In recent years, a large amount of work has been done in the field of computational linguistics which should be of immediate interest to translators, and perhaps also to interpreters. It is the purpose of this paper to bring to the attention of professional linguists the usefulness of and benefit to be derived from work currently in progress in this field, and to allay the wariness and scepticism that they may feel towards computers. We hope to demonstrate that, first, given its nature and purpose, automatic translation does not work in competition with human translators, but rather alongside them; and secondly, that there are several other computer aids for translators as yet only available to large bodies employing translators, but which in the very near future should be available to the individual freelance translator working at home on the slimmest of budgets.

1. Machine Translation (M.T.)

Most readers will be aware that research into automatic translation is a major activity in the field of serious non-numeric computing. Some may be less aware of its somewhat chequered history, beginning in the United States in the 1950s, through its loss of academic respectability following the disastrous ALPAC report of 1966, which stated that 'there is no immediate or predictable prospect of useful machine translation', to its renaissance in the 1970s, thanks largely to the fact that research is now based on much sounder linguistic theory. Excellent and up-to-date histories of
M.T. are given by Hutchins and by Vauquois, and Bruderer's work is, though not very readable, most comprehensive; further historical details would be out of place in this paper.

Some mention should be made, however, of a few currently operational systems. Research in North America has produced several such systems, of which perhaps the best known are the following. In Canada, TAUM-Meteo produces French translations of weather reports written in English, and the TAUM-Aviation system, again translating from English into French, is nearing full operational implementation. Systran was originally developed for translating articles in Russian scientific journals into English, and is now used by several large corporations (for example, Rank-Xerox) to translate between other language pairs, as well as by the Commission of the European Communities. In Europe, Grenoble's GETA system handles several language pairs, as does Saarbrücken's SUSY, albeit experimentally. Lastly, a Chinese–English translation system, CULT, is worthy of mention. This is described by Loh and by Loh & Kong, and is used extensively to produce translations of the mathematical Journal Acta Mathematica Sinica. A ready example of CULT's output is given in another article, the English version of which was produced by CULT itself. Mention may also be made of EUROTRA, the M.T. system proposed for the European Commission. Although this is some way from being operational, researchers from several European countries are already collaborating in earnest on the system, which is being designed to translate between (currently) thirty language pairs (involving the six languages recognized for official use by the Commission).

The quality of the translations produced by M.T. systems varies enormously. Some are of very poor quality, while others begin to approach a standard which may be described as useful for information purposes. What is the point of M.T. if the translations thus produced do not compare with human translations? The most obvious advantage of the inferior M.T. product is that it is both cheaper and quicker to produce. These factors alone would not justify the research done on M.T., however: increase in efficiency is not acceptable if it results in an inferior and therefore undesirable product. Yet there is evidence that automatically produced translations, even if of fairly low quality, are acceptable and desirable in certain circumstances. It can often be the case that if a text were not translated by machine, it would not be translated at all, so that end-users are pleased to have rough translations quickly and cheaply, from which they can judge whether the original is worth the bother of translating 'properly'. In a survey of seventeen end-users of Systran translations, it was found that '71% would be interested in receiving unrevised machine translations — from languages they do not know, — for their personal documentation and for working documents'.

We would not like to give the impression that all output from
M.T. systems is of low quality. Much depends on the input. One idea that has been used with some success is to limit the complexity of the input, and thus ensure good quality automatic translation requiring no post-editing. This limited-syntax and limited-vocabulary approach has been used extensively by, for example, Rank-Xerox, to produce technical manuals in a variety of languages. One result of this technique, as Lawson points out, may be that the source texts are written in language which is 'more elegant than their standard jargon' and which 'reads easily and fast' (pp. 81-82). The French TITUS system similarly produces fully automatic translations of texts the syntax and vocabulary of which is controlled and limited.

