Introduction

Pamela Mayorcas

It was a pleasure and a privilege to chair the committee which planned this tenth anniversary conference and to chair the first session.

Those of you who are mathematically inclined will notice that this is the tenth conference as well as the tenth anniversary, since the first meeting was held in 1978 and the series skipped a year in 1979.

In planning the programme the committee thought it would be interesting to take a rearward look at what was being said and predicted about MT in 1978, compare and contrast this with the MT scene in 1988, and finally take a look at trends for the future.

The conference programme was therefore conceived as a tenth anniversary celebration, and we also thought it would be appropriate to pay tribute to Margaret Masterman who was such a colourful figure at the earlier conferences, and a major contributor to AI and MT research in this country. Session 4 was therefore dedicated to her and the paper by her former colleague Yorick Wilks is an eloquent and, at the same time, entertaining discussion of some of her ideas.

However, the Committee was aware that a historical review, while being a meritorious exercise, was of no great practical value to working translators and their managers. And so we ensured that the meat of the programme concentrated on some practical examples of the harnessing of computer power to translation production – and regular delegates will be aware that the scene is indeed very different from what might have been thought.

In the mid-70s there were visions of banks of terminals and powerful CPUs churning out millions of words of translation, with the occasional intervention of a linguist, like a technician adding a drop of oil to the machine, changing a word here, or entering an odd phrase into the computer dictionary there. At least that is how some of the science journalists like to represent it.

The current scene is very different indeed. First of all, there are very

few commercially viable systems and very few in commercial operation. However, many of the benefits of MT – machine-readable input, integration of text and graphics, networking, high quality output, dictionary compilation – have been incorporated into computer-assisted human translation programmes. So MAT or CAT is now the name of the game, and MT is the translation equivalent of the dinosaur. A glance over the titles of the preceding conferences in the series shows the different angles under which the central theme has been explored.

The first meeting, held in 1978, gave its name to the series *Translation and the Computer.* The second event was held in November 1980 under the title *Machine Aids for Translators,* and was the result of an initiative of the Aslib Technical Translation Group and the Translators' Guild of the Institute of Linguists in collaboration with Aslib. In the words of Barbara Snell, chairperson of the 1980 meeting, the papers showed that translators were keeping their feet firmly on the ground, and seeing what they could do to increase the quality and quantity of their output with the aid of modern technology. In the intervening years, many of the subsequent MT initiatives seemed to have lost sight of that perspective, but I think there is no doubt that the current stress is very much on the aids, and less on the machine.

The conferences followed on a yearly basis each November and have attracted speakers, chairmen and women, exhibitors and delegates from all over the world: *Practical Experience of Machine Translation* (1981): *Term Banks for Tomorrow's World* (1982); *Tools for the Trade* (1983); *Translation and Communication* (1984); *Translating and the Computer 7* (1985).

In 1986 the conference bore the title *A Profession on the Move,* and that year was indeed a professional landmark, at least in the United Kingdom, with the establishment of the Institute of Translation and Interpreting: ITI took over from the Translators' Guild as co-sponsor of the conference series with Aslib and the Aslib Technical Translation Group. The Programme Planning committees are drawn from representatives of ITI and Aslib TTG.

The 1988 tenth anniversary programme highlighted four themes: how the scene was perceived 10 years ago, and what did, in fact, happen (keynote paper): what is actually happening (Sessions 1 and 2): some new areas which translators need to investigate (Session 3): past and future assessments of MT and NLP research and new openings for the profession (Session 4).

In addition to the conference programme, there were three half-day workshops on lexicography, postediting and software translation on the preceding Wednesday afternoon attended by several of the conference delegates and other translators interested in these specialised topics. The workshop on translation of software organised by Ulla Magnusson Murray was particularly interesting: after short presentations by repre-

sentatives from Digital, Lotus, ICL and Ashton Tate we heard how these large software firms integrate translation into the earliest stages of software development. There followed a very open discussion as to how translators and software firms could best co-operate. For a full report on this and other workshops, readers are referred to *ITI News* vol 3 no 3 (December) 1988.

The keynote paper was given by Professor Sager, a frequent speaker and chair at these conferences. He explained where MT research had made mistakes and what should be done to improve matters. He also deplored the current dearth of qualified linguists which means fewer students entering the field of computational linguists and fewer to teach future generations of researchers.

He also advanced a novel proposal that as we move to multilingual document production, natural language and text processing techniques should be exploited in order to build translation into the very earliest stages of the document creation process, so that multilingual versions of text are built up from the very beginning. While this has great appeal and is quite probably possible at the theoretical and technical level (relational databases are used for precisely this kind of operation), I suspect that organisational and managerial complexities would put huge obstacles in the way of such a desirable end.

