[*Translating and the Computer 9*. Proceedings of a conference… 12-13 November 1987, ed. Catriona Picken (London: Aslib, 1988)]

A survey of termbanks worldwide

John McNaught

Centre for Computational Linguistics, UMIST, Manchester, UK

INTRODUCTION

In the short space available, any survey must of necessity be brief, thus the reader should not expect a detailed, exhaustive survey of the state of termbanks across the world. In what follows, therefore, I will emphasise what I perceive as general trends in computer-based processing, storing and dissemination of terminologies, concentrating on how recent developments regarding termbanks relate to the translator.

As far as I am aware, there is no central register of termbanks. This alone makes a complete survey impossible. There are organisations which endeavour to collection information on new termbanks and changes in established ones, but this largely depends on individual termbanks signalling their existence and being willing to participate in such an activity. This was not the state of affairs some years ago. When I last had the privilege of addressing this conference, the number of established termbanks in the world could be confidently said to be few: they could be counted on the fingers of two hands. There were a number of projects planned or under development, and much talk of a future where there would be numerous termbanks serving a variety of clients, each with large stores of terminology, in a variety of languages. Predictions abounded that translators, having the greatest requirements of all groups for terminological information, would be able to access such rich stores of knowledge easily from their place of work, and hence achieve greater efficiency, accuracy and consistency in their translations. There was every expectation that termbanks would cooperate to exchange data, develop data in common, and provide a flexible, networked service for their envisaged worldwide user community. There was further hope that translators would come to play a key role in the development of termbanks,

by providing feedback or participating in the elaboration of new collections of terminology. There were no doubt expectations that, in the near future, someone would stand up in front of an audience and describe a harmonious worldwide terminology information service based on a network of termbanks. There were even plans that the United Kingdom should participate in such activities.

It is painfully evident that little of this Utopian state of affairs has come to pass. What we have seen over the past six or seven years, however, is a separate, but related, set of developments in terminology processing. I will examine these new trends below, in relation to the impact of termbanks on the translating profession. First, however, I will indeed attempt to survey what has been going on in termbanks over the last few years, and examine what benefits the individual translator may expect to derive from these developments.

RECENT DEVELOPMENTS IN TERMBANKS

Our first task, perhaps surprisingly, is to determine what constitutes a termbank. A number of years ago, when UMIST carried out its survey of existing major termbanks, as part of the feasibility study of a British termbank, supported by the British Library Research and Development Department, we were faced with a similar problem, and elaborated a definition that seemed to us to encompass the essential characteristics of an 'ideal' termbank. At the time, we used the term 'linguistic databank' (LDB), which seems to have gained some currency, as during our 1980 survey, we discovered that there were several different kinds of termbank, serving distinct purposes and clearly identifiable user groups, with different structures and modes of operation, and, importantly, processing data that were more than strictly terminological in nature. Our definition was as follows:

> 'a collection, stored in a computer, of special language vocabularies, including nomenclatures, standardized terms and phrases, together with the information required for their identification, which can be used as a mono- or multilingual dictionary for direct consultation, as a basis for dictionary production, as a control instrument for consistency of usage and term creation and as an ancillary tool in information and documentation.'
>
> Sager and McNaught (1981a: 1)

The classification of termbanks we developed then is still largely valid today. One class of termbanks serves mainly the translation needs of large organisations. Examples are LEXIS, of the Bundessprachenamt (FRG), Eurodicautom of the CEC, TEAM of Siemens (FRG), Termium of the Federal Canadian Government, MULTILEX of the All-Union Centre for Translation (USSR). LEXIS and Termium are described in companion papers in this volume, by Hoffmann and Landry, respectively. A second group of termbanks is motivated by the needs of language planning, the prime example here being

the termbank of the Office de la langue française in Quebec, which is seen as a vital method of maintaining French in Quebec as an authentic, living language with the ability to expand and regenerate itself in a dominant English-language macro-environment. A third group is constituted by termbanks associated with standardisation agencies. The main ones are NORMATERM of AFNOR (France), the termbank of DIN (FRG), which is based on TEAM, and in the USSR, the ASITO termbank of VNIIKI/GOSSTANDART.

Some of these termbanks have expanded their roles and now attempt to serve a wider public having differing needs. Most major termbanks, however, continue to serve primarily the clients they were set up for in the first place. This is an important characteristic of the major termbanks, and one we shall return to in due course. A detailed survey of the main European banks formed part of the above-mentioned feasibility study, and is published as Sager and McNaught (1981b). A more up-to-date description is available in Bennett, *et al.* (1986).

We may today add three further types of termbank:

(a) the national termbank, which attempts to serve a general purpose role in coordinating the creation and use of terminologies within a country, and hence is theoretically multifunctional, multilingual and exploited by widely differing kinds of user;
(b) at the opposite extreme, the termbank which may, today, be available on microcomputer for individual or company use;
(c) projects at universities and other research centres which have grown into, or will shortly become, fully-fledged banks. This type of bank covers a wide variety of different banks, and in fact often will cut across one or several other types. DANTERM is a typical example, as is the important and long-established EWF Bank of the Technische Universität Dresden (GDR). Often, banks sited in universities will work in close association with national language planning agencies, standardisation bodies, professional associations or national academies.

