

Electronic translation?

Barbara Snell

Language Services, Rank Xerox Ltd,
Gloucester

THE AGE OF MIRACLES is over, they say, attributing the wonders of today's world to technology. But if the advertisement for a 'Language Translator'¹ is to be believed, we must continue to accept marvels. With this personal pocket computer it is claimed that, 'You can communicate in a foreign country without speaking the language, or you can learn the language more easily thanks to a new electronic miracle.' We are further told that the gadget translates words and phrases into and out of six languages.

The makers of 'products that think' might have impressed me more had I not been a practising translator. This does not imply scepticism but a certain wariness. An article entitled 'Adventures in translation' supporting my suspicions appeared in *Business Traveller*² a month after the above advertisement. It described the trials of the proud owner of a pocket computer which flashed the puzzling message in German 'Make you to take credit playing cards?' when enquiring if he could pay for a meal with an American Express card. 'I want menu' was a little better though hardly likely to get the best out of the waiter; but a request to mind luggage—'Look at my paper bag'—met with a blank stare. Continuing his travels, the reporter tried Italian. The machine spelled out FARE IO PRENDERE CENA?, 'Do I take supper'; while in Spanish an enquiry for the price of a book came out as COMO MUCHO EL LIBRO?, 'I eat a lot the book!' The toy generated great interest but the practical aspect of typing out phrases, letter by letter, and obtaining dubious answers, compared unfavourably with even the worst of phrase books.

If this is the state of computer translation why are multinational companies and government organizations paying huge sums for machine translation projects in the hope of obtaining practical results? The answer must lie somewhere between such a pocket data bank, wrongly called a translator, and the automatic conversion of unedited texts from one natural language to another at a cost far exceeding the employment of human beings for the same task,

In the above examples the second request rendered as ICH WOLLEN SPEISE-KARTE (I want menu) is intelligible if abrupt and we can overlook the plural verb *wollen* instead of the first person singular form. The order to 'Watch my bag' which came out as BETRACHTEN MEIN TUETE (Look at my paper bag), presented another problem as the computer contains 'watch = see' and 'watch = clock'. A more sophisticated machine translation system would know to look for a verb, but when faced with a choice, this device flashes a series of question marks asking the user to go into the search mode and he can then select the right word. Even so, the result was scarcely satisfactory.

The Spanish example contains an ambiguity in the target language: *Como* means both 'how' and 'I eat' and in this sentence we need *cuanto* (how much) not *mucho*

(much). In a shop, *Cuanto el libro?* would suffice for most tourist requirements but to put the question in acceptable Spanish form 'How much costs the book?' either the computer or the user must know something of the language. In German the question would need to be 'What costs the book?' as 'how much' might come out as 'how very' (*wie sehr*) instead of the required *wie viel*. Reversing the process, a German inputting *wieviel* would be as likely to get 'how many' as 'how much' and, as any foreigner who has struggled to learn English can appreciate, it would take some programming to differentiate correctly between these two alternatives.

The first request, intended as 'Do you take credit cards?', appeared as MACHEN SIE NEHMEN KREDIT SPIELKARTEN? (Make you to take, etc.) because the user had not learned to ask questions without an auxiliary. In Italian we come nearer an equivalent for our 'do' with *fare*, but *prendere cena* suggests physically snatching the meal and their use of 'take' differs from ours; furthermore, the Italians would employ a verb *cenare* to eat supper.

As for Japanese, in which one can be ridiculous or even offensive by using wrong registers that involve different vocabularies for men and women, I doubt if the miracle machine will foster international relations except by creating novelty interest and drawing out conversations longer than before!

So it goes on—in other words, to use the pocket computer properly, the user needs to know the language it claims to teach before he starts. But someone who knows even only a little of the language is unlikely to find such basic information any more helpful than a pocket dictionary with its more straightforward access.

