

The translator as information user

Pamela Mayorcas-Cohen

Commission of the European Communities, Luxembourg

INTRODUCTION

As we know to our cost, non-translators have little understanding of the translation process, even labouring under the delusion that translation is a matter of looking up words in a dictionary and writing them down in the right order. Translators are not automatons; in a discussion on machine translation, Andreyewsky¹ noted the error of ‘categorising translators as “wordsorters”’.

Assuming that the basic requisite linguistic skills have been developed, information is the key to high-quality technical translation: information in the form of subject and background knowledge and documentation, terminological information, reference materials and precedent.

The translator needs *information* in order to:

- do the translation at all
- understand the translation
- find terms in the translation
- find references to dictionaries/sources which will help with terms and information
- know where to look for and obtain sources of information on terminology and subject
- know more about changes and developments in techniques
- know more about the evolution and benefits of IT
- know more about professional status
- know more about their own profession

Figure 1. The translator's information needs

THE LESSONS OF TRANSLATION THEORY

If we consider briefly some of the research into translation theory, we find plenty of evidence that the translation task is one where information is at least as significant as language knowledge.

The early theories, whereby language was seen as a series of codes which could be uniquely identified and transcoded into other languages, ideas which influenced the very earliest theories on machine translation (MT), have given way to the view that meaning is bound into the grammatical, syntactical and lexical structures of a language, and that this meaning has to be extracted through a process of understanding and familiarity with the *subject* of the text.

Perhaps one of the greatest benefits brought about by the research into automatic language handling systems and natural language (which according to many translation experts has had little bearing on practical translation problems),² is that it has highlighted just what it is that the translator does. Thus modern translation teaching and training concentrates on sense and content, rather than translation drills.

The transfer phase which occurs between the reading and understanding stages, and composition in the final target language (TL) results from the bringing together of knowledge and information about the source (SL) and target languages. The translator adds or makes use of information on the nature and subject matter of the text.

Thus theorists tell us that concentration on translation theory has been to the detriment of a closer study and resolution of the problems which arise in translation in practice, and reassure us that the failure of translation theory to understand fully the complexities of the translation process has produced the paradoxical conclusion that translation is, at least in theory, impossible. One is reminded of Dr Johnson's view of women preachers.

Hence Margaret Masterman's appeal to translators, also at one of these conferences in 1981, that they should come forward and explain to the computational linguistics experts just what it is that they do, and how they go about the translation task.

Nor is it surprising to learn that at American Translators' Association (ATA) conventions which offer a wide range of theoretical and practical seminars and workshops, working translators regularly eschew the sessions on translation and linguistic theory for the practical hard subject information workshops, generally led by subject experts in such topics as civil engineering or chemistry. While not one word in a foreign language may be discussed, translators can gain valuable insight into the subject and useful guidance on the best reference books.³

The translation process is one which relies on the totality of the human experience of the translator. The whole problem of translation resides in

the lack of experience, the difference between internal and external information. There is a lack of correlation between words and thought. Translators use their experience to look between the two. If they lack the experience or information they cannot do the looking.

A further theoretical definition states that translation is the ‘replacement of textual material in one language (SL) by equivalent textual material in another language, with a view to transferring information from the scientific writer in the SL to the expert in the TL’.

The model in Figure 2 is adapted from the literature on translation theory.² Theorists point out that since the author writes for the reader, and not for the translator, and since the translator does not have the culture of either author or reader, it is extremely difficult for the translator to have a full understanding of signifier A. But, provided the information on which signifier A is based is valid, the translator can use internal and external resources to transpose it into signifier B. (The failure of machine translation (MT) is due to the fact that the computer can only look at the signifiers, and that it has no conception of the significant.) Where information is missing the translator cannot adequately perform the analysis of the SL and synthesis of the TL but has to paraphrase in a way which offers the ‘best match’.

If we apply the model to technical translation, substituting content for culture, and if we say that content derives from information, then if the translator lacks the information available to both author and reader, translation will be impossible.

It has also been pointed out that the translator has to act as a substitute for the author and write as if he or she knew everything the author knew. The problem can be further complicated by the fact that the audience which the original author had in mind may not be the same audience for whom the translation is intended. For example, a report describing a complex piece of research on the health hazards of certain chemicals may be distributed to a working party of government representatives, trade

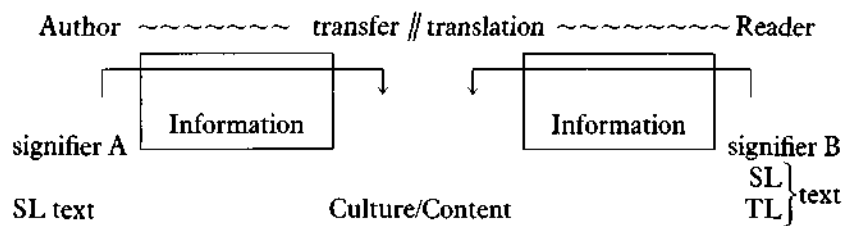


Figure 2. A theoretical model of the translation process

union officials, and manufacturers, in order to determine changes in industrial safety regulations.

THE RELEVANCE OF INFORMATION WORK AND INFORMATION TECHNOLOGY

In the information science literature we find that knowledge or information work involves activities in which the *acquisition*, *processing*, *storing*, *representing* and *communication* of information (knowledge) play a major role.⁴ I think it can be argued that translation encompasses all these activities, which would mean that translators should also be classed as information workers.

The translation process itself necessitates the *acquisition* and *processing* of language information, in addition to preliminary research in order to *acquire* and *process* subject information. Subject and language information is subsequently *stored*, either in one's head, or preferably in some hard and recoverable form.

Representation corresponds to the linguistic conversion process itself, while *communication* is achieved with the end product, the target text. While I may be stretching this parallel somewhat, I think you will allow that these activities are all present in translation.

Knowledge work also involves the use of electronic mail, text processing, teleconferencing, telefacsimile, videotex, teletext and cable TV. The information profession has for some years been debating the merits and demerits of these jewels in the information technology crown. It would be interesting to know how many of these will, in the next ten years, become standard tools for the translator.

