THE TRANSLATOR IN THE COMMUNICATIONS WEB

ASLIB Translating and the Computer, 17.
9-10 November 1995

Anna Cordon, British Telecommunications plc.
Advanced Applications and Technologies, BT Laboratories, Martlesham Heath,
Ipswich IP5 7RE
anna.cordon@bt-sys.bt.co.uk

ABSTRACT

Today's translators are already finding the Internet an increasingly useful tool of their trade. Access and hardware are becoming cheaper, and the increasing numbers of services and materials available on the Net are encouraging many more users to join. BT has already launched a number of services and is trialling others, including tools for managing the vast amount of information which businesses and individuals now have to manage in order to survive. Changes in the global economy, together with the increasing usefulness of technology, are changing the way business is done. This will provide new opportunities for translators.

INTRODUCTION

The usefulness, and the user acceptance, of online services have grown exponentially over the past ten years. The Internet, which started in the US during the 60s as a network of university computers, predominantly of interest to computer science departments, is now a network of networks worldwide. With the advent of the World-Wide Web and several user-friendly browsers, the Internet has now become a well-known and increasingly popular tool for business. The number of users is currently over 30 million and is still growing exponentially, with no sign of slowing down. It is not a superhighway - yet. It is more like a national road system, with a few main roads, a lot of minor roads, some unmade tracks and some footpaths. You do not need to have access to all of it to make use of the parts which are relevant to you.

SIX THINGS YOU CAN DO ON THE NET

The Internet offers many useful possibilities to the Translator, of which these are the most common half-dozen:

1- Sending - and sharing

The Internet is used effectively for delivery and receipt of documents, for instance between customer, supplier and subcontractor or between service and end-user. This means that they do not have to be based in the same town or country. A British customer may be able to get translations done in Boston, Mass, just as easily as in Boston, Lines. This means the very best translator, worldwide, can be selected for the job. It may also save elapsed time on deadlines: as I finish my working day, my translator can be starting work on hers. The sun never sets on the Internet.

The traditional distinctions between freelances and staff translators are already blurring in response to economic factors alone, but this trend is accelerating now that technology permits many kinds of worker to work from home or from a remote outstation. Businesses may be formed which, making use of communications technology, gather a team of people who act as employees or partners for the duration of a finite project. The possibilities of these Virtual Businesses, as they are called, which may be established in response to short-term projects or market opportunities, have yet to be fully explored.

Within traditional service structures such as today's agencies or services, large translation tasks often used to have to be split between several translators and then somehow homogenised later. Now, however, groupware enables several people to work at once on a shared file, perhaps even using the Web's audio and video facilities to collaborate on terminology, layout or semantic problems.

A video link with other members of a team or with employers has been shown to reduce human isolation problems arising from working from home (1). This can create synergy in several ways: you can share enthusiasm and as well as information.

The quality of the audio and video facilities is still low but they may be usable even now, At the cost of the local call they create a possibility of collaborating, live, with people on different continents. As bandwidth becomes cheaper in the future, the quality will rise.

2- Finding Subject Information

The Internet is fast becoming a useful source of information on a variety of subjects. Translators need to see terms in context - they can search on a term and see every occurrence of that term which the search engine can find. There are more articles every day, and an increasing number of search engines too. A translator may need to find out whether the technical word they have come across means a small round object like a washer or a large square object as big as a house or is not an object at all but an abstract concept. There may be a page about it on the World-Wide Web, possibly even including a video of it actually in use. So the subcontractor need not risk embarrassment by phoning the client, with all the disclosure problems that can cause. The subcontractor can even visit the customer's premises remotely for background information, by looking at their Web "home page". You may find the answers you are looking for in the 2000 public, academic or corporate libraries on the Internet. There is of course no guarantee of quality, since anyone can say anything on the Net. The very ease with which pages can be created, and messages posted, means that there is a lot of rubbish there. But there may be just the gem you need to save you from wasted journeys to local libraries, or from having to take a guess in order to finish your translation.[1]

3 - Maintaining Your Knowledge

The Internet is now becoming useful for general knowledge - it even carries real newspapers. Some are foreign language material, useful for people wishing to keep up their foreign language skills. There are many pages for teaching languages, though most of these are at beginner level.