In the majority of systems however, humans and machine co-operate to produce the better translations. The amount of pre- and post-editing by humans differs from system to system. Texts being translated by the C.E.C.'s Systran are passed to humans for post-editing only, and it is often the case that the output requires only a little correction. The bulk of texts that must be translated in the Commission represent a drudge for the professional translator - texts for which 'style, and even nuances, are of secondary importance' (Lawson, p. 84) - and the ability to off-load this drudge on to a machine that can do the job almost as well, more quickly, and without fatigue, will surely be welcomed by all concerned. This would then leave the human translator free to concentrate on the sorts of texts that, for the foreseeable future at least, will never be handled by an M.T. system, and which require distinctly human capabilities. Lawson calls these 'texts with a message for the less-than-committed reader' (as opposed to 'routine information texts') - literary, political, and educational texts, the style of which is marked by deliberate use of ambiguity, idiomatic turns of phrase, complex figures of speech - just those features of natural language that are the most problematic for automatic analysis and translation. These texts require all the ingenuity and creativity of an experienced human if a reasonable translation is to be achieved. Another class of texts that cannot be trusted to a machine are those for which an especially high quality of translation is required, for example diplomatic and legal documents. So the off-loading of a huge volume of routine paperwork will free the translator for the more interesting, demanding, rewarding, and creative work of translating these sensitive texts.

Let there be no fear that a sufficient volume of translation work does not exist. Generally, there is a considerable backlog of texts requiring translation. In the C.E.C., for example, where every document must be produced in six languages, often in a hurry, 'the style of our translations often suffers because of the very short deadlines against which we have to work' (Arthern, p. 81). This in spite of the fact that a substantial part of the Institutions' budget is for linguistic services, and approximately 1700 translators are employed full or part time; furthermore, in 1973 Anderla suggested that by 1987 the world's annual production of scientific
and technical information would be six times the then current rate.\textsuperscript{20} A veritable 'translation mountain'!

The translator does have something to offer M.T. research in return, especially as part of his job may entail revision of machine output. Feedback detailing the kinds of errors that the machine makes (they will be very different from those typically made by humans) together with any linguistic insights they may inspire, will always be of use to the M.T. system designers, too few of whom, alas, at present at least, have much formal linguistic training. As Arthern notes, while perhaps few linguists become involved in programming and preparation of dictionaries, others are involved in revising M.T. output. As a result of this co-operation, 'for them machine translation has become a colleague' (p. 79).


As we require the human to take a more and more active part in the mechanized translation process, we move away from M.T. proper into the area of Machine-aided Translation. Whether the dividing line between M.T. and Machine-aided Translation is clear cut or not is of little consequence. For us, the difference lies in whether the translation is of whole structures (M.T.) or of lexical items (Machine-aided). (Indeed, the machine may aid only in cutting down the translator's paperwork, but more of that later.) In this section, we will consider the use of the computer to provide for the user translations of individual items in the source text, leaving the syntactic transfer into the target language entirely to the human translator. In this way, the machine is used as a type of automatic dictionary.

One such system is described by Canisius.\textsuperscript{21} Briefly, the idea is that technical terms are stored in a multilingual dictionary, and that the source text is automatically scanned for an occurrence of any of these items. The result of the treatment is eventually a text in which the grammatical structures of the original remain, but with the vocabulary largely translated, as in Figure 1. Canisius points out that 'users with basic [grammatical] knowledge of the original language of a document can effortlessly understand the text on account of the interspersed words' (p. 264), and that the partial translation gives the reader some idea of the subject matter of the text. Generally, however, the mixed-language output must be seen as one stage in the process of translating a given text, not as an end in itself. The whole system does need the firm support of a dependable dictionary (as well as a lemmatization program to strip affixes), and the next section deals with the use by the translator of the many different types of computerized dictionaries and term banks.
3. Other computer aids for translators

Most translators will no doubt not be immediately involved with the various M.T. systems described above. However, it cannot be denied that there is a growing interest in M.T. — witness the recent commitment of the C.E.C. to a European translation project, noted in Rolling. This fact, when considered in the light of continuing developments in long-distance information communication systems (compare the various systems being implemented by national Post Offices and television companies) should prompt the professional linguist at least to investigate the field of M.T., as it is quite probable that in the future, the individual will be able to access an M.T. system from his own sitting-room or office via his own television set. The C.E.C. plans to establish very shortly a European information network for scientific and technical information (EURONET) which will link data bases in several member countries and will be accessed via post office communications. One of the first data banks to be made available to EURONET users will be the C.E.C.'s EURODICAUTOM. A freelance translator will then be able to avail himself of facilities previously reserved for staff translators.

