[*Terminologie et Traduction*, no.1, 1986]

MACHINE TRANSLATION IN BUREAU SERVICE

José Mendez
Mendez Translations, Brussels.

For the past thirty years or so, linguists, computer scientists and translators have been researching and developing machine-aided translation systems. We all believe, by our presence in this hall, that Systran offers the most potential. And by granting user rights to bureaux serving the private sector, and as a result of the advent of modern telecommunications, Systran is opening up the market - offering fascinating new applications for industry and commerce.

Undoubtedly, the service bureau's single largest potential market is that of multilingual documentation production. Mendez is already providing several multinational companies with a complete documentation service package - from translation through to printing and distribution. In terms of cost, our experience has shown that translation only constitutes a small percentage of the overall budget. As far as technical brochures are concerned, translation represents, on average, a mere 7% of the total production costs.

Using normal translating methods, i.e. without MT, we have found that production time is divided roughly as follows: 35% is spent in translation, client revision requires 15%, typesetting and studio work both take up 20% and the final 10% is spent in printing and distribution. In a traditional environment, therefore, translation is one of the most cost-effective processes but also the slowest. Since translation costs are low, the major considerations for a bureau or inhouse department are productivity and quality of performance. Quality of translation being, of course, the most difficult element to control since, left to the human translator, so much depends on subjective analysis.

Our client's profile, and consequently his needs, have also changed over the past few years. Stiff competition and rapid communications have forced commercial companies, operating internationally, to launch a new product simultaneously throughout the world market, along with relevant back-up material (for both internal and external use), including multilingual sales and advertising literature, brochures, reference manuals etc.

Delays in marketing a new product in one particular market could cost the company irreparable damage, especially if its market lead over competition in that particular area is marginal. Speed and accuracy are the name of the game.

The criteria have changed: to guarantee a constant standard and improve overall quality, the industry has been forced to adopt a more scientific approach to translation. This requires an investment in the preparatory stage to avoid unnecessary delays and increase accuracy as well as to integrate machine-aided translation in the working environment, thus improving output. Modern technology has advanced so far that machine-aided translation now incorporates both image and text processing. Machines exist which can fuse the product of the latter scanners and produce a colour page layout. Investment in such an integrated system can already improve a company's overall productivity by a factor of three. But the initial investment is high.

At Mendez we are in the process of developing just such a system. It is too early to give a precise cost analysis of the benefits of Systran to the service bureau but the advantages are demonstrated by the fact that our freelance translators, using Systran as an additional translating tool, actually approached us, proposing a reduction in their fee because the system rendered their task so much simpler and faster!

The complete service - pre-editing, raw translation and post-editing - is currently being sold by Mendez at the same cost as traditional translation methods. This has meant a slight reduction in our profit margin but a considerable improvement in productivity and accuracy of output. Experience so far at Mendez has also shown an evolution in the clients' translation requirements. Because of Systran's speed and scope, clients are now differentiating the levels of translation required. If only a quick rendering of sense is needed, very often a client will be satisfied with a raw machine translation. If the document being processed is to form part of a sales speech it requires fast post-editing. A text destined for a prestige sales brochure will require additional 'stylistic gloss'.

Systran's potential is limitless but its real value can only be enhanced through global improvement, regular input and with all its users speaking with one voice. A firm advocate of machine-aided translation and spokeswoman for the Pan American Health Regional Office (Washington DC), puts the current status of Systran into perspective:
"Machine translation systems require constant nursing".

The industry, it is true, is still in its relative infancy but gone are the days when translators debated on whether or not machine translation was possible. Automatic translation has undergone a qualitative change. The questions now being asked are: "Do we like what MT systems are giving us? Are MT products economical compared with traditional translating methods? or How can MT be improved in the future?"

Increasingly the role of the service bureau is that of consultant and interface between the translating system and user. We at Mendez, for example, are asked more and more by our clients to evaluate their linguistic and literature requirements. We are also asked to look into ways and means of rationalising the production of their documentation and provide feasibility studies for automating their office structure (linking for example, word processing, data processing and telecommunications).

Until recently, prohibitive costs meant that only a limited number of large companies and organisations were able to finance the ongoing development of an MT system. But now drastic reductions in software and hardware costs have provided a means of rationalising output, thus constituting a speedier and more reliable form of processing for routine translation jobs.

The marketplace is still looking for successful reference sites before it decides on committing itself to one particular system. But spiralling costs in terms of human translators, an inundation of highly specialised documentation requiring rapid processing into other target languages are just some of the reasons why more and more large international organisations are examining additional translation tools.