There are other possible axes of description, which we have by no means exhausted. However, for present purposes, it is enough to realise that 'termbank' covers a number of widely differing concepts.

Since our 1980 survey, we have noted an increasing level of terminological activity in the Nordic countries, based on termbanks. At that time, there was a major termbank at TNC in Stockholm, supporting the TERMDOK initiative, and a nascent termbank, DANTERM, at the Copenhagen School of Economics. Both these banks have been described at this forum in past years. Since 1980, TNC's termbank has moved from being largely a means of aiding the production of glossaries to a full-scale online terminological service; DANTERM, which was designed from the start as an online facility, has grown rapidly in size and coverage, and various other initiatives have taken place in the other Nordic countries. Thus there are now termbanks in Iceland, Finland and Norway. A special issue of *TermNet News* (No. 12, 1985) is devoted to

terminological work in the Nordic countries. The Norwegian termbank, NoTe, based at the University of Bergen, has made an especially important contribution to the development and recording of terminology related to the oil industry (Ute, 1987). A small, independent termbank project is sited at Oslo University. The Nordic banks typically record terminology in several languages, covering a variety of domains, and have collaborated closely together over the years. TNC and DANTERM in particular have further developed close links with industry and commerce in their respective countries, DANTERM, for example, sending out terminologists into companies to aid in the elaboration of terminologies which are then made available on the termbank. The Nordic experience is a valuable and instructive one, for it demonstrates what can be achieved with a concerted coordinated effort, both at national and local international level. In many ways, the developments in the Nordic countries are unique, being the result of the terminological needs of small countries, with minority languages, and a strong requirement for establishing correspondences with the major languages of sciences and technology. However, this does not detract from their experience, far from it: we would contend that the developments in the Nordic countries constitute the closest thing to a success story in the history of terminology and termbanks. Nevertheless, we would not wish to give the impression that there is a smoothly functioning global Nordic terminology service. There is not, each termbank in each country has its own set of users, its own purposes, structure and so on. For some time now, there have been discussions going on among these termbanks regarding a common classification for terminological data, and a common data format, which appear to be moving only slowly towards fruition. This alone is an indication of the individual nature of termbanks. However, if there were a worldwide cooperative network (in human terms) on the Nordic model, this would advance the state of terminology and termbanks at a rapid pace. Unfortunately, such a state of affairs is unlikely, even though there have been attempts in the past to institute a global network, and no doubt will continue to be.

Over the last seven or eight years, a large number of small termbanks has grown up. It is not possible to know the extent of such growth with any degree of certainty. Few of these banks can claim a full user service as yet. Some of these banks are, or lay claim to be, national centres. Some exist as prototype systems, other as projects, others as designs. Some have shown promise at the planning stage, and have nevertheless disappeared as a result of political decisions, or lack of funds. Many, especially those playing some kind of central role in a country, fortunately show every prospect of growing into large, multilingual, multifunctional termbanks.

No major (with reference to size) termbank has appeared in recent years, for rather obvious reasons: it takes a great deal of time and effort to build up stocks of terminology, especially for languages which have no great terminological tradition, or access to machine-readable terminological data. It is not normally possible simply to buy or otherwise acquire sets of terminologies from existing

major termbanks and install them: these are inevitably unsuited for local purposes, containing information oriented towards different types of user, which for local use is too narrow or too broad, or undifferentiated, or not of sufficiently high quality, and so on. There are many reasons for incompatibility. A major reason is that the information is highly dependent on the termbank system used to store and retrieve it. Ruhrgas AG have opted for the relatively simple solution of acquiring Eurodicautom software and mounting it on their own computers, thereby making exchange of data with Eurodicautom itself straightforward. Such a solution is not generally satisfactory in other situations. Despite the problems involved in actually building up rapidly to a reasonably broad, detailed subject coverage, let alone the other problems involved in the creation of a termbank, there is a great degree of interest in many countries in setting up termbanks, at both national and lower levels. UMIST has seen a constant stream of visitors from all over the world, in particular from the Arab, Asian and Latin American countries. No doubt UMIST was just one of the ports of call of such visitors, as they sought advice and information on setting up their own termbank. Probably the country that is most advanced at present in setting up termbanks is Japan: there, terminological and lexicographical information in general is seen as playing a key role in aiding the advance of Japanese technology and trade. Japan has invested large sums in promoting natural language processing research, and is currently developing many aids to language processing, in government, university and industrial research centres. Natural language processing systems are ultimately dependent on large stores of high-quality terminology. Terminological work in Japan is therefore geared to producing termbanks that may be exploited by both humans and systems. The most important work is being carried out by the Japan Electronic Dictionary Research Institute, which is supported by numerous Japanese companies. A comprehensive statement of terminology work in Japan is provided by Fujikawa and Ishikawa (1987).