Another leading writer and consultant on information problems, Alan Gilchrist, has warned against being prematurely seduced by the new technology.⁵ He provided some useful benchmark tests to help the user to decide whether and in what manner information technology provided real benefits. He suggested that the introduction of new technology occurs in three stages, each of which can only be regarded as a *good thing* if certain conditions can be identified:

- in the *first* stage, it enables you to do the same thing more cheaply, faster and better;
- in the *second* stage, it enables you to do things you could not do before;
- in the *third* stage, use of technology changes your behaviour and ways of doing things to match the new capabilities bestowed by that technology.

I think you will agree that the use of word processing corresponds to the first condition, i.e., it enables you to do the same thing faster and better. Whether or not it is cheaper depends a lot on what system you buy, at the moment.

Coming to the second stage — doing things which could not have been done before. One example is use of the public information networks. It is possible to access information services worldwide: databases for bibliographic references, abstracts and, in some cases, full text; online terminology databanks; library services for finding and ordering reference books; photocopy services and document delivery; and Prestel, used privately or through a library service.

In theory, the fact that such services are remote from the workplace, and no longer only accessible to those living in the major metropolises where such information has always been available in traditional hard copy form, should not matter. In practice, telecommunications costs and the lack of suitable hardware still constitute significant obstacles.

Information technology facilitates the reception and transmission of texts, tables and graphics between translator and client, between in-house department and out-house workers, between translator and reviser, or between the translation department and the graphics department in the same organisation, at a speed which was unattainable using conventional post and delivery services. Unfortunately both file transfer and facsimile transmission still leave a lot to be desired. Nélide Depiante's paper details some of the hazards that can arise in trying to transmit text between word processors, while telefax is only currently viable for relatively short and urgent documents. At a declared transmission rate of 20 seconds per page for Group 3 machines, but which is closer to 30-40 seconds in practice, it can take nearly an hour to transmit a sixty-page document, while the basic equipment costs in the region of £3,000.

In the European Commission we use telefax to transmit urgent documents and translations between requestors and translators, or between translation departments and the European Parliament when in session, and between Community institutions. Since a network of telefax machines has been installed to cope with the urgent transmission of documents between geographically-dispersed offices, especially in Brussels, it makes sense to optimise the investment by using it for translations as well. But the relatively high costs of installation and use may be a handicap for smaller organisations or individuals.

I am also concerned that the use of technology can mask certain inadequacies in the quality of the product. Word processing, a judicious choice of print style and layout, and high-class reproduction can produce work which looks marvellous in form, but which is rather poor in substance.

As regards the third stage (technology changing the user's behaviour),

clearly post-editing of machine translation does require a change in approach by the translator, who has first to acknowledge the level of quality to be provided and then to develop specific techniques for correcting the typical errors generated by the particular MT system.

Electronic publishing systems can be used to produce off line prints from online terminology databanks or from the archives of dictionary publishers, to produce cheap and timely updates to printed dictionaries, or to dispense with hard copy altogether.

Conferencing software could be used between a group of translators in-house or working for an agency to update and consult an editing or technical manual. This would mean that the most up-to-date version would be available at all times and would eliminate the distribution and filing of loose-leaf sheets.

Leaving the machine translation issue out of the discussion, I think that the most significant way in which technology will change our behaviour and way of doing things is that we will cease to regard each translation as a new task to be tackled on an *ad hoc* basis. There are now new things we can do which were not possible before: word processing as an adjunct to the intellectual task of translation for enhancing the appearance of the text; transmission and remote access; information storage and retrieval. The whole approach to translation will change: it will be possible to organise translation on a more systematic basis, and to make far better use of all the know-how that the community of translators has acquired over the years.

It should also create a climate where the client has a better understanding of the work of the translator. In the past, translators were shadowy figures scribbling away in dark corners, with dusty books. They can now be recognised as part of the information and communication challenge to business and industry, using the same kind of tools as are familiarly seen in the general office. This should lead towards a more meaningful dialogue between translator and client, as regards both the function of translation and the support required for it to be done properly.

In 1981 Alan Negus¹ predicted that all translators would be using word processors or workstations as a clever typewriter, and online systems for obtaining information, and that the falling cost of disk storage would enable them to have their own private or local terminology banks. It is in fact taking quite a lot longer than had at first been thought. And I venture to suggest that the problems of price, choice of equipment, compatibility and communications, the rapid developments in techniques and systems, and the unsuitability of many systems for a specialist task are only part of the reason. No one has looked sufficiently closely at the types of information that translators need, or how they want to access it, organise it, store and retrieve it. So no one is really certain how the hardware and

technology can be used. Being able to store large chunks of information is not the point; you need to know how to structure it in a way which matches your particular requirements.

We must be cautious about the advantages of information technology and the likely timescale. At the moment there is no real evidence that the paperless office or remote access is around the corner. While a terminal is now standard equipment for the airline or bank clerk, I would be particularly interested to know how many translators working for the large financial institutions, or indeed for the major telecommunications authorities, have access to the latest text-processing, data-processing reproduction and transmission equipment.

It is nevertheless a curious irony that so much interest in translation has been generated since technology came into it; machine translation has given translation a mystique, aura and status it has not known before. While interpreting and translating is considered to be the second oldest profession, and in terms of reward for effort, we may feel it has much in common with the oldest profession, the modern-day translator has little of the prestige or status awarded to the court interpreter or official translator in ancient or pre-Renaissance periods. It is only in countries which have a bilingual regime, such as Belgium or Canada, or in the 'minority' language countries, such as the Netherlands, that the profession of translator is officially recognised. (See Eveline Sleebos' paper on translators in Holland.) The predominance of English as a world language has certainly been to the detriment of the translation profession in this country and in the United States. Only now do I feel there is a glimmer of improvement; a quick glance at the delegates' list shows the range and variety of firms and organisations which now employ translators.

Unless we can educate first our direct clients, and then other professionals especially in the information field, to realise that we do have a specific need for information, we are not likely to get the necessary help and support in obtaining it. How we pay for it is another matter which I will also touch on later.

