Translating and online

Pauline Duckitt

Pharmaceutical Society of Great Britain

Paper presented at the Aslib Technical Translation Group conference and exhibition, London, 20 November 1980

Information can be retrieved by direct interrogation of a remote computer by means of a keyboard terminal and a telephone. The advantages of such an online system are fast access to large quantities of data and the opportunity to refine the enquiry by conversing with the computer. At present, data bases can be used to determine if a translation of a document, or an alternative, already exists. They can assist with translating particular words or phrases, especially in new subject areas. In the future, online systems may be exploited to produce more sophisticated aids, reflecting the structure of language.

THE ENGLISH WORD is 'online' and its French equivalent 'conversationnel'. Although the main topic of this paper is how information transfer via online searching can enhance translation, I cannot resist the opportunity to start with this example of the reverse case—how translation may enhance information transfer. 'Online' merely states that you are in direct contact with a computer; 'conversationnel' implies that two-way communication is possible between you and the computer. The latter is nearer the truth. Online searching can be defined as the process of interactively searching for and retrieving information by computer from a machine-readable database.[1]

Why should interaction be a desirable feature of computer use? Our chairman remarked during the 1978 seminar on 'Translating and the Computer' that 'if translators are to coexist with computers, we must become actively involved in directing their uses, let us be the masters and they the tools'.[2] Using the computer in an online mode helps us to achieve this control. It is well known that computers will only do what they are told. If an enquiry has to be formulated in a very complex language before it can be put to the computer, and this formulation involves computer scientists and a long delay, then *we,* non-computer scientists, do not have control of computers. Online searching allows us to put the questions in simple language. If you happen to phrase the question in the wrong way or if the computer does not have the information you require it will tell you so, immediately and, more or less, politely. You are then in a position to correct the question. Online searching is understandable and therefore controllable.

In this paper I shall be asking the following questions: what is online? why is it useful? how does it help translators? where can you obtain access from? and how much does it cost?

Online information retrieval

For the benefit of those who are not familiar with the concepts underlying online information retrieval I will give a brief explanation and hope that those to whom this may appear simplistic will bear with me.

The user of the online system constructs an enquiry to be matched against the collection of information in a database. Some people have drawn the distinction between databases, containing bibliographic references, and data banks, containing non-bibliographic information. For our purposes today, I will not bother about this distinction and will refer solely to databases, meaning by this structured collections of any kind of information in machine readable form. Depending on the response to the enquiry, the user reformats his enquiry or modifies its scope until he is satisfied with the information retrieved.

Communication with the database takes place via a keyboard terminal with a screen and/or a printer to display the interaction and the results. The data bases exist on computer discs made available by search system suppliers. Typically, each search system supplier will have a computer and will offer access to a number of data bases. The interposition of a telecommunications network between the user's terminal and the computers has revolutionized the accessibility of the databases. Users make what is usually a local telephone call and link their terminals to a network, which allows them to connect to computers anywhere in the world. One network, the International Packet Switching Service (IPSS) links UK users with computers in the USA (and vice-versa) and another, EURONET, links users and computers throughout the countries of the European Community.

Advantages of online

Among the chief advantages of online is the ability to modify the enquiry as the search progresses. Another advantage is the large amount of information that can be stored on computer discs. More access points to the information are possible when it is in this form rather than when it is subject to the limiting factors of the printed form, physical volume and type-setting costs. Thus in a printed book, access to the contents is made via the index; with an online equivalent, every word in the book could be searched. Retrieval of information online is also fast, a matter of minutes rather than hours or days. The information can be kept up to date easily because of the possibility of merging new items into the existing database.

Online as a tool for translators

The crucial question from your viewpoint is how online systems may help translators. Professor Sager in 'Translating and the Computer' outlined the stages of decisions and actions involved in the production of a translation.[3] Online can be used in the first two of these: deciding whether to translate or request a translation, and preparing the rough translation. The data bases which can be exploited for these purposes are diverse in nature. They include the currently small but ever growing number designed specifically for people concerned with translations and the hundreds primarily aimed at information retrieval specialists in specific subject areas. The relevance of the former

group will be obvious but the usefulness of the latter group resides in the extent to which data base producers gather and make searchable multilingual information.