4 - Tracking Down Terms

There are several dictionaries on the net, but not many as yet. Dictionaries are costly to build and publishers haven't yet worked out what to do about getting paid for use of them. The most obvious example is Eurodicautom which is big, multilingual and free. It is at *http://www.uni-frankfurt.de/~felix/eurodicautom.html.* There are also a couple of computing glossaries eg The Free On-line Dictionary of Computing which is in the very useful pages created by Imperial College, London, at *http://wombat.doc.ic.ac.uk/.* There are also one or two useful ones on chemistry and spectroscopy.

There is also a number of general, non-technical dictionaries for learning languages. Lesser-known languages are particularly well-served, because universities and private enthusiasts have made them available for the good of humanity.

5 - Problem-Solving

One of the most obviously-useful resources on the Internet is the existence of news groups. There are thousands of these on specific subject areas, including more than a dozen which are of specific interest to linguists. These enable you to ask for help with terminology problems, for instance, or to read other people's experiences of a particular gadget or tax problem. Again, of course, some of the opinion may be worthless, but it is certainly better than always pestering the only other translator you know, for advice!

6 - Finding New Customers

Lastly, you can use the Internet for self-advertisement. The philosophy of the Internet did not initially encourage commercials and certainly if you post one on a discussion forum you are likely to get a flood of complaints from other readers. But narrow bands of advertising often appear now, embedded in popular items such as directories. There is nothing to stop you creating a welcoming and informative Home Page on the Web, where interested people, including potential clients, can find out more about you. This could be a place to put a cv and a sample of your work. The style of the Internet actually encourages the addition of personal details such as photos or information on hobbies. Furthermore you can provide hypertext links which, when clicked on by users, will take them to other pages on the Web, so that they can read more on particular aspects of your service. A translator's page might have links to the page of the university where he graduated, or to some general information, maps, etc about the country whose language he translates, or to the professional institution to which he belongs.

So that is a good half-dozen of the useful things the translator can do right now with the Internet. I believe it will be an essential tool for many businesses within the decade. There are now plenty of books where you can read more about it. BT itself will be publishing one next year (2). BT's best-known Internet service, to date, is Btnet, which offers a range of high-speed and high-quality Internet access options for businesses. Within the next few weeks BT will be launching a dialled access service more suitable for smaller businesses. BT CampusWorld is an educational service already in use on the Net. It carries a wealth of learning material and activities for schools and collages. We are currently trialling Interactive Home services which give our customers access to entertainment services, home shopping and home banking. BT Wireplay will enable multi-user game playing between people who have never met. There are many other multimedia-based services in the pipeline.

FIRST STEPS ON THE INTERNET

For the beginner who wishes to keep an eye on the development of the Internet from time to time, without committing himself, one good way to do so is to drop in on an Internet cafe, where half an hour on the net, including friendly guidance, will only cost him about £2.50. There are more than thirty Internet cafes in the UK. There are a dozen in France, six in Belgium, five in Germany, and of course dozens in the USA. And more are opening all the time. BT occasionally operates a temporary Internet Cafe itself, free of charge, as part of its Local Partnership campaigns in regional cities. At the BT Technology Centre in the World Trade Centre in Cardiff, business people can explore a wide range of multimedia and communications technology for themselves. A number of Multimedia kiosks will be installed in London in the course of 1996. These will use a simple touch-screen interface to allow users to access information services (sport, traffic, weather) and teleshopping.

There are also one- or two-day courses on how to exploit the Internet, run by practically everyone who runs business courses - including ASLIB!