Mention of data banks introduces an area which is perhaps more familiar to the professional linguist: the storing and accessing of terminological information by computer.
There are two main ways of storing and accessing terminological information via computer that every translator should be aware of, bearing in mind the spread of long-distance information systems and the tremendous impact (present and future) of the microprocessor: data may be stored on a vast scale in a terminological data bank, the services of which are made available to a wide range of users; or, restricted field data may be stored on a medium such as a floppy disk, which can be thought of as akin in function to magnetic tape, and which constitutes a permanent record of information that may be accessed by a mini-computer. That is, in the near future, a translator could conceivably possess the equivalent of several bilingual and multilingual dictionaries on disk, at a fraction of the cost of a conventional published dictionary. Mini-computer prices are decreasing rapidly, to the extent that the 'home computer' is a definite reality.

4. **Terminological data banks (Term banks)**

Term banks provide a wide range of different services to many kinds of user: translator, terminologist, specialist librarian, technical writer, language teacher, abstracter, lexicographer, etc. They came about because of a desire not just to store information for its own sake, but to store information on technical terms, industry standards, and the like, and to provide means of accessing these data in a variety of illuminating and rapid ways. This automatically implies that term banks are not tailor-made for the translator, and indeed some dissatisfaction over services provided by term banks has been expressed. Several criticisms may be found in Rondeau et al. (rather damning) and Arthern (pros and cons).22 Nevertheless, the number of term banks is increasing constantly, and the quality of their services is such as to allay criticism. There exist several comprehensive listings of established and proposed term banks and accounts of their organization and services.23 Term banks catering specifically for the translator include EURODI-CAUTOM,24 Bundessprachenamt,25 Siemens,26 Montréal,27 and Québec.28

In Britain, UMIST has recently been awarded a grant from the British Library to conduct a feasibility study into the establishment of a British Term Bank.29 One important point to note is that if a term bank becomes a reality in the U.K. then the architects of the bank will be able not only to draw on the best aspects of existing banks, and incorporate the latest in storing and accessing systems, but also to tailor services to provide users with the kind of information they require. Figure 2 shows a scheme for a model term bank, as proposed by Krollmann, which meets the requirements of translators.

Is there however a real need for term banks from a translator's point of view? Let us look first at some consequences of not using a term bank, and let us take a simple everyday task that of
Figure 2. Model of a linguistic data bank (from F. Krollmann, 'Linguistic data banks and the technical translator', META, 16(1971), p. 118.)
discovering the appropriate translation of a new term. The translator's trusty companion, the bilingual dictionary, proves to be of little help in this instance, since, when he bought it, it was probably already out of date. The translator may have to wait several years for the next edition to find the required term - of scant use when the translation is required immediately! And moreover, by the time the term does appear in print, a revolution in the technology of the appropriate field may well have made the term obsolete. The translator may then have to resort to other methods: consulting other translators' files, or asking an expert in the field. This may provide no help, or produce several different translations. A further method would be to look for the term in the specialized vocabulary of a different field, with the danger of procuring a misleading or totally false rendering. As a last resort the translator may coin a neologism. (Coining of neologisms may have been the reason for several renderings being found in other translators' files.) Several translations for a single term together with the proliferation of neologisms raises the problem of consistency of translation.

Further, it has been found that a translator spends as much as 60% of his time searching for terminological information in published dictionaries.\(^{30}\) What is more worrying, in view of this inordinate amount of valuable time devoted to research, is that, according to one study, 62.1% of translation errors were found to be caused by terminological inaccuracies.\(^{31}\) With the information explosion already upon us, and the ever-increasing demand for translations, it is clearly in the interests of all participants in the translation process to ensure that translations can be produced quickly, efficiently and accurately. And at low cost.