INFORMATION AIDS FOR TRANSLATORS – CASE STUDIES

I should like to provide some practical illustration of the aids used by translators, in the form of three case studies. The first is typical of problems encountered in organising technical translation work in the United States, described in a series of articles published in a special issue of *Technical Communication* devoted to technical translation.⁵ I have picked out some of the key points regarded as typifying the ways in which high-quality technical translation can be assured and reflecting some of the problems we have been discussing so far. The other two are

based on my own personal experience working with the European Commission.

Case 1: Ralph McElroy Company

The company, which is one of the largest translation agencies in the US, employing over 100 translators and editors, owes its success to a deliberate limitation of subject area and the building-up of customer information. The company concentrates on work for large corporations and in specific areas of chemistry, medicine, engineering, biology, geology, and physics. The translators get to know the appropriate house style and the political background of the clients, and are able to collect suitable background material.

The best translator is [one] ... who has worked for a particular customer on a particular subject.

It is also considered essential that translators devote time to building up stores of knowledge acquired in the course of their translation work; this includes banal but 'straight, objective information' such as the correct form of trade names, or uncommon abbreviations. Translators must also be equipped with the necessary background material, and standard technical references.

Satisfactory translation can only be assured if translators are able to acquire context and subject information. They may need help with passages which appear to be technically incorrect, which require confirmation from a subject expert, or perhaps the source used by the author for drafting his foreign language document. Where a document cites new technology, it is essential to check with the customer to ensure correct understanding; it may also be necessary to explain linguistic ambiguities to the customer.

The possibility of such consultation is a translator's ideal but is not always possible.

In the opinion of a senior member of the agency:

In addition to his skills as a technical translator, the person should have all the necessary resources for professional translation, and it is essential that he or she should have a technical background in the field in question.

These two qualities are regarded as interdependent and essential to fully professional performance.

The following quote struck a real chord with me, and I am sure you will appreciate it too:

Translation is a skill forever influenced by the winds of change; unlike learning to ride a bicycle, a translator of technical information must continually relearn his or her field based on technological advances and terminological changes. Areas such as plastics, textiles and wood chemistry challenge the translator to be informed, to judiciously and creatively discover the redefined present tense.

That seems to be it in a nutshell! The syllabus of the Ecole Supérieure de Traducteurs et d'Interprètes⁶ in Brussels points out that:

le traducteur travaille un matériau en perpétuelle mutation en constante expansion, si l'on considère le nombre incroyable de termes qui naissent, associés à de nouvelles découvertes.

Both these quotations accurately and succinctly describe a situation and environment which we all recognise to be true, though we often have difficulty in getting customers, or people outside the world of translation, to understand this.

I should now like to come closer to home, both geographically and personally, and refer to my own working experience.

Case 2: European Integration Department (Special Translation Unit), Foreign Office

From 1970 to 1973, I was involved in the translation into English of EC secondary legislation prior to UK accession. I was particularly concerned with the translation of texts on non-tariff barriers to trade, but the team employed at the Foreign Office covered the full range of Community legislation from agriculture to external relations, through social legislation, the iron and steel industries, nuclear technology, transport, industrial and commercial policy, and consumer protection.

It would have been quite impossible to translate those secondary instruments without knowing the full corpus to be translated, or without having sight of all major documents, primary instruments, treaties, and other related Community texts which filled the gaps in our ignorance about the Community political and institutional framework and which helped to explain the Community decision-making procedures. There were special problems to do with Community terminology, particularly as regards the system of law which is quite different from the English one.

Liaison with Whitehall departments, manufacturers and research associations ensured that the technical language used would enable UK industry to implement the new laws properly.

On the whole the people we approached for information were extremely helpful, and appreciated that without their help and

information the resulting English regulations would be unsatisfactory, and potentially difficult to implement.

Case 3: EC Commission Language Service

Turning now to my more recent experience in the Commission's Language Service and looking at the range of subjects which require information, I can single out as examples: harmonisation of technical legislation, nuclear engineering, new energy sources, environment, pollution, food science, measurement and testing, science policy, customs duty evaluations, and educational reforms.

Terminology (for source language and target language)

1. Dictionaries
 - (a) bilingual
 - (b) monolingual
 - (c) multilingual
 - (d) of abbreviations
2. Glossaries
 - (a) internal (in-house; translators' own; group effort)
 - (b) external
 - (i) privately produced (professional and trade associations; specific industries; public utilities)
 - (ii) commercially published
 - (c) part works (published as appendices to monographs, in official documents or publicity material)
 - (d) glossary sections of translating journals (e.g., *Babel*; *Lebende Sprachen*; *Technical Translation Bulletin*; *Banque des Mots*)
3. Online terminology databanks (public and private)
 - (a) bilingual records (e.g., TEAM, Siemens AG, FR Germany)
 - (b) multilingual records (e.g., Eurodicautom - available on ECHO host service)
4. Terminology information services
 - (a) internal/in-house
 - (b) external (e.g., other companies or organisations with translation/information departments)
5. Standard vocabularies and nomenclatures (e.g., ISO; CCITT; IEC; ISTC), customs nomenclatures (e.g., BN; CCT)
6. Card indexes
 - (a) internal (in-house; translators' own)
 - (b) external (commercially available, e.g., Institut Marie-Haps, Belgium)

Figure 3. Support documentation (English Translation Division)

The types of text for which background information is essential are: policy documents and statements, proposals for major multi-annual programmes, tender documents, commissioned studies, minutes, speeches, and official publications.

We are backed up by a substantial Terminology and Documentation Service, which maintains card indexes for all source and target languages, and stocks copies of translations, official publications, and glossaries; the Service provides a phone-in Help Desk and supplies copies of reference documents and subject literature.