The first two stages of translation referred to above cover five different areas in which online data bases can aid translation.

(1) *Deciding whether a translation is required*

During the production of many data bases, material in many source languages is abstracted by linguists into the target language of the data base, usually but not always English. If the candidate document for translation is published, it is worth checking to see if an abstract exists in one of the online data bases. If so, the information content of the abstract might render a full translation unnecessary. Of course, there is nothing new in this approach. However, whereas previously searching for an abstract might have taken several hours and might therefore have been deemed not worthwhile, an online search can in a matter of minutes determine whether or not an abstract exists. Millions of references can be scanned for the combination of authors' names, title words and journal title which identify a particular paper. Because publicly-available data bases usually cover only published material and because, it must be admitted, the information content of abstracts varies between data bases, the usefulness of this approach is limited yet should not be overlooked.

(2) *Finding if a translation already exists*

The same databases may also be used in finding whether a translation already exists, if it is likely to be in one of the cover-to-cover translated journals that exist in certain subject areas.

More relevant, however, may be the World Transindex (WTI) which has recently been made publicly available online as file 33 of the Information Retrieval Service (IRS) in Frascati, Italy. This is accessed from the UK via the Dialtech Service of the Department of Industry. WTI holds details of translations collected since 1978 by the International Translation Centre in the Netherlands and the French Centre National de la Recherche Scientifique (CNRS). These include translations of scientific and technical literature from East European and Asiatic languages into Western languages and also translations of other Western languages into French. Because the information is online it is easy to provide a large number of access points to the information such as type of documents, target and source languages, publication year, title and subject index terms. Figure 1 shows a typical record.

PRESENT AND FUTURE UTILISATION OF SOLAR ENERGY IN THE USSR
Russian into English
MALEVSKI (I.M.); TARNIZHEVSKI (B.V.)
GBR; Date: 1979;
UK-USSR ENERGY SYMPOSIUM/1979-10/MOSCOW; SUN; Date: 1979;
Type: TS, LM, MV, XC, ZM
CC: 903.J.02
DS: ENERGY/ENVIRONMENTAL SCIENCES

FIG. 1 *Example of a record in the data base World Transindex*

An example of a search would be to find details of translations of articles published in 1980 on solar energy from Russian into French. At present this is a fairly small database but it illustrates the possibilities inherent in online systems for sharing the collections of specialist centres and making them available over a wide area.

(3) *Gathering information in the subject area of a translation*
Information gathering may form one of the preliminaries of a translation if the translator is working in an unfamiliar area or if the subject of the translation is particularly recondite. An abstract of the paper, if it exists, may provide a valuable starting point, even if it is not considered to be full enough to act as a document surrogate. Interrogation of the same subject specific-data bases may reveal a review that will provide the necessary background.

(4) *Terminology*
By providing access to remote terminology banks, online makes its potentially greatest contribution to aiding translation. These data bases are like computerised dictionaries in that they provide equivalents of terms in a number of languages. However, not being limited by having to be produced in printed form, they can also include descriptive information for each term-equivalent pair, including usage samples, synonyms, definitions and grammatical information. Inclusion of the context of a term is particularly important in scientific and technical fields.

The other major advantage is that terminology banks can be frequently updated with new technical terms. Translators have to deal with newly-developed situations, processes and materials. Dictionaries cannot provide this sort of information as the time lag between editions is too long, a minimum of 2.4 years, even in such fast-moving fields as electronics.[4] The alternative is to consult other translators or foreign specialists or research the topic in detail to be able to deduce the meaning, a process which can take up to 60 per cent of the total translation time.[4] So the dissemination of new terminology via online data bases can provide a much-needed aid to translation.