We would argue that this may be achieved through the use of a term bank. Consider some of the services provided by a term bank and the advantages it represents over the use of conventional dictionaries. A properly constituted term bank should provide the translator with, first, a rapid means of accessing a far wider and potentially more informative range of terminological data than that contained in a published dictionary (which has the added disadvantages of being bulky, time-consuming to use, soon obsolete, prone to wear and tear, and not particularly cheap); and secondly, various means of presenting data to suit special needs.

The ideal term bank will be equipped with computer dictionaries and accessing systems that are organized in such a fashion that the translator receives just the type and amount of information he wants. Criticisms levelled at early term banks were that they provided either too much information, which was distracting and misleading, or not enough. The systems of the Bundessprachenamt and of Siemens, for example, provide several, useful means of organizing information. At the Bundessprachenamt, the user may request among other services the following.\(^{32}\) firstly, printout of a text-oriented glossary, which
provides translations for specified terms in the order they occur in the text. A survey of the use of such an aid found that translators spent a third less time on a translation and made only two-thirds of the errors registered by translators working conventionally.\textsuperscript{33} Secondly, alphabetic printout of a text-oriented glossary is useful for ensuring terminological consistency when work is split amongst several translators. Thirdly, microfiches of any part of the term bank are provided. These are cheap and easy to produce and are ideal for short translation jobs. They may be replaced at regular intervals, which will ensure that the translator has a fairly up-to-date record of the term bank’s holdings in his particular field. About 8,000 terms can be stored on a microfiche of 10 x 15 cms.

The Siemens system likewise produces text-oriented glossaries and various alphabetical lists, on microfiche or hard copy.\textsuperscript{34} Moreover, it provides an ideal example of how a term bank may centralize information, yet make it widely available, for a co-operative has been formed, which includes Siemens staff, freelance translators, private bureaux, etc. Terminology is stored by Siemens and is made available through various communications systems to other members, who in return play a vital part in contributing data to the bank.

Those translators who work in large organizations will be interested to learn that much research is currently being carried out on on-line interactive methods of interrogating term banks. Such systems implemented on main-frame computers are at the moment of practical use only in large organizations with their own internal communications system, for example, Siemens. A well-defined interactive interrogation system provides the translator with several options. It allows him to pursue a term’s meaning and translation, and the relationships it contracts with other terms, by permitting him progressively to widen or narrow the search space. It provides him with contexts, definitions, sample usages of the term, or a display of the term in a thesaurus-type hierarchy with related terms. It also provides synonyms, antonyms, and near equivalents, indications of style and register associated with the term, and statistics on, for example, the frequency of use of the term.

Arthern makes a plea for what he calls a 'translator package' which would provide on the screen of a Visual Display Unit the minimum amount of information usually required by translators (p. 93). This plea arose from his experience of EURODICAUTOM. Carestia-Greenfield & Serain likewise suggest that user resistance to computers would be overcome if systems were designed with the end-users' needs in mind (p. 93).

So far we have concentrated on the types of services provided by a term bank. Little has been said on how a term bank is set up, or indeed maintained. Data for a term bank are culled from a variety of sources: standard specialized dictionaries, both monolingual and bilingual glossaries, thesauri, national and international standards
One of the prime motivations for setting up term banks is to ensure that the most up-to-date information is available to users. Two main methods are employed here. New terms are obtained through automatic keyword extraction techniques upon machine perusal of the latest literature. These terms are then processed by terminologists and inserted into the bank. Alternatively, reports are received of new terms adopted or coined by translators who use the bank.

This last provides an extremely rich source of information and will not be neglected by the manager of a term bank who wishes to keep or attract customers. A personal interest in any system is a strong motivating factor for using that system. Thus the translator will be able to furnish new terms and translations he has come across, either by filling in specially formatted cards, or, when working interactively, by entering proposed modifications into an intermediate store, which will be inspected regularly by a terminologist. (This constraint is necessary as the main data base will under no circumstances be susceptible to direct modification by the user, for obvious reasons.) Arthern's objection (pp. 86-87) that a translator working interactively cannot provide as much feedback as one annotating a printout is thus overcome. The terminologist will likewise be responsible for deleting obsolete terms.