Fred Zirkle was most encouraging, I felt, in that he explained how a large organisation which had invested heavily in MAT had come to terms with the lack of commercial viability but was now, sensibly, building on the valuable experience it had gained by incorporating this into a highly sophisticated, technology-driven translation production network. He repeated the view expressed by Veronica Lawson, who played an important role in organising the early conferences, that the advent of machine translation would not deprive translators of work; if anything it would increase the amount of work since the increased volume handling capabilities would encourage managers to request translations of all kinds of texts previously thought impractical, or too expensive. Also technology brings translation into the forefront of managerial view, and puts it in the same league as other marketing and customer literature.

Fred Zirkle also endorsed earlier assertions that technology will not replace the translator but rather will help translation to move onto a higher plane, and improve the status of the translator. There is an opposite view, of course, which is that since all translators (or nearly all) now use word processors and PCs they are considered on a par with the office typist. I feel that much remains to be done to improve the status and consideration awarded to a highly professional group of people who are as necessary to the health of any firm or organisation as the armies of accountants, financial analysts, marketing managers and research personnel who are still more highly valued (and paid) than the average

translator. The Institute of Translation and Interpreting and its sister FIT organisations play an important part here, and there is visible progress.

Isabella Moore confirmed what many translators have found, namely that all but the most expensive OCR equipment is quite unsuitable for converting non-standard characters and less-than-perfect paper copies into machine-readable form. Since the economic viability of MT/MAT systems depends, in large measure, on the source text being in machine-readable form and already loaded into the system on which you are working, this is a fairly major handicap. She also had some disturbing experiences to recount of the difficulties which small users may experience in their relationships with suppliers who, apparently, may lack the linguistic knowledge for explaining the functioning of certain features of the system, sorting out translation-related difficulties and assisting with customisation.

She also provided some salutary warnings as to the cost effectiveness (or otherwise) of CAT in a small business environment – her opinion was that the market would probably not bear the real cost of such a service, so that it would have to be marketed as some kind of loss leader with stress laid on the improved quality of service (customer-oriented terminology, high-quality output); even more worrying were her findings that the true costs are probably very difficult to measure.

The importance of terminology, as a discipline in itself, as a tool for translators and as a key element in any CAT system was stressed in many of the papers, and this is perhaps one area where translators and their managers in the UK need to take a closer look at the more developed study of terminology in Europe.

Alain Paillet had no doubts that good terminology was at the heart of any decent system, and stressed how important it was for users to define very clearly what they expect from the system. He made the interesting point that the cut-and-paste facility promoted as a highly desirable feature of many systems was not perhaps so essential since one hardly ever uses the term in the form it occurs in the dictionary, or exactly at the place where the translation is interrupted to find a term. This will probably give the developers of terminology software some cause for thought. Alain Paillet and Isabella Moore seem to share a certain scepticism as to what constitutes commercial viability, and whether the CA or the T are the most significant part of any CAT system.

His company's decision to develop their own terminology system provided further evidence of a trend, which I noticed in the early 80s, towards a multiplicity of term banks and terminology systems as users perceived the universal terminology bank, like the universal translation machine, to be something of a holy grail. Users have begun to realise that they have to define their own needs and build a system to suit their

particular type of terminology, with the type of input and control procedures best suited to their working methods and, most especially, information retrieval procedures which best match their particular requirements. So, just as there are many different dictionaries, and even more specialised glossaries, we now have very many firms and companies building their own terminology stores. I believe that we shall be hearing more about this in the 1989 conference.

Dr Jackson treated us to a refreshingly honest appraisal of the current situation as regards file transfer, which is problematic largely because of the range of characters that occur in translated texts. Now that industry and commerce is becoming aware of the need for producing multilingual documentation, it may be that they will find solutions to these problems, but it is one translators have been struggling with for years, getting virtually no response from suppliers and being regarded as cranky and fanatical because they insist on being able to type, read and print an 'é' as just that and not '£e' or even 'e´'. (Is the recent French proposal to abolish the circumflex inspired by a desire for spelling reform or simply capitulation to the problems of technology?)