At the other end of the automatic translation spectrum we have SYSTRAN, a system operated by the US Government and Commission of the European Communities amongst others. Over half the budget of the European Parliament is swallowed up by translation costs and, given the equal validity of all member state languages, the accession of Greece, Spain and Portugal will strain the budget to a frightening degree by increasing the language pairs from thirty to seventy-two. An evaluation of the 'Systran Automatic Translation System' undertaken by Jacques Chaumier *et al* a few years ago³ indicating a future role for the system said '... that, in terms of economics, its (SYSTRAN's) performance—whilst also capable of improvement—allowed its use to be envisaged. ...' But leaving aside these examples at the extremes of the machine translation spectrum, how can staff and freelance translators benefit from electronic equipment now on the market to enable them not to become redundant but to work more effectively?

What do we mean by 'effectively'? We will not follow the red herring by asking if translators, or anyone else in this country, could work harder, we will look into the efficiency of the job. A commercial translation is a compromise with time. Given unlimited scope for reflection, second thoughts, revision, consultation and returning to the job with a fresh eye, we could in theory, produce the perfect translation. In fact, if you take faithful representation of the original in another language as 100 per cent success, a translation is often more than 100 per cent effective, being an improvement on the source text which may have been written in great haste or with less skill in handling words. Translators, however, in earning their daily bread, are constantly constrained to sacrifice quality for time. Accuracy is less likely to suffer as we know a mistake is a greater crime than a poor turn of phrase. We can also work more quickly if we keep close to the original, thereby reducing the risk of error, than

if we step back from the job and rethink whole sentences or paragraphs in the target language. So if equipment can help us to save time we should be no less accurate, but the results should be better. We cannot hope to compete with computers for speed but they will have to run hard to catch us on quality.

Let us examine some of the 'electronic miracles' which can save our time and reduce routine activities.

Facsimile transmission

The strange garbles spewed out by telex machines often turn simple English into ambiguous nonsense and play havoc with foreign language texts, providing translators with insoluble headaches. Even an impeccable rendering is without often vital clues such as capital letters and accents, and strange abbreviated jargon phrases easily become meaningless to the uninitiated. Worst of all, the source is suspect; one knows it was written by a machine and is likely to be wrong.

Using a telecopier or other form of facsimile transmission you receive a copy of an original free of the apparently random quirks of a machine and spared mutilation by typists working in a language they do not understand. The need for transliteration is removed—a boon for users of Cyrillic scripts, Arabic and the ideographic languages which can now be transmitted by telephone line. A short straightforward translation can be given within minutes to a client or distant branch of a company, together with drawings or supplementary information; similarly the layout of the text can be provided. Any particular legibility problems, such as those of a handwritten text, can be clarified quickly, verbally, at the time of transmission. There is a variety of machines on the market with transmission speeds down to one minute for an A4 page.

Word processors

Not all translations are revised or even rewritten, but few translators would not welcome the chance of amending their first versions on a screen. Even the most conscientious translator avoids searching for neater turns of phrase when this means retyping a 700-word page. Errors, when spotted, will of course be corrected; erasing fluids are useful but far from perfect and the results are usually a compromise between what one would like to see and the time one can afford to spend achieving it. A translation typed on a word processor is stored on card, tape or floppy disc and it can be displayed on a screen, typed out by the machine or transmitted to another word processor via a telephone line.

To the user a word processor is like a typewriter, but the typed text can be revised or edited as required and retyped by the machine. A disc contains up to 150 pages which can be played back on paper in any required format with automatic centring, spacing and justified right hand margin ready for printing. Revision features include the ability of the machine to change a word or phrase each time it occurs in the text; e.g., if 'smoke-grey matt finish' had been written throughout a contract and it was subsequently decided that 'grey high-gloss coating' was preferred, on command the word processor would change the phrase every time it occurred. With a communicating word processor your text can be transmitted in its final form at a speed of 120 characters per second to another communicating word processor at the end of a

telephone line anywhere in the world. Not only do the interchangeable print wheels allow the use of ten and twelve pitch, proportional spacing or different typefaces on the same machine, but there are different wheels for foreign languages. Linked to a code computer your word processor can act as a terminal for receiving or sending information. A translator with a word processor could therefore use it as a terminal to link up with a data bank. He could also have a store of 'memory cards' for his personal glossaries or they might be provided by a central source to encourage consistency of terminology (especially useful for a number of translators working on the same or similar jobs in different locations). The scope for such equipment as a back-up to translation is enormous.