As for the rest of the world, we give here the briefest of details, mainly to illustrate the fast growth that has taken place in the field in the mid-1980s.

In Europe, the number of termbanks has steadily grown over the past decade. France has seen the demise of efforts coordinated by AFTERM, but has recently launched new initiatives within the framework of 'les industries de la langue'. A small, but important, termbank has been independently created at the Université de Clermont II, concentrating mainly on soil mechanics in English and French (Henning, 1985). It has recently become available on the Minitel network, and provides therefore a de facto high-quality national service in this field. CILF (Conseil international de la langue française), besides automated terminological glossary and dictionary production activities, using the ALEXIS software package (Clemencin, 1985), has made available on the Minitel network a database (ORTHOtel) of general vocabulary and associated exercises essentially as a teaching aid, but useful for checking of spelling and style also. GACHOT s.a., who offer a machine translation service on Minitel

using SYSTRAN, plan to exploit the SYSTRAN dictionaries to their fullest extent by converting them into a suitable format for online consultation. An extensive review of terminology-oriented work in France has been recently undertaken by the Agence Linguistique Européenne, whose report (ALF, 1986) also gives details of the degree of computerisation in terminology in the various centres investigated.

In West Germany, the Institut für Arbeitswirtschaft und Organisation of the Fraunhofer Gesellschaft had an ESPRIT project which was projected to lead to the development of the GLOT termbank. Ruhrgas AG acquired the Eurodicautom software some years ago, as mentioned above, and is in the process of building up its own stocks of terminology. It has a particularly liberal and cooperative approach to collaboration with other terminology centres, having offered to make both its term holdings and software available free of charge (the latter subject to CEC agreement), and to process term holdings of others if supplied on magnetic tape.

The termbank of the Technische Universität Wien, Austria, has been growing steadily since its inception in 1979. In Italy, the Gruppo DIMA in Turin has been involved in termbank projects for a number of years now, particularly in collaboration with FIAT. In the Basque Country, a termbank is used as an aid to glossary production at the Unibertsitate Zerbitzuetarako Euskal Ikastetxea (UZEI), the Basque Center for University Services in Donostia/San Sebastian. In Catalonia, TERMCAT (Centre de Terminologia Catalana) in Barcelona has established the BTERM bank. A variety of subject fields is covered, in four languages: Catalan, Spanish, French and English. The termbank will shortly be available for public consultation, as it has reached what is considered a viable size (over 100,000 term records). It is envisaged that this bank will fulfil the role of a national termbank for Spain (Cabré i Castellvi, personal communication). The Escuela Técnica Superior de Arquitectura, Valladolid, is collaborating with the Universidad de Valladolid to establish a small termbank dealing exclusively with construction terminology whose results, in published dictionary form, will be distributed in all Spanish-speaking countries.

In the Netherlands, Philips in Eindhoven has set up its own termbank (K.W, Schneider, personal communication), to allow greater flexibility in its terminology work. Previously, Philips had used the TEAM system of Siemens, being one of the partners in the TEAM cooperative network. Continuing collaboration will be ensured due to use of a term record format which is entirely compatible with that of the TEAM system.

In the Celtic countries, there are signs of growing collaboration in terminology. The Irish standards agency, IIRS, together with the National Board for Science and Technology, has been encouraging the development of terminology projects for some time now. Collaboration has been recently established with a small termbank on the Isle of Skye, in Scotland, and talks are underway with groups in Wales and Brittany. Significantly, the Eurotra

terminology work is coordinated by the Irish Eurotra Group, who are in the process of installing a multilingual termbank for the needs of the project. Ireland in particular is therefore developing much expertise in automated terminology, and we look forward to interesting results there in the near future.

In Czechoslovakia, a Czech termbank is being developed at the Charles University of Prague, at present largely in Czech and English, but with plans to expand to six languages and a wide variety of subject fields (Minihofer and Machová, personal communication). It is used at present as an aid to glossary production, and has attracted the interest of the largest Czech technical dictionary publisher. In Bulgaria, a national service is under development, called TERMSERVICE, at the Bulgarian Academy of Sciences (Nikolova and Nenova, 1982). The Polish Academy of Sciences has likewise promoted the establishment of a termbank, and COINiM, an agency of the Polish Committee for Standardisation, has established a bank of standardised terms. The Federation of Scientific and Specialised Translators of Serbia in Yugoslavia is in the process of establishing a termbank also.

Among the Arab nations, there is a great deal of interest in setting up automated terminological services, regular meetings being held by numerous Pan-Arabic associations to investigate Arab cooperation in terminology. A national termbank (BASM) is under development at the Saudi Arabian National Center for Science and Technology, Riyadh, which has prompted cooperative efforts among various Arab states, and there are other banks in Morocco and Tunisia. The paper by Sieny in this volume gives further details on Arabic activities.

In Israel, the Academy of the Hebrew Language has promoted the development of a distributed termbank in its various centres, including Haifa and Jerusalem.