INFORMATION OVERLOAD

That is a brief introduction to today's Internet. Unfortunately, in the process of bringing you whole libraries-worths of information, the Net can easily reduce rather than improve your productivity because, there's just too much to read. I have mentioned the use of search engines on the Internet for finding useful information out of the huge mass that is available. These engines require you to choose key words which, together with expressions such as AND and OR, will describe the kind of article you are looking for. They do not currently find everything, and they often bring you a large number of articles which are not at all relevant. It can waste a lot of your time wading through 90 articles which happen to mention the word Bloomsbury for instance, in search of something about the Bloomsbury set of literati. And then again, maybe the information you want is actually in an article about Lytton Strachey and does not use the word Bloomsbury at all. At BT we have been developing several tools which attempt to solve this problem (3), (4).

THE INFORMATION AGENT

Suppose you are a translator specialising in telecommunications, for instance. You go into the Internet, get a search engine to search on the keyword "laser", and behold, 180 references to read. Unfortunately the first twenty you look at are about laser dentistry, or laser lightshows, or laser surgery. We at BT thought that rather than using keyword searching it ought to be possible to get the computer to find articles which will be of interest to the user BY WATCHING WHAT THE USER WAS INTERESTED IN LAST TIME. We have built a research prototype which does just that. We call it the Information Agent.

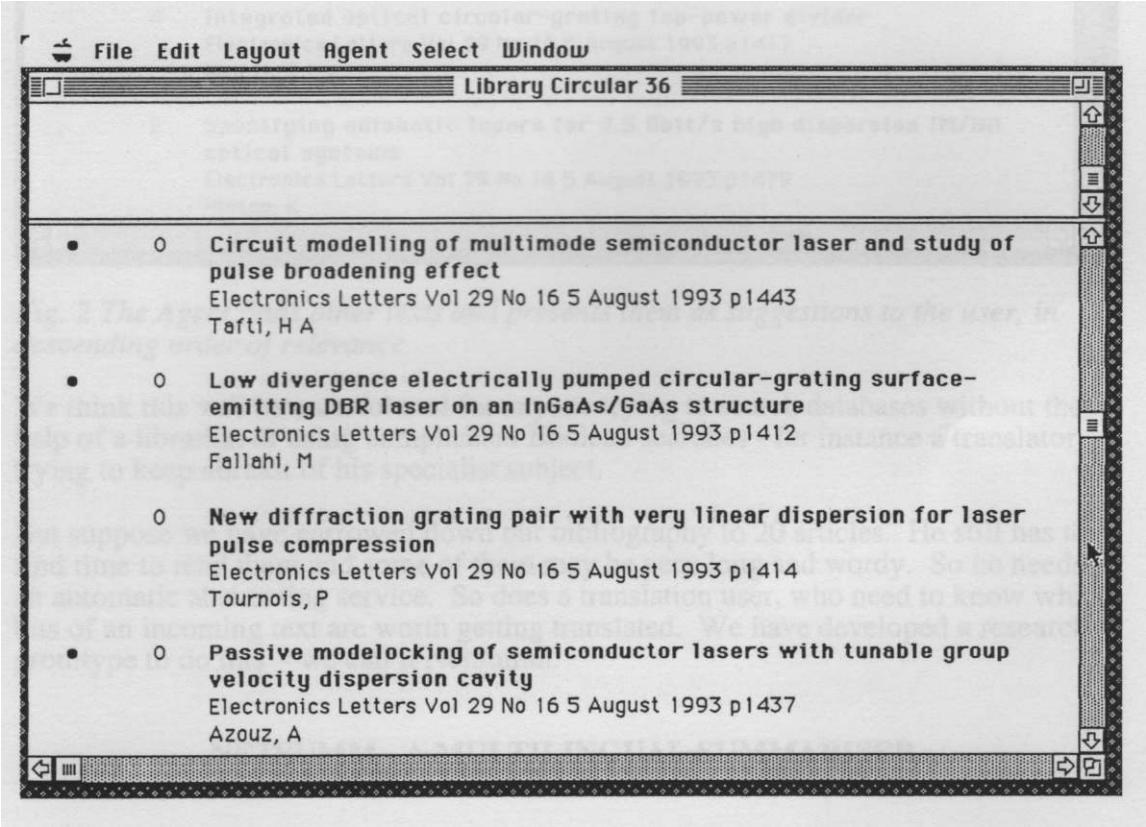

Fig. 1 The Agent observes which items the user has already clicked on (black dots)

This is a page of assorted new accessions received by a library. The user clicks on a few articles she thinks are interesting. She then asks the computer to "learn" those preferences, effectively saying "bring me more like these". The system re-orders the list with what it thinks to be the most similar articles at the top.