Data banks

If we start with the assumption that a translator is more intelligent—or should we say discerning?—than a computer, it will be easy to understand that a data bank for the use of translators is not the same as one designed for computers.

Common words with a plurality of meanings provide the greatest volume of work for lexicographers and the greatest headaches for programmers of machine translations.

Let us take the word 'head' as an example. It is easy to see that with no context to indicate the required sense, the word could mean (Beachy) Head, the head of a firm, a head of hair, head a ball, head for trouble, a young head on old shoulders, a head for heights, lose one's head, and so on. *Harraps English-French Dictionary* devotes a two-column page to the eleven meanings of the noun, with as many again for the verb:

- tête*,
- esprit* (mind), 'it never entered my head';
- 'head of celery' *piéd de céleri*—surely more logical to say foot than head for the base?—or alternatively 'bring the matter to a head' *faire aboutir une affaire*,
- rubrique* (heading); —a selection of naval and geographical terms, 'to collide with a ship head-on' *aborder un navire par l'avant*, or *cap* (headland);
- the range of *chef* 'head of a family, company, etc.';
- pièces de gibier* (head of game); or *bon nombre* as a collective;
- face* (of a coin)—once again French logic wins;
- technical uses such as *colonne* (of water), *volant de vapeur* 'head of steam'; and others.

Differentiation between these French and English terms would present a mammoth task for the programmer but none of the uses is likely to confuse an experienced translator whose immediate recognition of the synonym for 'head' from the context and background indicates the sense, even if he is not familiar with the idiom.

It may be that the supplementary knowledge necessary for a computer programme to compete with the translator (this is the field of artificial intelligence) will be less than that which the man has gleaned on his road through life, but the human approaches his job as a ready programmed unit, aware that celery grows not upside down but from a club-shaped base and that footballers use their heads as well as their feet to propel balls.

Even the most experienced specialist in his field consults reference books at times

We are told that up to 60 per cent of a translator's time is spent on this, task⁶ and it would be sad indeed if the work never exceeded his readily available knowledge, depriving him of the stimulus of stretching his experience into new territory. When a system such as the European Community's EURODICAUTOM, described by Jacques Goetschalckx⁴ is widely available and more comprehensive, search time will be considerably reduced. The Siemens TEAM System⁵, which links text processing and editing systems with the final phototypesetting, takes the process still further, providing a sophisticated back-up to the human operator whose expertise it uses to the full by eliminating routine mechanical operations such as dictionary search and retyping.

A data bank somewhere is all very well but I need it at my fingertips, you may say. In 'Designing a multilingual terminology bank for US Translators'⁶, Sara E. Morton tells us that the results of a survey of members of the American Translators Association indicated that on-line access to a terminology bank was no more favoured than loose-leaf updates arriving by post. This is surprising, but as a display terminal is clearly more convenient than a printout, once this facility is available and the cost can be eliminated, there is little doubt that this will be the preferred form of accessing a data bank. The more a system is used the cheaper it becomes and hardware costs come down when the equipment becomes commonplace. The day cannot be far off when every office and home will be equipped with a Viewdata system linking a TV screen via the telephone to an infinite variety of accessible information. The cost of funding a dictionary for translators in the form of an information store before it has been expanded into a commercially viable entity is clearly a problem, as both Morton and Goetschalckx point out. It would need to be fully comprehensive and highly effective to function properly and gain acceptance and income, but this is not an insuperable difficulty.