In the United States, there has been little development in the area of termbanks. An early project at Carnegie-Mellon University foundered from lack of support. The development of a regional termbank for South Carolina was reported in 1983 (Turpin, 1983). A small termbank has been created by the American Association for Medical Systems and Informatics, with a view to supporting applied research in knowledge-based activities, the medical domain being a ‘traditional’ focus of such activities. The US National Bureau of Standards has a computerised terminological information system for use by its experts.

In Venezuela, computerised terminology work is centred on the Banco de datos terminológicos de la Universidad Simon Bolivar (BTUSB) in Caracas. In Mexico, the Centro de documentación y terminología para traductores, Mexico City, has installed the Eurodicautom software.

In the People’s Republic of China, a small project exists at the Institute for Scientific and Technological Information of China (ISTIC), aimed at automated dictionary production.

Among international organisations, the UN has been promoting the

establishment of new banks and the harmonisation of existing ones in its various agencies, particularly through the adoption of a common term record format. The IMF and the World Bank have well established banks now, and the UN Documentation, Reference and Terminology section in New York has intensified computerisation of TIBPS, its terminological information service. The World Meteorological Organisation and the International Telecommunication Union, in Geneva, both have important, though modest, machine-readable collections of terms in their primary fields of interest. Multinational computer companies have instituted automated services for their translators and technical writers. IBM, for example, have such systems in their Paris and Montreal offices.

A growing number of translator schools are investigating the possibility of developing termbanks and other machine aids to translation as a means of training their students. Leaders in this area are the schools in Geneva and Maastricht. The latter school further intends its termbank to function as a means of supporting a network of former students in a translating cooperative.

This rapid survey is meant to be indicative only: we certainly do not claim to have accounted for all termbanks, established, under development or planned. It should further be noted that details regarding coverage, languages, user population, modes of access and so on are not given. There are three reasons for this:

(a) much of the information on subject field coverage, languages and size of holdings is hard to come by, largely ephemeral, and in not a few cases the result of wishful thinking or projected plans. It is often in the latter case difficult to separate out descriptions of an actual state of affairs from a desirable state of affairs;
(b) it is questionable how useful such knowledge is, in the light of what will be discussed below regarding relevance of termbanks to translators;
(c) collation and analysis of this kind of information is not possible in the context of a paper of this nature: normally, such an activity would form the data gathering topic of a feasibility study and necessitate reasonable funding over an extended period of time.

As for the United Kingdom, the situation is little better than it was eight years ago. After the initial feasibility study of a UK termbank was completed (Sager and McNaught, 1981a), which demonstrated the need for a multilingual, multifunctional bank covering a wide variety of subject fields, and the existence of a large potential user community, the project ran into severe funding difficulties. No governmental funds were forthcoming, and attempts to set up a consortium of parties to fund development even of a prototype system came to nothing. UMIST over the years has tried several different approaches to solve the funding problem, and has tried to generate interest by developing its own prototype system, on the grounds that people were more likely to support a full-scale project if they could see a variety of facilities demonstrated on a small scale.

In a recent attempt, UMIST coordinated the establishment of an inter-university termbank project, which brought universities with major interests in terminology and translation together with a common aim to set up a distributed termbank project. The universities involved each dedicated a member of staff to the project, and provided modest funds for the purchase of equipment to run termbank software (at a time of more severe than usual financial stringency in the universities, it may be added). As some universities, notably Aston and Bradford, had already developed termbank software of their own, cooperation was centred on establishing a common format for the recording of terminology, the idea being that at least terminology might be stored in machine-readable form against the day when a UK termbank might be created. In parallel with this activity, UMIST coordinated several applications for grants to fund a major termbank effort.

Today, in the face of a succession of disappointing setbacks, the inter-university project is no more: UMIST has ceased dissipating its energies on what is a project of manifestly little importance to funding bodies, and the other universities involved have either given up in a similar fashion, or have attempted to develop initiatives of their own, with no attendant success at the national level. The United Kingdom and the United States, therefore, remain the only major industrialised countries of the world without a termbank of national importance. It would appear that the interest in developing a termbank in these countries is indirectly proportional to the status of English as the major language of scientific and technological communication. The situation in the United Kingdom, with its European outlook, is much worse than in the United States, as there is an exponentially growing requirement for multilingual terminology in this country.

The outlook for the development of a UK termbank is therefore highly pessimistic. One might think that the problem lies with attitudes towards translation and foreign languages in general in this country, and, especially in this forum, one might be easily led into believing that. However, we would do well to remember that the primary task of any termbank is to support the description of the special languages of science and technology of a linguistic community. Normally, this means an emphasis in the first instance on the description of the terms of one's own language (the cases of bilingual countries such as Canada, or multilingual environments such as the CEC, call for different measures, as we have seen). Without a coherent description of the terminological resources of one's own language, there is little point in engaging in a search for term equivalents for other languages. The United Kingdom has patently abrogated its responsibility in describing the special languages of English, and is apparently content to let this be done by other agencies. Of the English language speaking countries, only Canada is in a position to exert control over terminological developments in English. It is clear that the special languages of English will be increasingly influenced by the contents of termbanks scattered throughout the world: invariably, English as the language

of scientific communication is first on the list of every termbank manager or designer, after his or her own language.