As one of this afternoon's papers will deal with a terminology bank in detail, I will limit any further comments on the subject to mention of the European Community's terminology bank EURODICAUTOM, which is now publicly available via EURONET from the ECHO service of the CEC. Although there is currently much international activity in the area of producing standardized terminology in machine-readable form, encouraged and co-ordinated by Infoterm,[5] EURODICAUTOM appears to be the only terminology bank publicly and easily available at the moment.

The bibliographic databases, however, contain a great deal of multilingual information which can be used in a similar way and have the advantage of being ready now. Databases covering such diverse areas as sociology, engineering and agriculture, (such as Sociological Abstracts, COMPENDEX and Commonwealth Agricultural Bureaux Abstracts), all carry article titles in the original language of publication, each significant word of which is searchable, as well as their translations (see Fig. 2). The pairs of titles show the term in question in context, which provides an important check on meaning. Data bases in languages other than English, if they still contain English titles, provide extra help in the form of non-English abstracts and indexing terms. PASCAL, a multidisciplinary French data base, is useful for such purposes (see

0F041-00533 Forestry Abs
Investigations on the pine weevil (Hylobius abietis). I
Snytbaggeutredningen. I

0F040-04660 Forestry Abs
A collar against Hylobius abietis
Med kragen mot snytbaggen

FIG. 2 *Examples of translated titles in the data base Commonwealth Agricultural Bureaux Abstracts*

PROSTACYCLIN INHIBITS MOBILISATION OF FIBRINOGEN-BINDING SITES ON HUMAN ADP-AND THROMBIN-TREATED PLATELETS
HAWIGER (J.); PARKINSON (S.); TIMMONS (S.)
AFF: VANDERBILT UNIV., SCH. MED., NASHVILLE TN 37232, USA
NATURE: GBR: Date: 1980; Vol: 283; No.: 5743; p.: 195-197; 23 REF.;
Cote: 142; Langue: Anglais Type: TP, LA
CC: 361.B.22.B.02
DS: PROSTAGLANDINE 12/INHIBITION/FIXATION/FIBRINOGENE/ THROMBOCYTE/COAGULATION SANGUINE/METABOLISME/ HORMONE LIPIDE/HOMME
L'ETUDE DES EFFETS DE LA PROSTACYCLINE SUR L'INTERACTION DU FIBRINOGENE AVEC LES PLAQUETTES HUMAINES MONTRE QUE LA PROSTACYCLINE INHIBE LA MOBILISATION DES SITES DE FIXATION SPECIFIQUES ('RECEPTEURS') POUR LE FIBRINOGENE SUR LES PLAQUETTES HUMAINES ET QUE CET EFFET EST PARALLELE A L'INHIBITION DE L'AGRETATION INDUITE PAR ADP OU LA THROMBINE. L'EFFET INHIBITEUR DE LA PROSTACYCLINE PEUT LIMITER L'ETENDUE DE L'INTERACTION PLAQUETTES-FIBRINOGENE IN VIVO ET DANS LA CIRCULATION EXTRA-CORPORELLE

FIG. 3 *Example of a multilingual record from the data base PASCAL*

Fig. 3) and also provides access to language pairs which do not include English. Such databases also provide detailed subject classifications and indexing which are not available for multi-purpose terminology banks.

(5) *Portraying language structure*
Translators are not always searching for term equivalents but sometimes want to find words related in a different way to the one they are starting with. Working within the English language we would go to Roget's Thesaurus or a thesaurus in a special subject area which would guide us to broader, narrower and related terms. Multilingual thesauri do exist and are good candidates for online treatment. Several large volumes are required if all the possible structural relationships as well as alphabetic indexes are

to be provided in printed form. Online to such a data base, you could wander freely through a language, choosing from a multiplicity of entry points and tracing a conceptual path at will.

Of course, this assumes that the relationships between the concepts have been identified in the first place. To quote our chairman again 'it is "ideas" not "words" that we transpose from one language and culture to another'.[2] When organizing any data base for computer searching, you soon find that you have to think very clearly about the ideas behind the information. Every relationship must be made explicit in order to allow automatic processing by the computer.