His eloquent phrase, that there is an 'anarchy of data character conversion codes and character representation which inhibits data transmission' struck an all too meaningful chord with many in the audience. But at least forewarned is forearmed. It is perhaps a positive sign, as we learnt at the software translation workshop, that computer software suppliers are themselves facing up to these problems and searching for solutions to machine and systems incompatibility, and this may filter through to translators. This was reassuring because it proved that translators are not fanatical, or cranky, that the problems are real, and that even though no immediate, easy or fast solutions are available, it was better to face up to an understanding of the problems rather than try to whitewash over them.

Translators can sometimes seem a little paranoid, reacting as if the whole world of hardware and software suppliers are acting in some kind of malicious conspiracy against them. So it was in some ways comforting, although still depressing, to learn from Barry Mahon that we are not alone and that other professions suffer equally with telecoms problems. But it is also true that a proper understanding of what goes on and of the various aspects of telecoms and file transfer puts us in a much better position to discuss our requirements with suppliers, or troubleshoot problems. Both these papers merit repeated reading.

Large suggested that as translation becomes increasingly specialised and new terms and concepts cannot be found in dictionaries, translators should learn the skills of information scientists to locate information and also to find translation equivalents. While translation is essentially about disseminating information, translators themselves are information users

and need to use the same skills employed by information specialists in their job. Information retrieval skills are also valuable for determining the optimum design for terminology systems incorporating database-type functions in order to enhance the collection, storage and retrieval of collected terms. Equally, the boilerplating techniques described by Knowles for repetitive work or where texts have to follow strict formulae fall within the realm of database technology, rather than machine translation.

Doug Arnold, in his introduction to Session 3 enlarged on Professor Sager's innovative paper of 1981 *(Practical Experience of Machine Translation,* North Holland, 1982) in which he put forward a theory of text typology – suggesting that before MT could be effective, users had to make a careful assessment and selection of text types and then select and adapt the appropriate system. This survey is essential reading for those who tend to lump all translation together into one big box, and do not think in terms of different types of text and different types of language all of which need to be handled differently. Translators know instinctively even if they have not deliberately thought about it, that they adopt different modes, different gears if you like, for different types of text (minutes, laboratory reports, contracts, journal articles) and that MT systems also need to take this into account.

However, Arnold joins with Knowles and Wilks in suggesting that this truth should not put a halt to research into natural language processing and computational linguistics which can yield valuable insights into the nature and behaviour of language, and provide new avenues for MAT.

The papers by Scott and Pym describe practical, and highly sophistic-ated procedures based on MAT systems which have been customised to suit particular types of text, both demonstrate the absolutely essential part which careful and detailed planning and strong management have to play in ensuring a viable and workable system which is cost-effective, achieves the desired result and is user-friendly.

While I was listening to these papers, however, I could not help recall a passage in *After Babel,* in which George Steiner considered the Kabbalistic vision of a world without translation, a world returning to the pre-Babel state in which a universal communication and understanding obtained. However, the Kabbalah also foresees the desolation of a world in which language loses its meaning and no longer serves as a vehicle for communication; many of us feel such concern in regard to the develop-ment of restricted language systems.

Their experience seems to confirm the very early evaluations of Systran (circa 1978) which found that MT is only cost effective where there are high volumes in a narrow and specific subject area and text type, the input is already available in machine-readable form and costly pre-editing is eliminated (the PACE controlled language is in effect a form of pre-

editing but is an integral part of the document creation process and not a separate operation).

One significant change that translators are facing is the increasing demand for them to be involved in all aspects of document production. Here again, Professor Knowles foresees further anarchy since translators are not being involved in the design of systems created to produce multilingual documents. There are no good practice standards for highly formatted documents embodying typographical variety and interspersed graphics, all normal fare for technical translators.

Professor Wilks constructed his tribute to Margaret Masterman around a survey of some of her ideas, some of her brilliant insights, and her passionate and almost naive desire to get into the translators brain – an area at that time ignored by most developers who concentrated on the analytical mathematical approach.

He could not deny that there were some curious contradictions in her work. While she based her actual research on the premise that MT did not require inference knowledge of the world and highly complex analysis, but only required a superficial processing of language, this was in contradiction with her statements that researchers had to get into the brain of, and work closely with the translator. This last was a brave position to adopt at the time, and of course now proves to be the only viable one, yet many in the MT/MAT field continue to ignore this so-obvious premise.

Finally, we were treated to a most exciting and novel account from Dr Luyken of the exciting new areas opening up for translators and interpreters in the field of media translations; this requires an entirely new approach, the learning of basic skills and integration with a range of other professionals in film making, television and advertising. It was generally agreed that this was one of the most lively and refreshing parts of the conference.