Each target language division is divided into special subject groups in

Subject information and terminology

1. General reference works
 - (a) encyclopaedias
 - (i) monolingual
 - (ii) multilingual
 - (b) yearbooks
 - (c) directories (company and trade; international)
 - (d) literature guides
 2. (a) EC legal texts: Official Journal; Budget; secondary legislation
 - (b) EC official texts: Bulletin; Annual Report
 3. Official and legal texts of other institutions: ISO; EFTA; UN; UNCTAD; FAO; CCC; ECC; CCITT
 4. Subject-specific reference works (generally monolingual, and especially for TL)
 - (a) handbooks and manuals
 - (b) monographs and textbooks
 - (c) manufacturers' product and equipment guides
 - (d) technical guides and specifications
 - (e) national and international standards
 - (f) trade literature
 - (g) scientific, technical and commercial journals and periodicals
 - (h) daily and weekly 'quality' press
 5. Consultation of authors, experts or subject specialists
 6. Visits; lectures; films; exhibitions and conferences
 7. Professional bodies and official bodies, especially in TL country
-

order to generate stores of knowledge and build up expertise in a particular area. Over the years the English Translation Division in Brussels has built up a large collection of original English reference material: dictionaries, handbooks, guides, catalogues, government publications, standards, treaties and agreements, specialist periodicals, newspaper cuttings, special supplements from newspapers and the 'better' weeklies, and some highly-specialised subject literature. Figure 3 shows some of the support documentation used in the Division.

Part of the art of drafting Community or indeed any other institutional or policy documents is careful wording and phrasing to highlight or allude to or camouflage a particular issue. Therefore the translator needs to be fully aware of the political significance and context of the text.

Another common dilemma which Commission translators frequently encounter is the request for quick translations, amendments to texts out of context or over the phone, all of which highlight the need for subject knowledge and background information.

Translators working abroad have a particular need to keep in touch with their mother-tongue country; unless one has access to the daily and specialised press, radio and TV, or can make regular visits, it is all too easy to lose touch. Apart from preventing the pollution of one's syntax, it is essential to keep track of changes in usage, in grammar and spelling, the shifts from proscribed to accepted use of jargon or slang, the birth of neologisms. It is also essential to keep up to date with new techniques and technologies, institutional and economic developments, political and legal changes.

Precisely to overcome this kind of difficulty, it was recommended by staff representatives at the Commission that translators should be given the opportunity to spend time in their home country to re-immense themselves in the language, and also to attend conferences, symposia, special lectures, exhibitions and shows. Thus, for example, members of the English Translation Division's agriculture group have attended agricultural shows in order to see equipment, talk to experts, and collect technical and trade literature.

We also have the advantage of French, Italian, and Dutch colleagues close at hand who can help to solve problems of interpreting the source text.

To complete this look at the aids used in support of translation, it might be useful to look at the kind of environments in which translation is done (Figure 4) and a random list of the kinds of texts which translators handle (Figure 5). Information and support requirements will vary, depending on the nature of the organisation, the type of client, and the type of text. The material shown in Figure 3 is typical of that used by all Commission translators, but readers can probably identify material which is of rele-

vance to them, and supplement the list with other material specific to their area of operation.

1. General (no order of size)
 - (a) government departments, institutions and NGOs
 - (b) manufacturing industry
 - (c) trading and advertising sectors
 - (d) scientific and research establishments; research and trade associations
 - (e) data-processing firms, primarily translating software, sales and user literature
 - (f) financial, legal and insurance bodies
 - (g) translation bureaux services
2. Specific organisations
 - (a) the European Commission, together with the other EC institutions, represents the largest multilingual and institutional translation service in Europe, very probably worldwide
 - in 1984, the Commission of the European Communities' Language Service translated some 570,000 pages, equivalent to 171 million million words per year!
 - (b) Siemens AG, probably the largest private language service in Europe (electrical and electronics engineering and manufacturing)
 - over one million pages of documentation per year
 - (c) Bundessprachenamt, the Federal Office of Languages, Bonn, probably the largest government service in Western Europe
 - (d) The Translation Bureau of the Secretary of State, Ottawa, Canada
 - (e) Teleglobe Canada, the Canadian telecommunications authority
 - translates some one million words per year
 - (f) see also this year's list of delegates (see p. 183) for the large number of smaller manufacturing and commercial firms which now have a translation department

Figure 4. Where translation is done

Advertising and sales literature
 Product information
 Instruction and operating manuals
 Training course materials
 Patents
 Standards
 Nomenclatures
 Contracts
 Cost estimates
 Legal documents
 Invitations to tender
 Specifications
 R&D results
 Laboratory reports
 Learned articles
 Reports
 Minutes
 Site and inspection reports
 Product documentation
 Instructions for use
 Books and monographs
 Commercial and marketing information
 Financial information
 Handbooks
 Internal and working documents
 e.g. discussion papers
 minutes and agendas
 opinions
 administrative and staff matters
 Public information literature
 Studies and reports

Figure 5. Types of translation

COST

We have to consider the cost of providing information for translators. We can also raise the perennial chestnut: what is the cost of not having information?

It is almost impossible to put a price on information deprivation. The same is true of information for managers and researchers. It is quite hard to justify to management what the cost would be if the work were not done, what the cost would be if the translation were not perfectly attuned

to the latest idioms and terms of art, if references and quotations were not perfectly cited. However, it should be possible to evaluate the time spent by translators on researching reference material. Surveys have found that translators may spend 25-60 per cent of their time on such research.

After all, a translation badly done, or not done at all could cost a company or organisation hundreds of thousands of pounds in lost revenue or poor customer relations.

Since it is a well-recognised principle of business that time is money, time taken by a translator over weary research in order to track down a vital piece of information which might have been more readily available with the right kind of support is time lost, time wasted, time not spent on producing more translations, getting through the workload, and meeting vital deadlines. While in the case of the freelance translator, such lost time and money is his or her own in the first instance, the client will suffer in the long run, if the work has to be skimped because the translator cannot afford to spend too much time on checking material which should have been supplied by that client in the first place.

If we accept that information and background material *is* an essential requirement, what is the cost of providing it? Well, I haven't produced any figures, but I would like to itemise the main cost issues pertaining to various aspects of support for translation, and offer some considerations on which cost-benefit decisions might be based.