Newly coined terms present experienced translators with some of their greatest problems, to which must be added unusual combinations of words and special or possibly incorrect use. Such terms are absent from dictionaries because the lexicographer rejected them as not being sufficiently authenticated, permanent or acceptable on his publication deadline—to say nothing of words that have come into use since that date. Such material can easily be incorporated into a data bank with appropriate comments and be deleted or updated as usage crystallizes or changes. The amount and type of information provided for each entry will be defined, refined and determined in the course of time, possibly on the lines of the current system operated by the Brussels Terminology Bureau⁷.

The growing use of microprocessors and their consequent reduced cost can only serve to make information storage and retrieval more accessible to us all. Who would have ventured, when calculators first appeared, to predict that every schoolchild would have one as a matter of course, but who would lay money against that now?

If we take the principle of the 'Language Translator' which I derided at the start of this article, we have a pocket computer fitted with preprogrammed cartridges which can be bought like dictionaries; these may be supplemented with memory capsules to add personal or specialized glossary notes. Here then is the technical capability for off-line data banks. On-line information retrieval has been operating with increasingly widespread success and efficiency for many years already but when will publishers

start compiling dictionaries for this market? They surely do not expect us to continue with the abacus in the age of the calculator merely for the sake of tradition!

Automatic translation

I shall make no attempt to describe the numerous areas and projects involving automatic, machine or computer translation, as it is variously described; this has been done more competently and comprehensively than I could hope to, both admirably succinctly⁸ and in remarkable and fascinating detail⁹. What can, however, be clearly and briefly stated with regard to the possibilities of machine translation, is that for the present it is simply a matter of how much or how little pre- and/or post-editing of texts is acceptable. This is most clearly reflected in the cost of man-hours, but there are other sides to the question. Words, as we have seen, can have many meanings and even more nuances, but if we control our input by using a restricted vocabulary, limiting each word to a defined usage, the computer should be able to handle this material with a high percentage of accuracy.

Returning to 'head', our example of a word with many connotations, we list it in our controlled vocabulary as valid only in the prime meaning as part of an animate being; in its second sense we will use 'mind', which in turn must never be a verb, and we will have to say 'look after the baby', 'heed the step', etc.; we will call a 'head of celery' a 'root', and so on through the twenty-two meanings given by Harrap with the possibility of many further subdivisions. All words must be similarly treated, e.g. 'lift' may only be a verb, 'elevator' perhaps taking the noun function, or we could have 'raise' for the verb leaving 'lift' for the noun; but once it is established, writers must adhere to the discipline of the vocabulary. Such a form of English is currently on the market¹⁰ and in use by a number of companies in variously tailored forms, proving especially useful for instruction and service manuals. Whether it will ever reach a general level of acceptance remains to be seen. Habits die hard and I can foresee a considerable degree of psychological resistance from experienced writers to any such imposition of vocabulary. The way we express ourselves is more than just a matter of convenience; it is surely at the very root of our personalities. Consider the measure of hostility engendered by countries when they try to change or abolish the language of a minority group!

A text controlled in this way makes computer translation much easier and it also becomes child's play for a human¹¹. This form of pre-editing does not require a translator or even a skilled linguist, although post-editing, to be effective, almost certainly needs to be done by a translator. It will no doubt be possible to find people prepared to write in controlled language but it may be more difficult to persuade translators, used to working with 'real' language, to undertake the task of unravelling and correcting machine output.