One essential fact remains however. Translation systems will never replace the human translator but will put an end to amateurism. After all, photography has never seriously threatened the artist, nor has the advent of records and the compact disc destroyed concert-going. If anything, the translator will be motivated to produce a better finished product as he, or she, is freed from certain tedious and time-consuming translation chores. He can also increase his linguistic versatility and specialise in new terminology and subject fields to a greater degree. His output will increase since he can input a text at the end of the working day, send it to for MT processing during the night and post-edit the following day.

It is true that the man-machine interface is not always going to be without problems but the human interface will always play an essential role in any future system. Perhaps the translator, as we traditionally know him, will change. Inevitably his role will become more that of post-editor or reviser.

We are entering the age of digital communications, where ease of information exchange, processing and transmission will be increasingly at our fingertips - be it at work or in the home. Breakthroughs in technology are happening so fast that it is difficult to keep abreast of changes. For this reason alone, translation will always represent a dynamic business growth area.

The advent of new terminology, the creation of new fields of specialisation and the corresponding acronyms, etc. will continue to provide professional translators with all the work they can handle. There may well be a shift in emphasis in the translator's method of working, but with the application of MT, linguists will be needed to update and create new lexicons, glossaries and develop syntactical and grammatical software.

This new breed of translator - combining linguistics, lexicology, translation and data processing - is already working at Mendez. We like to be called 'linguistic engineers'. Since last January, when we were given commercial rights to act as an official Systran service centre, our team of translators have successfully made the transition from traditional working methods to using MT as a means of boosting productivity and accuracy.

Whichever system is used, no one can expect 100% perfect translation. Language is, after all, a very subjective medium, shadowed by overtones, ambiguity, stylistic idiosyncrasies, etc. No translator can be perfect, any more than the copy he produces - especially when the latter is based on a poorly written original. An old translator's adage, 'the style is the man, for both an original writer and a translator, and very often the twain do not and could never meet' applies just as much to human translation as it does to machine-aided processing. (Personal likes or stylistic criteria, of course, are totally redundant for MT.) Likewise, a translating system, such as Systran, can only produce 'accurate' translations if the original text is '<u>prepared</u>' or written in a simple, unambiguous manner.

Consistent format, vocabulary and sentence construction in the source language are of paramount importance. That is why legal documents, standard forms, official circulars and technical specifications lend themselves, perhaps, more to machine-aided translation. If we are to use MT output efficiently and cost-effectively, we must first educate the client or author into writing coherent texts - free from ambiguity and anomalies which could confuse the system - unless we leave these to the post-editing stage.

This simpler form of writing acts as a type of control. Large corporations, who are satisfied Systran users, all resort to such a technique. Xerox processes a very narrow range of subjects and the source text is written in MCE - Multinational Customised English - with limited vocabulary and stringent writing rules. The Caterpillar giant has developed its own inhouse source language to harmonise and ease the translation of some 20,000 company publications into 50 languages - the Caterpillar Fundamental English.

To the layman the base language used may seem like standard English but it uses simplified sentence construction, coherent alphabetical listings, restricted vocabulary and is readily comprehended. Caterpillar claims that adjustment to its basic English takes a foreigner approximately 60 hours training.

But with all these multinational corporations - to the list must be added such names as IBM, ITT, Digital Equipment Corporation, Eastman Kodak, Ford Motor Company, Prime Computers Inc - developing their own base language surely there is a danger that the original MT datadictionary will migrate in different corners of the world, within the same industrial sector and possibly cause duplication of work - so that the original rationale behind creating a unified data dictionary is lost. This is one argument in favour of a full service bureau, such as ours, which can also double up as a machine translation service centre for a particular country. Continual feedback, updating of glossaries, regular consultation and harmonisation with different bureaux - accessing the same host - is one way of improving the system, making it more flexible, consistent and, last but not least, commercially viable.

One advantage that MT has over the traditional translator, which cannot be overlooked, is that humans do not like working under pressure - a situation which is aggravated when a person does not know what the purpose of the translation is. The fact that all input and output from MT systems is machine-readable and that the MT itself is available, on-line, 24 hours a day, 365 days a year, with immediate feedback, instant correction, revision and transmission capabilities it is the ideal solution to rush translations, with yesterday's deadline!

As I have mentioned before, experience at Mendez has shown that more often than not the client needs a fast turnaround (and rough translation as opposed to a more time-consuming, stylistically and grammatically polished rendering of the text). Client satisfaction is increased further if he is provided with a full package of related services. At Mendez we have decided to invest in this direction. We have capitalised on our translating experiences gained in a wide range of fields over the past twenty years and cashed into the future. We have developed from being a purely translation oriented bureau into a specialist in multi-lingual document production.