The cost of traditional information sources

Standard reference works, specialised dictionaries, manuals and guides are very expensive, and the problem is further exacerbated by the need to update them with the most recent editions.

Freelancers have to spend this money out of their own income, staff translators have to persuade or educate the powers that be into acknowledging the need and providing funds for such material.

The cost of introducing information technology

It is repeatedly being said that the cost of machines and mass storage is decreasing, while that of software and programs is increasing, as is the cost of setting up a system. Thus the objections raised to storing large quantities of data, for example previous translations or extracts from reference manuals, are no longer valid, provided that someone is prepared to pay for systems analysis, design and installation and user support.

But the larger the collection, the greater the problem of how to organise it, and further time, money and resources will need to be allocated in order to maintain it.

Then again, for translators who do not have access to some form of central storage at their own workplace, the low cost of mass storage is less

relevant than the high cost of installation, of communications equipment, and of telecommunications and connection time.

An average online enquiry can cost between £20 and £30 for a fifteen-minute search which may produce only a bibliographic reference or at most an abstract. The full text still has to be ordered, obtained and paid for.

The cost of equipment

For the individual translator or translation department or agency it is not cheap to install word processors or telecommunications equipment, leaving aside the problems of compatibility with clients or other translation agents, character sets and keyboards, maintenance contracts, special stationery etc.

Are you likely to generate over £7,000-worth of additional income, or save the equivalent in salary costs by investing in a word processor?

It largely depends on how the equipment is used, and it is unfortunately true that a lot of time can be wasted in the beginning getting to grips with the machine. The cost of the annual maintenance contract must be added to the calculation of the initial outlay. But the work will look better, and therefore the client should be more satisfied; this should generate more work and enhance the image of the translation service.

For an organisation, the cost of equipping translators' offices can be absorbed into general office automation costs and overheads. While there is probably strong resistance to money being spent on sophisticated and expensive equipment for translators, company accountants should be asked to look at the cost of translators' time spent photocopying, or on long-distance calls to the author or subject expert, all of which represent lost translation time.

As an added bonus, the microcomputer or word processor can be used to improve the organisation of work, and to keep track of valuable information which would otherwise have to be searched for afresh with each new job.

The cost of terminology databanks and MT or CAT systems

The cost of researching, coding and checking terminology in a computerised dictionary or CAT system has been estimated at some \$32 per word pair, and the major recurring expenditure in an MT system is the preparation of online glossaries and dictionaries. This is glossed over very quickly by the promoters of CAT systems. While it is true that systems will list missing terms, the cost of researching in the first place, inputting, coding and verifying the terms (and each and every word, even if it appears only once in the system, must be dealt with) is very high. It is impossible to amortise this unless it can be guaranteed that subsequent

translations of sufficient size and using very similar terms will occur.

You might perhaps be interested to know that in the second evaluation of Systran, it was found that not even the EC (perhaps the largest translation client in the world) generates enough text in one language pair in one specialised subject for MT to be cost-effective.

The 'time is money' aphorism is a two-edged sword, depending largely on your standpoint, as translator or user of translation, and needs some careful thought.

Time wasted on unsuccessful research into essential information for the translation task is time that could be spent doing more translation. It's a vicious circle. Equally time and money misspent on the wrong kind of research is also time and money wasted. However, time spent on setting up information systems, on preparing and coding term and reference cards, on indexing and filing background documents, and on maintaining the collection, is time saved in the long run.

The cost of redoing translations

As far as the large company or institution is concerned it would seem quite ridiculous to waste money on producing a translation if one already exists.

However, if tracking down the translation and then obtaining it proves to be a lengthy process, if an in-house translator or editor has then to check it and perhaps extensively modify it, it may prove cheaper and quicker to redo the job. This is all a matter of experience, knowing the reliability of the sources available and the likelihood of finding an existing translation in a particular domain and target language.

A company or organisation with its own library or documentation service should be able to carry out the necessary research using the known sources of existing translations, and obtain the translation quite quickly. This can save unnecessary expenditure on the commissioning of translations in a difficult subject area and difficult language combinations for which the results might in any case be unsatisfactory.

For freelancers the issue is slightly different. If they suspect that the translation may have been done before, should they let on or do it again or perhaps crib from the original version – *shock, horror* - and take the credit, and money, for the *new* version? You could say that it serves the clients right if they haven't bothered to make enquiries; or alternatively call it part of the service if they do not have the resources to establish that a translation already exists. If they need the translation, they do not care how it is supplied.

But the freelance translator may decide to be virtuous and tell the client, especially if the existing translation can be obtained quickly and at reasonable cost; he or she could edit it, if necessary, and charge an appropriate fee, plus commission for obtaining the translation.

The result could well be a happy and impressed client who will entrust the translator with further such enquiries. My feeling is that the freelancer would benefit, since this would produce good client relations and more time for further work and further income.

Some time ago, members of the Translators' Guild launched an initiative to pool translations of legal documents since it was felt that there was nothing to be gained by duplicating the translation of highly complex texts. I believe the intention was to establish some system of payment and referral commission for those who notified translations to the register, and those who drew on this pool on behalf of their client. I do not know if this initiative materialised, but there is clearly scope here for both freelance and staff translators, in conjunction with an information provider.

We have an example of such a pooling arrangement at the Commission. An index is kept of the titles of all national legal instruments translated by the Commission's different language divisions. Further titles are supplied by the Ministry of Justice in Denmark, which, for reasons which may be obvious, has to translate a great deal of legislation both into and out of Danish. Copies of the texts, in the original version and in translation, are filed by the Terminology Division. Any new request for translation of a national legal instrument is first checked against this index, to avoid duplication of effort, or to supply translation in the same area or of the same source text but into a different target language.

It is unfortunately inevitable that certain EC texts will be translated concurrently by the EC translation departments, by government agencies, and perhaps by a private body. Where political or industrial confidentiality is involved, different bodies will want to obtain a translation of a particular text without it being known by other parties that they are interested in the document. It is also inevitable that different organisations will need translations at different times, and cannot wait for one of the other interested parties to produce a text.