Given the commitment of the UK government to achieve advances in information technology, in the form of the Alvey programme, to the expansion of trade with other countries, to the building up of UK industry to compete in foreign markets, it is astonishing that so little attention has been paid to the cornerstone of all such activity: communication of information at its basic level, that of the terms used by specialists. Other countries have recognised the importance of efficient communication, and have accorded language and terminology research in particular prominent positions. Japan has emphasised natural language processing and lexicographical/terminological projects for some time now. One need only look at the high number of machine translation projects under way there, or at the efforts of the Electronic Dictionary Research Institute. In many ways, with terminology, the Japanese have a much harder task than we here in the United Kingdom, but that is no excuse for the blatant disregard so evident in this country. In France, language and terminology have become the focus of a new industry, and we look forward to rapid advances there. Here in the United Kingdom, projects on natural language processing are few; on lexicography, projects are to be counted in low figures in research centres, although UK publishers are among world leaders in the application of computational techniques; in terminology, there are no projects funded by central bodies. The amounts allocated to research into language under the Alvey programme were paltry in comparison to those being invested by other countries in similar initiatives. The centres capable of carrying out multilingual natural language processing work in this country are desperately few. The United Kingdom in short has an inadequately funded, small group of people who have talents in natural language processing or termbanks. It has no major or even minor national project in termbank development, and has no national programme for machine translation, although it participates fully in Eurotra, through substantial funding from the Department of Trade and Industry, and partially funded, through Alvey, the development of a prototype English-Japanese MT system at UMIST (the industrial partner being ICL), which showed much promise, but is now discontinued through lack of funding. Eurotra will not provide everything that a national programme would provide, is devoted to carrying out research and development in areas impinging on machine translation, and in any case is not expected to achieve usable results in the near future, in the sense of industrial developments. The Alvey natural language sub-initiative concentrated, in its short lifetime, rightly, we believe, on research into general purpose natural language processing tools for English. The awaited follow-up to Alvey will concentrate on applications, where such tools are expected to be used to build actual systems. However, in the case of the natural language processing tools, these are not at a sufficiently advanced stage of development to be used in an applications-oriented environment. This is not the fault of the researchers involved, who have produced high-quality prototype

tools, rather it is the fault of funding bodies and government seeking to push forward to applications, while neglecting the effort that is needed in natural language processing to achieve directly usable results. In short, there is a tendency to want to run before walking among governmental, funding bodies, while those who actually do the work are well aware that, if the running shoes are effectively a prototype, untested pair of walking shoes with a sporty label, then they will be inadequate for real situations. It is evident that, in the foreseeable future, the United Kingdom has no real wish to set up an adequately-funded national programme to carry out the required foundational research into natural language processing; no interest in developing multilingual natural language processing aids; no interest in attempting to overcome language barriers, whether internal (monolingual) or external (multilingual), to enable the growth of UK trade and industry; no interest whatsoever in promoting the development of what is in point of fact one of the simplest yet most effective tools for overcoming such linguistic barriers: a termbank.

Both the French and the Japanese have realised the importance of language and of overcoming of information barriers. The United Kingdom is in a unique position, having at its disposal one of the richest assets any nation could hope for, namely the English special languages of science and technology, yet it has failed to exploit this asset. In so doing, the United Kingdom has failed the rest of the world, whose processes of information transfer are largely predicated upon English special languages, on establishing equivalences between English special languages and their own, and it has particularly failed those in the United Kingdom who are at the heart of the multilingual information transfer process, and who mediate and therefore control information, and in so doing control the economic, industrial, technological, and social wellbeing of the nation.

The above digression has taken us away from our main topic; however, it is, we believe, necessary to understand why, in this country, there is no likelihood of a termbank being set up at a national level.

RELEVANCE FOR THE TRANSLATOR

What, the translator may legitimately ask, will this increase in growth of termbanks over the last few years mean for me? How will my daily work be affected by the existence of a bank with exactly the bilingual terminology I need in India, or Sweden? How can I best exploit termbanks in Mexico, Spain and Argentina that all contain terminology I would like to use? The answer is, unfortunately, as one might expect, that the majority of the many termbanks in the world today are inaccessible for the ordinary translator, at least insofar as online access is concerned. Termbanks, like any other databank, are only as useful as their communication links. Some banks are private institutions, in the sense that they serve a deliberately restricted number of people, and are not open to general public access. An interesting variant of this is the TEAM network of

partners, where other companies use Siemens's system in return for supplying terminology. Other banks are genuinely public services, but can only offer a phone-in query service. Yet others are available locally or nationally on telecommunications networks, but are unavailable internationally. Some, very few, are available internationally, if one has the money to pay for the luxury of international telecommunications. Perhaps the most widely available online termbank is Eurodicautom: it certainly appears to be the most popular among translators who actually have access to a termbank, but then that popularity may be entirely due to its accessibility.