We employ some 20 full-time translators inhouse with a further 150 freelance translators working under contract. We have an additional staff of 40 working in our production department - ranging from photography, layout, phototypesetting, proofing, printing, EDP processing, packaging and distribution.

We can provide a full package from translation of the source document to the finished product. All the client has to do is sign off the final proof, provide the original input and help on any new technical glossaries.

We took the plunge three years ago when we set up Dataset SA in collaboration with the Antwerp-based printers, Graphica. Our investment has more than paid off. As far as machine-aided translation is concerned, Mendez first came into contact with the Systran system through our extensive contacts and work undertaken with the European Commission. Last January we signed a momentous agreement with Orda-B to become Belgium's official Systran service centre.

The transition to Systran went smoothly as we had already invested heavily in office automation equipment, including the Wang OIS system, which gives us access to the host Systran mainframe and clients whenever applicable. Often to save time our client - providing he has the necessary equipment - sends us the source text via electronic mail or modem link-up. If necessary, our translating staff can revise the text - to remove any possible ambiguities for the Systran system, or input any new vocabulary for the Systran dictionary. The document is then sent for batch processing and any post-editing can be carried out on screen at Mendez and then returned to the client either as hard copy or on-line for approval. If it is a question of printing a brochure - once we have received the sign-off, we can transmit directly to our printing unit for typesetting. We have found this working method to be most effective and rapid. The 'rush' job is no longer a problem.

Our belief in the Systran system lies in its scope, proven performance and future potential. The fact that Systran has been around for such a long time and that it has been sold by DP specialists rather than qualified linguists with DP competence, has made the market hesitant.

Now that Systran is available on the commercial market and its linguistic capabilities and dictionaries are being expanded by competent authorities, such as translating bureaux, the future looks very bright indeed.

At Mendez we first tested Systran's effectiveness by making it available to our permanent staff as an additional translating tool. We did this initially rather than try to sell a service we were relatively unfamiliar with. Our experience showed an improvement in the overall standard of work and profitability. We then carried out field trials with a cross-section of clients. At Mendez we have some 950 different companies on our books - whose scope is quite representative of market demand throughout Europe. The major sectors represented are: information science, automation, office equipment, mechanical construction, medicine, engineering, chemistry, international institutions, government agencies, publishing, hifi and video, ministries, telephony, federations, insurance, legal, oil and related industries, domestic appliances.

The trials, so far, have proved most successful. Of course Systran's potential is only as good as the glosses and dictionaries put into the system. Only by having access to a wide range of 'market' experience through service bureaux such as Mendez - who can provide the latest state-of-the-art linguistic and technical advice and who are in regular contact with the industrial and commercial fabric of society - can Systran develop into a more flexible and viable product.

The advantages of using a service bureau, offering Systran in conjunction with a variety of other services, as opposed to an in-house department working in isolation or at odds with its industrial contemporaries, are obvious. If the modern translating bureau has opted for the future, it will already have invested in modern communication and DP/WP equipment. The transition from such a working environment to the Systran interface does not come, therefore, as a drastic change. The public sector - be it on a national or international level - is far less geared to assimilating such a transition. The inherent hierarchy is less flexible, sometimes hostile to change, and very often lacking in the necessary DP or WP facilities.

The current non-availability of certain language pairs - French-Dutch, Dutch-French - on Systran is another reason why it is difficult to interest national institutions in Belgium in investing in such a system.

As far as international institutions are concerned, the language pairs currently available on Systran are of more interest, but here again we come across various elements that are hostile to innovation and change. Undoubtedly the best potential market for Systran service centres lies with multinational industrial concerns.

Only time will tell what form machine-aided translation will take in the future. Current developments seem to indicate that Systran may well be available on office computers and desktop PCs within the next couple of years, provided the Unix conversion goes ahead. With Systran running on Unix, the user would have access to a whole range of word processing packages but the major advantage would be text manoeuvrability. Since Unix was conceived originally for phototypesetting it offers a sophisticated means of transposing text and graphics (ideal for sorting, lists, pagination, creation of indexes, etc.) - a vital tool for final brochure production.

It is true that the industry is divided into two camps: one being the development of more user-friendly, cheaper desktop systems - offering a wide range of basic software/dictionary packages - which is easier to sell in the short-term but will lead to software migration and a disunited user base; and the second being the development of larger, more mature packages (Systran of course falls within this category) with centralised benefits for all its users - with the added advantage of constant expansion and diversification of common data dictionaries.

I firmly believe that this latter, unified approach to MT will have the best overall results. I also believe that both the private and public sectors should collaborate closely with service bureaux offering Systran (create even a Systran user association), thus encouraging more generalised access and diversified input to MT, until the day is reached when sophisticated voice analysis technology has put a computerised simultaneous interpreting system into the pocket of the man in the street.