of the final text depended on the abilities of the post-editors; MT was a tool, a means not an end. The requirement for post-editors was, above all, technical understanding; language skills were secondary - the user read to understand. For instance the best person to translate an IBM PC manual was someone who understood the IBM PC. The quality of the target language text was of

secondary importance in the market place (as the Japanese had long understood). Commercial pressures meant that MT was becoming essential if information was to be put where it was needed. However, Ulla Magnusson-Murray disagreed: language skills had to go hand in hand with technical knowledge.

A speaker from NATO thought that a major problem for MT was that different languages reflected different conceptualisations of the same objects and events; translating sentence by sentence does not work, and for this reason straight MT cannot adjust source language texts to the variety of target language environments. Veronica Lawson stressed that such problems often arose, even in technical texts such as standards (different countries might standardise quite different aspects of articles) and instructions; she recalled a cartoon in the popular magazine TV Times in which a husband and wife had been trying to get a washing machine to work: 'Now let's try the Italian instructions.'

A member of the Eurotra team at the University of Copenhagen was invited to report on the development of the Eurotra project. This, he said, was designed as a support to research in computational linguistics and as such was already a success. He believed a major achievement was the establishment of successful collaborative research in ten European countries, although whether collaboration would continue once a productive system was close was somewhat doubtful. Final contracts had only just been signed; now genuine national teams could be set up, where before only provisional project teams had been operating. Initially only previously analysed texts would be experimented with, but the aim was that by 1988 the experimental system should be capable of translating unanalysed texts within a specific limited field between all the language pairs of the European Community. Eurotra should be operational in the early 1990s. Loll Rolling of the CEC added that a team of six translators had now been seconded by the Commission to co-ordinate Eurotra research in Luxembourg. It was questioned why Systran would not be sufficient, i.e. why was Eurotra needed? The answer was that a genuine multilingual system brought many advantages for translating between 'minor' language pairs, e.g. Danish and Dutch, where there might well be shortages of competent translators, and for translating simultaneously from one to many European languages. On the other hand, there would clearly still be a role for Systran in the case of one-to-one translations.

A speaker from Logos said that the 'one-to-many' capabilities of the Logos system were under test; methods of source language analysis devised for one particular target

language were largely (about 80 per cent) applicable to another target language or a series of target languages.

These remarks prompted a speaker from Saudi Arabia to comment on the present limitations of MT systems arising from their concentration on Western European languages. When systems did incorporate Japanese, Arabic, Hebrew, etc, it was often as a by-product of methods developed for languages of quite different structure. Methods of source language analysis which worked for European target languages were not so successful for non-European target languages. He did not believe that universal MT systems were possible; MT systems devised for specific language pairs were likely to be more successful.

The question of post-editing was raised again at the end of the session. Translators had professional scruples about post-editing; whether they took to it depended on their attitude to word processors and the like. It was suggested that a distinction should be made between rapid and standard post-editing. Even with human translation, revisers tended to revise too heavily at first; they had to learn to edit without rewriting. In the CEC some revisers became very fast post-editors. It was recalled that the idea of 'rapid post-editing' of raw Systran output was proposed by the CEC translators themselves in recognition that not all users wanted 'purple prose'.

In response to an enquiry as to whether job descriptions for post-editors had been compiled, Linn Downs replied that Burroughs had done one, as had WCC, but she was not able to give a summary of it. Merle Tenney (ALPS) thought that post-editors ought to be good translators even though the jobs were not the same. It was reported that Bernard Vauquois (Grenoble University) had drawn up profiles of linguists, translators and post-editors. Training courses for post-editors had started in 1983. At the CEC in Luxembourg, where post-editors were expected to aim at four pages an hour (usually minutes of meetings), with a minimum of two, it had been found that 'rapid post-editing' demanded great skill: good subject knowledge, word processing ability and willingness to adapt to the MT tool. Experience and confidence in one's ability seemed to be very important. In short, the best translators often made the best post-editors.