It is interesting to note that the same messages emanate from the papers and discussions as they did ten years ago: translators must be involved in the development of MAT systems; the cost effectiveness of MAT systems depends mostly on non-linguistic functions, such as machine-readable input; dictionaries, which are the key to MAT systems, are costly and difficult to update; there is still a problem of compatibility, character conversion and file transfer; telecommunications are still a hazardous area.

One of the most serious complaints about MT research is the fact that the extremely costly, and labour-intensive work of compiling and updating dictionaries for MT systems has not been available outside of the MT environment for everyday use, by working translators.

Much of the real progress lies in the integration of telecommunications and desktop publishing technology with the translation production

process so that customers now receive a sophisticated product from a new generation of translation suppliers – the role of DTP and other document production technologies will be examined at the next and future conferences. The principal contribution of computers has been to improve volume but not necessarily quality of output. Or, as Isabella Moore put it, technology is best substituted for mechanical intervention, while the thinking process is still best left to the translator. According to Knowles, MT research has simply confirmed the need for human beings to save the machine from disaster. We hate to say we told you so . . . !

However, natural language processing tools can make a valuable contribution to the conversion of source to target text by providing the techniques for editing and text compilation, in other words contributing to the translation production exercise rather than to the act of translation itself.

Those who attended the first meetings in 1978 will notice that the accent has changed considerably from pure machine translation theory and research to a situation in which, at last, translators are closely involved in the development of systems. There have been disappointments as well as successes, notably perhaps the failure of the authorities to rise to the challenge of creating a British Term Bank. I think it is also true that many translators have been guilty of wanting things to be handed to them on a plate.

The picturesque description of MT/MAT software development as a 'juggernaut lurching from one financial deadline to another' must have struck sympathetic chords with many translators in the audience involved in research programmes.

The annual conference has become one of the highlights of the translation year and has helped to introduce a large and international audience to word processing, termbanks, thesauri, online data, information retrieval and provided a remarkable forum in which academics, developers, translation users and last, but definitely not least, translators have been able to debate new research and applications. As Professor Sager says in his keynote paper, suppliers no longer make exaggerated and unsubstantiated claims, firstly because there is nothing to be gained from so doing, secondly because the audience has become much wiser than it was in 1978.

There can be no doubt that there is now general agreement that the translator must be the focal point of any MAT system. However, as is made clear by Messrs Wilks, Knowles, Sager, and Zirkle much still needs to be done to achieve this both as regards understanding the intellectual basis of translation and understanding the translation production process. Zirkle predicts the development of mega translation companies With world networks, mother tongue translators (are there any other kind?), highly qualified specialised translators, all supported by state-of-the-art technology.

It is interesting, and this is particularly noticeable in a professional organisation such as the Institute of Translation and Interpreting for example, that translators, both staff and freelance, have begun to build their own translation systems from the bottom up, starting with the humble PC and word processing software and learning how to build on or in facilities and functions which help them to produce better translations faster.

Whether all this will lead to the emergence of MT/MAT systems as originally conceived of in the 60s it is not my place to say; certainly it will not happen without system developers following Margaret Masterman's injunction that they learn how translators do it, and that translators show how they do it. In order to develop into a viable product, MT needs to come out of research environment and the product needs to be commercialised, but as Fred Zirkle points out, this requires tremendous financial commitment. Research seems to lead to further research rather than reaching a stage where solutions to problems are found and can be implemented in an experimental industrial application. One important thing we have learned from these conferences is that the predicted lead times are highly elastic.

We are most fortunate in this excellent collection of papers, not only because they form a valuable record in themselves but in addition they are a most instructive, comprehensive and very frank appraisal of the way in which thinking on MT/MAT has evolved, the errors made, the real contribution of computers and systems to the translation production process.

We are also fortunate in the extensive bibliographies which have been provided with some of the papers, which together constitute a serious reading list for anyone interested in penetrating this fascinating subject still further.

I think it can be fairly said, in sum, that these past ten conferences have had tremendous informative and educative value. My hope now is that the next ten years will see some real practical benefits emerging from the proper dialogue between translators, clients, and suppliers of systems and software. This dialogue has begun as a result of initiatives taken by the Institute of Translation and Interpreting, and this conference has played a major part in bringing the various parties together.

Finally, I should like to end with a word of thanks to all those involved in the success of the conference – Aslib, members of the Conference Planning Committee, the chairpersons and speakers, and the audience without whom the conference would be of no value. I should also like to give special thanks to the staff of Aslib's Publications Department for the hard work that went into producing this volume.

REFERENCE

Steiner, G. *After Babel,* Oxford: Oxford University Press, 1975.