This does not merely reflect a possible lack of enthusiasm on their part but the realization that a text was produced by a mindless automaton violates a very fundamental principle of a translator's work. When we puzzle over obscure passages—maybe a semi-legible letter from a remote corner of the world, a painstaking effort for someone with little formal education, or at the other end of the scale, an abstruse philosophical passage, part of the life work of a thinker, or a research paper by a scientist, all possibly unskilled as writers—the thought that spurs us on in the

struggle to resolve the mystery of the text is that this is a message from one person to another or to others. Translators are the intermediaries, faced with the privilege and responsibility of interpreting and reproducing the essential meaning of the text. When the words fail to do this adequately we invest ourselves with imagination, like archaeologists trying to decipher hieroglyphics, and ask ourselves what the writer wanted to say. We know he would not have bothered to write unless he badly needed to say something and by putting ourselves in his shoes, ninety-nine times out of a hundred we find the key and our readers never know that the original did not spell out the message clearly. But, faced with machine output, we may feel from the start that perhaps the machine has lost the meaning and we might as well refer to the original human version. By the time we have done this once or twice, we will be tempted to ignore the machine's efforts altogether.

Conclusion

What conclusions can be drawn from all this? One cannot overlook the fact that those best qualified to use machine translation systems are the translators themselves, but they are unlikely to relish sharing their skills with a rival. As translators we must realize, however, that if we are not to be superseded by machines we should make the most of the facilities available to improve our performance and keep ahead on a cost-effective basis, it is on this, coupled with the quality of our work, that we will be judged.

The interests of the manufacturers of electronic equipment are geared to mass markets rather than specialist users such as translators, but current equipment might almost have been designed for our needs. The average staff or freelance translator of yesterday sat at his traditional office desk or kitchen table with a manual typewriter. He has been replaced by the user of an electric typewriter with carbon ribbon, backed by an increasingly wieldy collection of reference works. The next step will be a translator with a communicating word processor linked to a computer.

Let us look at this surely not-too-remote future: Reaching your desk in the morning you play out a text which has come in overnight from abroad and been stored in the machine. It is urgent. You type your translation. If there are any problem terms or if you wish to check usage in a particular context you access a data bank onto your VDU from a computer, on a time-sharing basis. You revise your work on the screen and decided on the layout; when you are satisfied, you phone your client and transmit the text at 120 characters per second. He takes your camera-ready translation out of his machine within minutes after you have finished working on it. This may be an electronic miracle but it is not a fantasy.

REFERENCES

- 1 *Advertisement, US News & World Report*, 30 April, 1979, 3; 13.
- 2 BLUMENTHAL, R. Adventures in translation *Business Traveller*, May/June, 1979, 75-7.
- 3 CHAUMIER, Jacques. *Synthesis of the economic and qualitative evaluations*. Paris: Bureau Marcel van Dijk, 1977. (Report No. 5) Study carried out for the Commission of the European Communities.
- 4 GOETSCHALCKX, Jacques. EURODJAUTOM. In: SNELL, B, ed. *Translating and the computer*. Oxford: North Holland, 1979. pp. 71-76.

- 5 TANKE, E. H. Implementing machine aids to translation. *In: SNELL, B. ed. Translating and the computer.* Oxford: North-Holland, 1979, pp. 45-70.
- 6 MORTON, Sara E. Designing a multilingual terminology bank for United States translators. *Journal of the American Society for Information Science*, November, 1978. 29 (6), 297-303.
- 7 A vocabulary designed for card index use, it covers a wide range of special fields. The EC languages are given with sources of information and a definition in one EC language is provided. Amendments are made on cards of a different colour. Circulation via the *BTB Bulletin* three times a year.
- 8 ROLLING, Loll. The facts about automatic translation. *In: Canadian Association for Information Science. Proceedings of the sixth Canadian conference on information science*, 1978. 267-9.
- 9 HUTCHINS, W.J. Machine translation and machine-aided translation. *Journal of Documentation*, June 1978, 34 (2), 119-59.
- 10 WHITE, E. N. An International Language for Servicing and Maintenance. *Cited in: Technical Translation Bulletin*. December 1979, 25 (3), 85.
- 11 KIRKMAN, John. Extending the ILSAM concept. *Technical Translation Bulletin*, December 1979, 25 (3), 80-86.