Some years ago the English Translation Division in Brussels set up an information exchange system with the Health and Safety Executive (HSE) in Derby in order to avoid unnecessary duplication of effort. The HSE translations now appear in *World Transindex* so there really is no excuse for anyone outside the EC institutions wanting an English version of a Community proposal (particularly in the areas of industrial and agricultural vehicles, health legislation, mines and steelworks) not checking this index.

The major sources of translations to my knowledge are: *ASELT*, published by the EC; *BISITS*, published by The Institute of Metals; the British Library's *British Reports, Translations and Theses*, and *World Transindex*, both available in hard copy and online; *SIGLE*, sponsored by the EC; and the National Translations Centre, Chicago. Both the CNRS in Paris and the CNDST in Brussels hold translations registers. Copies of the

translations notified to these registers can generally be obtained from the source body at reasonable cost. There is a good deal of overlapping between the indexes, and increasingly so as more and more data is put onto the various online systems, so it is worth double-checking. You can find further detailed information in *The Translator's Handbook*⁷ and in 'The role of translations in sci-tech libraries', published as a special issue of *Science and Technology Libraries*.⁸

ACQUISITION OF SPECIAL SUBJECT KNOWLEDGE AND INFORMATION SKILLS

There are two issues which seem to require some attention, once it has been established that the translator is an information user, and a user of specialist information at that. First, how is subject-specific information acquired, and secondly, how is the information handled? There is a great deal in the translation literature about the importance of subject specialisation. Can I simply say that I feel it is inevitable that the majority of the new generation of technical translators will have an arts background? It is essential, therefore, that translator training syllabi should incorporate modules on technical subjects and on the best routes for acquiring special subject knowledge, and that this should be followed up by in-house training courses, as well as personal initiative. There is evidence to suggest that a number of training courses in the UK are starting to put a heavy accent on technical learning.

As regards the second issue (how the information is handled), some earlier research carried out when I was preparing an MSc thesis² suggested that only the Canadian, French and Belgian translation schools included information and documentation techniques in their courses. I thought it might be useful to see if this situation had changed since 1981, and sent a questionnaire to a number of CIUTI members (Conférence des Institutions Universitaires de Traduction et d'Interprétation) as well as to some non-CIUTI universities and polytechnics in this country. Unfortunately I do not have space to give you more than the most summary conclusions of this survey.

The responses indicated that while most courses stress the importance of special subject literature, and while some courses afford students the opportunity to become familiar with classification systems and the evaluation and use of reference works, none offered a comprehensive information retrieval module. This is largely due to lack of time, rather than lack of perceived relevance. The nearest thing approaching an IR project on the syllabus is the terminology project which requires students to research a highly technical topic, and to write a 3,000-word explanatory text in the foreign source language, accompanied by a detailed 300-word glossary and

definition list, together with details of the specialist works and sources used, suitably identified and evaluated.

Full details of the results of the survey are being sent to Mrs Margareta Bowen, President of CIUTI, who has kindly offered to circulate the report to all members. A fuller report on the questionnaire and survey is being published in the *Technical Translators' Bulletin*, published on behalf of the Technical Translation Group by Aslib.⁹ May I simply record here my sincere thanks to all the course directors who took the trouble to reply, and in some detail, to the questionnaire.

AN INFORMATION SYSTEM FOR TRANSLATORS

Generally, translators are badly informed about the wealth of information sources available for solving many of their queries, and poorly versed in efficient methods for handling the information that does come their way. Even where a translation department is part of a library or information service, translators are often left to devise their own Heath Robinson methods.

A great deal has been written on the creation of personal information systems, using integrated software. While the manufacturers' claims for such systems tend to fall short of reality, it is also true that end users need to have a very clear idea of what they require of the system.

A translator-dedicated information system would necessitate extensive analysis of the ways in which translators search for, select and use information.

Personal information systems

Do you realise that you have already developed your own personal information system? The dictionaries, notebooks, diskettes, files, card indexes, manuals, handbooks, modified glossaries and margin jottings, names of experts, cuttings, photocopies of earlier translations and background documents, collections of journals and specialist publications, all constitute your own personal information system.

Surveys of working methods have shown that translators can spend 25-60 per cent of their time looking for information in order to be able to produce high-quality, accurate, and reliable translations which conform to in-house, legal and political norms.

By 'looking for information', I would like to suggest that whenever translators have to interrupt their work and either leave their desk or pick up the telephone, it is because they experience an information need which must be met before the translation work can continue. So I would include new information, i.e., on terms or subject content which the translator has not encountered before; information which needs to be verified: formulae

or nomenclatures which can be checked in manuals and handbooks, and terms or content for which the translator needs to go back to original sources in order to jog his memory, or confirm that the translation proposed is in fact correct; and finally, hard subject knowledge or information.

Some of this information will, therefore, already exist in dictionaries, in published or in-house glossaries, reference manuals, handbooks, encyclopaedias; further information can be elicited by consulting an expert, a subject specialist, or perhaps a fellow translator, a professional or trade organisation, specialist library or in-house documentation service. Much of the information required for translating a specialised text will require serious and sometimes difficult and time-consuming research as well as reading of the appropriate subject and background literature. The results of this research can be stored in notebooks or on index cards, on the word-processor glossary file or a small database program on a micro-computer. But pressure of work may preclude more than 5 per cent of it being noted down. Very often notes are scribbled on the backs of envelopes or the last telephone bill! So translators end up carrying a lot of this information in their heads.

Both individual translators and their colleagues (if they work in a team or unit) would benefit enormously if all the mentally-stored information, information which is to be found in the reference tools in the translator's own office or some other library, and the sources of other kinds of specialist information known to be available, were to be systematically stored, and organised and made accessible from one central source. Such information would then be accessible to the translator who had originally unearthed it. It could also be made available to other translators requiring information for the same text, or for other texts for which the same information was pertinent, and it would be very important as source material for new, inexperienced translators.

In the next twenty years, much of the information which is currently available in hard copy, and much of the in-house documentation which staff translators require for their work will be available in electronically-stored form. Worldwide telecommunications facilities also mean that translators working in any one place, or country, for any one firm or organisation, are no longer, in theory, limited to the information which is locally available. Since translators are working between pairs of languages, this obviously has enormous advantages: it will be possible to check original language sources for terminological and subject information in both the source and target languages.