If you will permit me to do a little star-gazing at this point, I would like to be able to see a time when all translators could have their personal files of information on microcomputers for online access, as some, doubtless, have at present. No longer would finding a piece of information be restricted by the alphabetical order of cards or the number of cross-references that the compiler could be bothered to write out. On the other hand, they would be forced to analyse the relationships between words and the different functions that the same word may have in varied contexts. This would indeed be an aid to translation.

Availability and costs

Coming back to earth, it must be said that online data bases will be no help unless they are readily available and cost-effective.

Translators work in very different conditions: within translation departments of large organizations, with one or two colleagues in a medium-sized company or freelance, often far from centres of information. Are online systems equally available to all?

I am afraid that I would be painting a false picture if I said this were so. Throughout this paper I have been using 'publicly available' to denote the accessibility of a particular data base to anyone who has signed contracts with the relevant system and telecommunication suppliers and who has a terminal and some means of connecting it to the public telephone network. There are many useful data bases which are not available in this way. The only thing that can be done here is to find out what exists and try to encourage the organizations who have developed data bases to share their expertise with others.

To some of you, the conditions for accessing publicly-available databases may be daunting enough. Within large organizations, it is likely that the information department will already have a terminal which can be used. A company with one or two translators might consider buying a terminal and organizing access specifically for translation purposes. The Online Information Centre at Aslib can provide details of what has to be done. For the isolated freelance translator, there is the possibility of using a middle-man, an information broker, to carry out searches on your behalf. The benefits are not as great as when you are present while the search is being conducted, but the process is essentially no different from telephoning a reference library for information. The Online Information Centre can provide a list of brokers.

Cost-effectiveness is obviously important. Please note that I did not say cheapness. The costs of using online systems are clearly visible and may seem high until you

realize, for example, that the time spent finding the words you want is also a significant cost factor. It might not be cost-effective for a freelance translator to have his own terminal until there is a great deal more information relevant to him from this source yet to a larger organization the increased speed of retrieval might make it worthwhile now.

What are the costs? They vary greatly and I can give only an order of magnitude. A simple terminal can be bought for under £1,000 but they can also be hired. Telecommunications charges will be about £2.50 per hour on EURONET or £10.00 per hour on IPSS. Access to the data base may cost up to £10 per hour if it is subsidized or about £30 per hour at commercial rates. Information brokers' rates will vary from organization to organization. Remember, though, that you may need only 5-10 minutes to find your information.

Conclusions

To summarize the current status of online data bases as aids for translators, there exists at the moment a small number of publicly-available data bases aimed specifically at translators or those responsible for the provision of translations. A much larger number of data bases is aimed at information retrieval specialists yet they provide multilingual information which is, at present, underutilized by translators. Availability of equipment and cost of these services may limit use for the time being but as the number of online sources directed at translators increases, as I am sure it will, it will become increasingly cost-effective to go online.

REFERENCES

1 HAWKINS, D. T. *and* BROWN, C. P. What is an online search? *Online* January 1980, p. 12.

2 SNELL, B. M. Introduction *in Translating and the Computer: Proceedings of a Seminar London, 14th November 1978* edited by Barbara M. Snell, 1979 North Holland Publishing Company, p.vi.

3 SAGER, J. C. Multilingual communication *in Translating and the Computer: Proceedings of a Seminar London 14th November 1978* edited by Barbara M. Snell, 1979 North Holland Publishing Company, p.9.

4 MORTON, S. E. Multilingual terminology banks here and abroad; online terminology dissemination as an aid to translation *in* BRENNER, E. H. *comp. The information age in perspective . . .* proceedings of the ASIS Annual Meeting 1978, vol. 15, pp.242-244.

5 SAGER, J. C. *Report of the First International Conference on Terminological Data Banks, 2-3 April 1979.* Vienna: UNISIST, Infoterm, 1979. 18 pp.