As the overwhelming majority of termbank software is derived from information retrieval software (a notable recent exception is the CEZEAUTERM bank in France, with its custom-designed software), the average termbank suffers from all the problems attendant on such software, which are well documented in the information retrieval literature. This has been pointed out on many occasions, including this very forum in 1982, when Kjell Åström presented a highly perceptive paper on aspects of termbank operation (Åström, 1983). He makes the important point that a user has to be effectively an information retrieval specialist in order to get the best out of a termbank, all the more so if he or she wishes to access several termbanks, each with different characteristics and query languages. It is all very well for users to browse in a termbank, or spend some time manipulating the system to achieve their goal, if that termbank is, say, company internal, where connect and search time is free. Even if the translator takes some time to satisfy the query, in any case a great deal of time will probably have been saved in relation to traditional methods. However, if the termbank is remote, and there are telecommunications, subscription and connect costs to pay, then it is not cost-effective for translators to carry out their own searches. In some organisations with in-house termbanks, translators are in fact not allowed to search online: they must pass their queries to a terminologist or termbank operative who will carry out the search for them. In his paper, Åström hopes that termbanks will become more user-friendly, and that a simple to use common or generic query language might be developed, such that individual users could interrogate one or more banks with the minimum of fuss. We have yet to see evidence of this happening to any great degree.

Thus, although, as we have seen, there are numerous termbanks throughout the world, the majority of translators do not benefit from them, in terms of being able to carry out personal online searches, or getting rapidly-produced text-oriented glossaries. Those banks that engage in the elaboration and publication of glossaries or dictionaries are probably among those having the highest impact upon the translating profession at large, as, assuming the translator knows in the first place about the existence of a published glossary, he can buy it and place it on a shelf for immediate access. The obvious disadvantage here is that the user of a published glossary is restricted to the inflexible medium of the printed page. Sophisticated searches to extract the maximum amount of information

regarding a term or a set of terms are just not possible. The best that can be said about such glossaries is that they are relatively up-to-date, assuming an ongoing data collection and publication activity on the part of the termbank. The use of a termbank in this manner to organise glossaries for publication is in essence no different to the activities carried out by a high-quality specialist dictionary publisher who employs computational techniques. A notable exception among termbanks is Termium, which offers free (as yet) online search (excluding telecommunications charges), publishes microfiches of its holdings, and is shortly to offer its holdings on CD-ROM for some $1500 (Canadian), together with suitable software to allow interrogation. Landry's paper in this volume contains further details regarding Termium.

The recent survey of translation practices (Smith and Tyldesley, 1986), although carried out on a relatively small population, nevertheless showed that few translators used a termbank: only 14 per cent of the sample used one. Eight termbanks were mentioned by respondents, but the most popular proved to be Eurodicautom: 8 per cent of the sample used this bank. Further significant results of that report were:

—translators rated terminology work as the greatest 'time waster'. This served to confirm earlier pre-termbank surveys, and one must surely ask the question: what impact do termbanks have on the translating profession in general, if translators still view terminology work as the most time-consuming task? The answer is alas obvious, as we have seen;
—translators would very much like dictionaries to be fully integrated into their office system software.

Despite intensive attempts to encourage termbanks to collaborate more closely, to exchange data, elaborate data together, specialise in particular subject fields and acquire data for other subject fields from partners, to harmonise term record formats, to develop common command languages or simpler, generalised search procedures, and so on, there has been remarkably little success. There are many reasons for this, not the least being that the major banks are large and inflexible, with idiosyncratic practices and associated software systems: after all, they were set up to respond to a particular demand by a particular user group. There is therefore little motivation to change, especially when change is in many cases prohibitively expensive, or may involve adapting to what is perceived as an unsatisfactory solution. Talks have been going on for many years among the major banks to determine ways and means of exchanging data, which have met with very little success, only one or two changes of any significance having been carried out. Problems of term record format, quality, orientation (e.g. whether concept-oriented or not), etc., all play a part in frustrating progress here. An exchange format has been decided upon, named MATER, but this merely lays down the logical characteristics for data exchange, i.e. it provides a template of data elements that banks may adhere to for exchange purposes. However, MATER has nothing to say about the content or interpretation of the record

format fields, and so is of little real help: all it does is to ensure that the receiving bank gets a tape with elements in a certain order, it cannot aid a bank in interpreting these elements with respect to its own holdings. A fundamental problem afflicting the majority of termbanks is that they appeal either to no underlying theory of terminology, or to one which has little relevance in the modern computational terminology world. Proposals for an exchange format which takes into account content and interpretation of data fields were made some years ago, but have been not as yet been adopted (McNaught and Nkwenti-Azeh, 1983).