Carestia-Greenfield & Serain and Arthern, among others, note that there is an unfortunate resistance among translators to computer aids, and especially to interactive interrogation. This is understandable if the objection is that the desired information is not supplied. However, resistance to the computer for subjective reasons must be overcome, for even today difficulty would be experienced if computer aids to translation were not available. Krollmann states that 'by and large it is true to say that despite initial opposition on the part of the translators, who at first were sceptical towards the computer, the day to day work of our translation service would now be inconceivable without the system'.

It should also be clearly understood that as far as the translator is concerned the computer, its dictionaries and accessing system are all part of a 'black box'. He need know nothing of how they operate. Above all, he need not know how the dictionary is organized. The accessing system, front-ended by a natural language interface, interprets user commands and requests and is responsible for extracting information and presenting it in as helpful and useful a manner as possible. A systematically organized computer dictionary contains much information, implied by the way it is organized. One of the main drawbacks to a published dictionary is that the user must be aware of how it is organized, if he is to draw full benefit from it. That is, he must know how entries are sub-divided, whether compound terms are to be found under the first or subsequent elements...
5. Personal dictionaries and text-processing equipment

The impact of the microprocessor has been such that today a freelance translator may aspire to possessing his own computer. Within a few years, the home computer will be commonplace. For the translator, this means that he will have access to the same kind of services as are at present provided by term banks, though on a smaller scale. Mention has already been made of the amount of information capable of being stored on a floppy disk. With his own system, the translator will of course be able to add or delete terms from his dictionaries at will. Note that access time is usually faster on a small computer with only one terminal than on a mainframe computer run on a time-sharing basis. Moreover, there already exist hand-held computers that provide a limited number of translations. While these are still very much at the novelty stage, interpreters in particular should follow their development closely.36

Another application of computer aids to translators may be seen in the wide range of text-editing and formatting equipment and systems now available. These range from word-processors to on-line text-editors and machine type-setters. Working with a text-editor for example, a translator may edit and amend a text at will, change the format, rearrange the sequence of paragraphs, and produce a final version which may then be printed out on a line-printer or sent down the line to a type-setting device. All this can be accomplished without setting pen to paper. The finished text may also be compressed and stored so that later, should the need arise, all or part of it may be reproduced. Tanke gives a comprehensive account of text-processing facilities. As for the advantages of text-processing, Sager gives some interesting statistics based on studies carried out by the C.E.C. in Brussels and Luxembourg, where some 400,000 rough and 130,000 clean pages are produced annually. He points out that 45% of the original effort involved in retyping revised translations could have been saved by the use of appropriate text-processing equipment.37

Lawson makes a plea for 'something written specifically for translators' (p. 83), and we trust that this article in some sense meets that need. It seems evident to us at least that the translator of the future will have to come to terms with technological advances that directly affect his day to day work, and that he will develop, as Sager puts it, 'new attitudes and working habits'.38


5. The list is by no means exhaustive. For more details of these and other systems, see Hutchins.


7. TAUM, Groupe de recherche en Traduction Automatique de l'Universite de Montreal (TAUM) (Montréal, TAUM, 1979).


The translator as Computer User


19. B. Vauquois, 'Aspects of mechanical translation in 1979'.


terminologies et linguistique computationnelle', in META, 22 (1977), pp. 184-91; Arthern, op. cit.


27. R. Dubuc, Description du système TERMIUM (Montréal, Université de Montréal Banque de Terminologie, 1975); M. Paré, The University of Montreal's banque de terminologie (Nice, International Federation of Translators, 1974).

28. J. M. Fortin, La Banque de terminologie du Québec (Québec, Banque de Terminologie, 1974).


32. Krollmann, 'User aspects of an automatic aid to translation', in Overcoming the Language Barrier.

33. ALPAC, op. cit.

34. Tanke, op. cit.

35. Krollmann, 'User aspects of an automatic aid to translation', in
Overcoming the Language Barrier, p. 245, with our emphasis.