But these electronic mammoths will inevitably suffer from disadvantages of size, and the fact that they may well have been designed for a different type of user, with different requirements.

Our information problems are very similar to those of other users of information systems, namely the authors of the scientific, technical or legal texts which we translate. I suggest, therefore, that what translators need is access to a suitably designed set of software tools which will enable them to set up their own personal information systems.

A few years ago in a paper published in *Journal of Information Science* S. Cooney¹⁰ suggested that the resources spent on producing very large information systems and big databases might rather more profitably have been spent on improving personal retrieval systems for scientists who require the means to store and retrieve their personal notes, reports, reprints (photocopies) and document references.

This seems to be analogous to the world of translation, and probably reflects the views held by many translators, namely that the money and resources spent on automatic translation systems, which in any case are only as good as the information fed into them, could have been better spent on a completely different type of translation aid.

In the case of the individual translators, such a system would comprise their own personal and individually generated information and sources. In the case of a translation company or staff department, the information store would represent a pooling of the resources generated by each individual; this would then be available to other individuals in the group, including new, inexperienced translators.

Considerable effort would be involved in collecting, selecting, verifying and inputting the information, but it would produce a system of real and growing value. Let us not forget that this effort has to be expended in any case by the individual translators, and that their energies would then effectively be recycled since the results of such efforts would be available to them and other translators in a generally available information system. This would result in the effective and efficient recycling of translation energy.

Characteristics of a personal information system

In his paper, Cooney defined a personal collection as being one which comprises approximately 500 to 5,000 documents or references. A recent survey amongst EC Commission translators indicated that each individual might produce between 300 and 600 terminology or reference cards for their own private use. If we multiply this by the total number of translators, and add hypothetical quotas of 20,000 additional documents, files or references held at divisional level, we soon exceed this total. However, the Commission's translation service is broken down into target language units and further subdivided into subject groups. The total number of references held by a group would probably be close to the figure suggested

by Cooney. Thus the conditions for a personal collection exist. Provided the indexing and retrieval characteristics are identical for each of the 'personal' information systems developed for each group of translators or individual translators, it would be quite feasible to provide links to permit the retrieval and interchange of information between systems.

Cooney noted some significant differences between personal collections and general collections which I feel are significant to our environment:

1. The user whose queries are put to the system and who requires the system to perform the retrieval function is also the one who selects the input.
2. The range of documents and the user's range of subject interests for retrieval are generally biased and selective, towards the specific needs and interests of that user or group of users.
3. Personal collections are generally physically and psychologically more accessible: physically in that they may consist of a card file on the desk or set of Leitz files on a shelf in the office; psychologically in that the translator personally has been instrumental in collecting, and indexing the documents.
4. Since personal collections are smaller than large general-purpose databases, the design and method of indexing can be simpler and more accessible to a non-professional information user.
5. Personal collections and indexes are generally kept by people who are amateurs in information techniques.

Does this begin to sound appealing, a system which can readily handle 500 to 5,000 documents, is simple and inexpensive to install, maintain and use, and can accommodate changes of interest and new fields of work?

While this list of requirements may seem straightforward at first glance, they do represent some fairly tall orders. But, by applying some of the basic rules of information storage and retrieval, good results can be achieved.

Decisions as to selection of the documents or records to be put into the system should be left to the individual or group of individuals. Only when the total of the collection starts to exceed the limit which has been defined for a personal collection, would it be necessary to have recourse to external sources for setting up an information retrieval system.

Alternatively, if the total collection is broken down into categories, none of which exceed the 5,000 ceiling, and provided a common indexing language is used for the different categories (terminology, bibliographic references, reprints and original documents, external sources of information), the system will remain personal and accessible.

Personal information software

A number of suppliers of large-scale information retrieval systems have designed micro versions of their information storage and retrieval software, for use on a range of personal microcomputers. The dBase II type of program generally permits only fairly simple record structure, limits the size and structure of fields within a record, and relies on lengthy and slow character-string match and serial searching of the entire set of records.

The new micro versions of the larger database systems are generally faster and more sophisticated, offering classic searching of inverted files, which contain data taken from individual fields and records. They will allow you to specify whether fields are to be of limited length (fixed length fields) or unlimited, and to specify different data structures in different fields. They allow for control or checking tables to ensure that formatted data such as dates, names and subject codes and controlled vocabulary terms have been entered correctly, and in accordance with rules defined by the user. You can use standard vocabulary for certain fields, i.e., in order to specify the type of information (book, photocopy, term card). So, if you know that the information you want on company takeovers is probably in a photocopy of a journal article, made during the last six months, you can specify 'Photocopy', select the appropriate subject code, and limit the search to 'new items entered since 85.06'.

This allows you to preselect certain types of record before pursuing a more detailed search. You can also use logical and value operators (AND, OR, AND NOT, GREATER THAN) or their symbolic equivalents, to combine search terms in a single field or in separate fields; alternatively queries can be put to the system via a suite of customised menus.

You can request screen displays of parts of records, and a variety of printouts, for example, a list of keywords in context as a handy manual reference to the information available, special subject lists or chronological lists.

These systems generally come with purpose-built menus which are used to set up the database, define the structure of the information and the way in which you wish to access it, and with further menus for updating old records. This is a simple and user-friendly approach which replaces the arduous programming required for setting up large-scale databases.

Such systems are not yet cheap enough for the individual, but are certainly within the means of a translation company or commercial or government organisation, especially if they are looking for an information retrieval system for other parts of the organisation. The cheapest are available for between £250 and £500; the top end of the market ranges from £2-4,000 to over £20,000 – the price of a small suburban flat! Further useful information on information retrieval systems and software

can be obtained from the Library Technology Centre, 309 Regent Street, London W1R8AZ.

Construction of a personal classification scheme

Cooney suggested a detailed but, even for a non-specialist, workable method for resolving the vexed problem of all retrieval systems – the method of describing a document in such a way as to ensure (one hopes) that it will always be retrieved when required and never retrieved when not relevant.