The situation is becoming worse by the minute even, with the fast growth in termbank numbers we are witnessing today. Having a termbank is rapidly becoming a status symbol, as pointed out by Thomas Schneider, of Siemens:

> 'other [termbanks] will be designed mainly for prestige purposes, and that implies: be different from the other banks, and incompatible. If each man is an island, a termbank shouldn't be, but it seems a hopeless task to keep superfluous projects from starting up if egos are involved.'
>
> Schneider (1987:210)

In the light of the above remarks, the reader may wonder why it is felt to be highly important to establish a termbank in the United Kingdom (even though the financial outlook is pessimistic, one still has hopes). Surely this would merely contribute to the proliferation of termbanks in the world, and increase the isolation of each one? By way of an answer, we may note the following:

—there is no central, authoritative termbank for the English language, which has a pre-eminent status among the languages of the world;
—we have the knowledge and expertise in the United Kingdom to ensure the development of a termbank that will be flexible and compatible to a high degree;
—we have no need of a prestige bank: everyone expects us to have one already, therefore having one will not increase our prestige any. It will be taken as a matter of course that the United Kingdom should have a termbank for the English language. We rather have a need for a high-quality bank that will serve as many users as possible, both monolingually and multilingually;
—we have a population of potential users who are becoming increasingly familiar with office systems technology, who dearly wish to have high-quality mono-, bi- and multilingual dictionaries at their fingertips, integrated in these systems. A UK termbank is the ideal and necessary addition to the equation to ensure that such dictionaries are provided, with guaranteed quality, and up-dated regularly.

Let us look a little more closely at this last point, for it is the key to the future of terminology, the translating profession and a national termbank in the United

Kingdom. While we refer purely to the United Kingdom here, the organisational model we propose is applicable in other countries.

It is our belief, shared by others in the field, that the best way to integrate the translating profession (and other users) and a termbank in a fruitful relationship is:

—to recognise that remote online access is a non-starter for most users, and that such aids as text-oriented glossaries are unlikely to be made rapidly available to the bulk of users on demand, given the overheads of telecommunications alone;
—to pursue a policy of providing glossaries and dictionaries for use by translators on their own machines;
—to encourage software manufacturers, or alternatively press for projects to be set up, to develop adequate tools for the translator to manipulate dictionaries in an office systems environment. Any such systems developed should, evidently, be compatible with respect to data formats produced by the term bank, or permit easy transformation to and from the recognised format;
—to encourage translators to cooperate in the identification, recording, evaluation of glossaries, by providing facilities for feedback, at least.

We recognise that there are many problems involved here, and many others that must also be addressed to bring such a venture to fruition. Perhaps the most important points concern the relationship between translators, software packages for glossary management, and the organisation of terminological work. There are numerous software packages around for the translator to organise his or her own dictionaries and glossaries. Linguatech's TERMEX (Wright, 1986), the multilingual terminological management system, is well established, due to solid and authoritative design by Alan Melby, who is respected both as a computational linguist and as a translator. A salient feature of this package is that it can be integrated with a number of different word processors, INK International's Text Tools is relatively new on the market. Published articles discussing the concept of mini-termbanks include Baudot, Clas and Gross (1984), Gouadec and LeMeur (1984), Gouadec (1987) and Schneider (1987). The details we give here remain brief, as other papers in this volume address this very subject (particularly the paper by Cay-Holger Stoll). However, when we consider the status of dictionary management packages, their use by translators, what do we find? Typically, the translator will use a package such as the ones described to organise his or her own terminological data files. In some cases, one may purchase ready-made files from the companies which supply the packages. However, in both cases, i.e. with translator-created files, and with company-created files, there is a danger, which is signalled by Thomas Schneider:

> 'The main danger in the new capabilities of small computers lies in the fact that an individual has all the tools at hand to work by himself. Of course, autonomy per se is nothing negative. But it is enticing, as many years of translation work in larger offices have proven, to keep one's knowledge to oneself and to disregard the work of others. The consequences, i.e. an abundance of pseudo-synonyms and the creation of terminological garbage, show that more efforts have to be made to shape cooperative attitudes in all persons associated with the field, be they translators, terminologists, researchers or teachers.'
>
> Schneider (1987:211)

What dictionary management packages will therefore achieve is simply better management of a translator's own resources. They will not allow a translator to participate in or benefit from terminology work at any higher level than his or her own facilities, if they continue to be used as at present. However, it is our belief that such systems have a highly important role to play in conjunction with glossaries and dictionaries supplied by a central termbank (or network of specialised termbanks): the termbank is able to ensure quality and consistency of data, through interaction with relevant bodies, other termbanks and the users themselves; the management systems allow the translator to integrate a termbank supplied specialised dictionary (or several, depending on requirements and capacity) with the existing office system software, and to work individually, without having to worry about constant interaction (or lack of it) with a remote termbank.

One final point brings us back to the concept of a national termbank, or network of banks functioning at a national level as a public service. We are in complete agreement with Kjell Åström, who is of the opinion that:

> 'If linguistic databanks have national coverage and fulfil an important need in any particular country's linguistic heritage and development it stands to reason that operations should in part be financed as a part of a cultural programme sponsored by the nation at large.'
>
> Åström (1983:185)

CONCLUSION

Most termbanks, we may conclude, are of little real benefit to the majority of people who wish to use them, for a variety of reasons. A survey of existing and planned termbanks reveals an astonishing degree of diversity, with little hope of international cooperation at present. The only workable solution in a UK context is for the translating profession, and other bodies interested in information transfer, whether mono- or multilingual, to take the initiative in pressing for a national terminological service that takes into account current trends in office systems and the working practices of the professional. Given the status of English as the primary language of science and technology, and the increasing need in this country for multilingual terminology, which are key factors for the development of the cultural, technological and economic life of

this country, steps should be taken immediately by government to fund the creation of a national terminology network, in close collaboration with the main users of such a service. It should be fully realised that such a service may not generate any profit perhaps for years to come from its immediate users, but its indirect impact on trade and industry on the larger scale will be highly significant, and it will thus pay for itself many times over.

REFERENCES

ALF (1986) Inventaire des ressources terminologiques en France. Rapport preparé pour le Ministère de la Recherche et le Commissariat Général de la Langue Française. Agence Linguistique Européenne, Paris, October, 1986.

Åström, K. (1983) Aspects of term bank operation. In: Snell, B. (Ed) *Term Banks for Tomorrow's World: Translating and the Computer 4.* London: Aslib.

Baudot, J. and Clas, A. (1984) A model for a bilingual terminology mini-bank. *Lebende Sprachen 2*, 1984.

Bennett, P.A., Johnson, R.L., McNaught, J., Pugh, J.M., Sager, J.C. and Somers, H.L. (1986) Multilingual Aspects of Information Technology. London: Gower Press.

Clas, A. (Ed) (1987) Vers 1'n 2000: la terminotique, bilan et prospectives/Objective Year 2000: terminotics, state of the art, prospects for the future. *META* Special Edition, 32(2) June 1987.

Clemencin, G. (1985) ALEXIS: système pour la gestion des bases de données terminologiques. *TermNet News* 14.

Czap, H. and Galinski, C. (Eds) (1987) *Terminology and Knowledge Engineering.* Frankfurt: INDEKS Verlag.

Fujikawa, M. and Ishikawa, T. (1987) Topics for 'terminology and machine readable dictionaries in Japan'. In: Czap, H. and Galinski, C. (Eds) (1987), pp. 85-122.

Gouadec, D. (1987) Les horizons de la terminotique. In: Clas, A, (Ed) (1987), pp. 130-138.

Gouadec, D. and Le Meur, A. (1984) Les micro-banques de terminologie. *META* (29) 4.

Henning, J.M. (1985) La banque de terminologie de 1'université de Clermont-Ferrand II. In: Henning, J.M. (Ed) (1985).

Henning, J.M. (Ed) (1985) *Coopérations en terminologie.* Proceedings of a workshop organised by CEZEAUTERM, University of Clermont-Ferrand II. Rapport no. 85/1.

Horecky, J. (Ed) (1982) *COLING-82.* Proceedings of the Ninth International Conference on Computational Linguistics. Amsterdam: North-Holland.

Marcuk, J.N. (1985) Translation of new terms: terminology work of the All-Union Translation Centre. *TermNet News* 13, pp. 32-34.

McNaught, J. and Nkwenti-Azeh, B. (1983) *European Cooperation in Terminological Databanks.* CCL/UMIST Report No. 83/13. Centre for Computational Linguistics, UMIST, Manchester, UK.

Nikolova, B. and Nenova, I. (1982) Termservice — An automated system for terminology services. In: Horecky, J. (Ed) (1982), pp. 265-270.

Oubine, I.I. and Tikhomirov, B.D. (1982) Machine translation systems and computer dictionaries in the information service. Ways of their development and operation. In: Horecky, J. (Ed) (1982), pp. 289-294.
Perstnev, I.P., and Volkova, I.N. (1985) The automated system of the terminological information service. *TermNet News* 13, pp. 30-31.
Sager, J.C. and McNaught, J. (1981a) *Feasibility study of the establishment of a terminological data bank in the UK.* BLRDD Report No. 5642. CLL/UMIST Report No. 81/8. Centre for Computational Linguistics, UMIST, Manchester, UK.
Sager, J.C. and McNaught, J. (1981b) *Selective survey of terminological databanks in Western Europe.* BLRDD Report No. 5643. CCL/UMIST Report No. 81/9. Centre for Computational Linguistics, UMIST, Manchester, UK.
Schneider, T. (1987) Terminotics of the future: three scenarios. In: Clas, A. (Ed) (1987), pp. 204-211.
Smith, D. and Tyldesley, D. (1986) *Translation Practices Report.* Reading: Digital Equipment Corporation.
Turpin, C.D. (1983) A language bank for South Carolina. *ATA Chronicle* (12) 7.
Ute, I. (1987) Technical terminology at the Norwegian termbank. In: Czap, H. and Galinski, C. (Eds) (1987), pp. 357-369.
Wright, L. D. (1986) The Mercury glossary management package. *Language Monthly* 31.

FURTHER READING

For announcements about recent developments, inserted by termbanks themselves, the reader is advised to consult especially *TermNet News.*

AUTHOR

The author may be contacted by e-mail at jock@uk.ac.umist.ccl (JANET) or (mcvax)!ukc!uk.ac.umist!jck (UUCP)