Keywords selected for indexing a small sample of documents are listed alphabetically. The frequency occurrence of each keyword is noted. Future users of the system provide a short text, briefly analysing the type of work in which they are engaged, or in the case of translators, the subjects covered by their translation work. Keywords are retained on the basis of the number of documents corresponding to any particular keyword: where the document number is too high or low, the keyword is either further divided, or grouped with others. The keywords are then assigned to between six and ten major topics or facets identified from the work description provided.

Each document or piece of information to be identified in the system is assigned a unique number. Keywords, taken from the personal classification scheme, are attributed to each document. Since the retrieval software described above allows you to have controlled vocabulary fields and free text fields, you can supplement terms taken from the classification with free indexing terms. Used in combination with controlled term searching, this will enhance the performance of the retrieval system. You start by selecting the main topic or facet or combinations thereof, corresponding to the enquiry. The system tells you which keywords correspond to that main heading, and how many documents correspond to each of those keywords. Depending on the data-handling capacity of your system, you can either continue to refine your search in order to limit the final number of documents which match your enquiry, or examine samples of documents from the initial set and rephrase the query.

If this type of exercise is beyond your means, you can use an existing classification scheme, for example UDC or Dewey for general texts, or something like the COSATI list of subject headings for general, scitech subjects. Those working in a very specialised subject area could investigate the thesaurus classifications which have been devised for particular industries or particular sectors of the economy. Aslib, which acts as the UK Thesaurus Centre, can give more information on the types of thesauri available.

It might also be a good idea to enquire into the indexing and classification scheme used by your company or organisation, and to see whether

you could use this or a suitably simplified or modified form to classify the documents and material which you have collected. Since this should be related to the nature of your translation work, it is logical to suppose that the classification scheme used by the organisation's reference and library collections could be applied to your material.

CONCLUSION

This brings me, I suppose, to the burden of my message, which is to urge you all to go and identify and talk to the company librarian or information officer. If you are working freelance, you should make friends with this person in the companies or organisations for whom you work; if you are in a translation company you should establish good contacts with your customers' information departments. By talking to them about the translation work that you are handling for them, I hope that you will discover that you are both working for the same client: the manager or the research worker.

This person needs material which has to be translated from foreign languages, as well as original material which will have been collected and organised by the library or information department. All this material can be of inestimable value to the translator.

By the very nature of their work, translators are inevitably thrown back on their own resources. Time is at a premium. But time spent on organising and setting up an information retrieval system is time saved in the long run as well as contributing to the production of a higher quality product, and improving job satisfaction.

May I end with something that many of you will recognise as my pet hobby horse. I sincerely believe that we should not be using valuable funds to research MT systems.

I do not believe that we shall be able to build a system (fifth-generation computers, inference systems and knowledge representation notwithstanding) which will be able to cope with the subtlety, complexity, ambiguity and constantly changing nature of language. I am strengthened in this view by something which Margaret King, a senior participant in the EC Eurotra project said at one of these conferences in 1981: that in order to compete with human translators' extraordinary capacity for storing complex language conversion codes, their universal knowledge of the world which they apply in order to resolve ambiguity, and the constantly changing nature of language, it would be necessary to feed in (to the Eurotra machine translation system) whole phrases and even whole sentences in order to ensure complete, satisfactory and fast translation by the machine. If you are going to do that, you might as well train more and better translators and develop general-purpose aids to translation and translators.

The expert systems of the future will need to be fed terminological,

subject and context-specific information before they can make the right choices of lexis, register, syntax. So the MT system is itself an information user!

And we must ask, 'Where will all this information come from? By whom, and how will it be collected and verified?' Alan Readett, until his retirement a group head in the English Division in Brussels, pointed out that it is the translator who encounters the unknown word or concept first, not the machine, and it is the translator and terminologist who have to solve the puzzle, before the machine can apply the newly found SL-TL conversion to a text.

The translator's ideal information system will be constructed by translators, information scientists and computer and systems experts working closely together to understand the phenomenon of translation, and will provide hardware, software and telecommunications equipment which will provide translators with all the facilities they need for handling the material they use in the translation process and which matches the way in which they perform the translation task.

REFERENCES

1. MAYORCAS-COHEN, Pamela. Evaluative report: a state-of-the-art review of computer-aided human translation. London: City University, September 1982 (unpublished).
2. MAYORCAS-COHEN, Pamela. MSc thesis: The translator as information user. London: City University, October 1982 (unpublished).
3. *Technical Communication*. Special issue on 'Technical Translation Today', 29 (4) fourth quarter, 1982.
4. JARVELIN, Kalervo. Knowledge work augmentation and human information seeking, *Journal of Information Science*, (5), 1982, pp. 79-86.
5. GILCHRIST, Alan. Patterns of information transfer, storage and retrieval in various types of organisation revealed by recent research, Part 2. In: Taylor, H., ed. *Information management and organisational change: proceedings of an Aslib conference held at the London Tara Hotel on 6-8 April 1981*. London: Aslib, 1982.
6. Ecole Supérieure de Traducteurs et d'Interprètes, Bruxelles, Belgium, Prospectus 1985.
7. PICKEN, C., ed. *The translator's handbook*. London: Aslib, 1983.
8. FEDUNOK, Suzanne. Printed and online sources for technical translations. In: The role of translations in sci-tech libraries, *Science and Technology Libraries*, 3 (2), Winter 1982.
9. MAYORCAS-COHEN, Pamela. Survey of information-gathering techniques taught to translation students, *Technical Translation Bulletin*, 32 (2), 1986.
10. COONEY, S. A standard procedure for generating personal classifications and indexes. *Journal of Information Science*, (2), 1980, pp. 81-90.

AUTHOR

Pamela Mayorcas-Cohen, Commission of the European Communities, PO
Box 1907, Bâtiment Jean Monnet, Luxembourg - 2920.

The opinions expressed in this paper are those of the author, and do not necessarily represent those of the Commission of the European Communities.