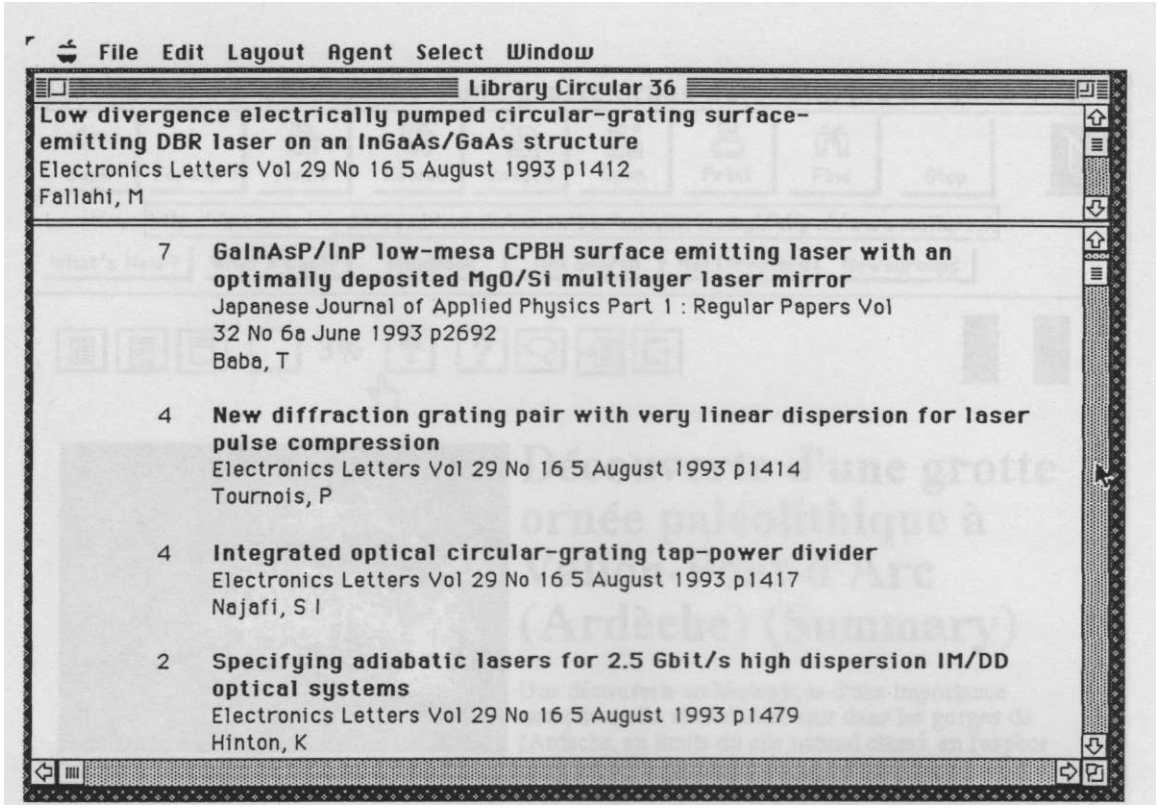

Fig. 2 The Agent finds other texts and presents them as suggestions to the user, in descending order of relevance

We think this will be a useful tool for anyone trying to search databases without the help of a librarian or using complicated Boolean searches - for instance a translator trying to keep abreast of his specialist subject.

But suppose we have narrowed down our bibliography to 20 articles. He still has to find time to read them and some of them may be very long and wordy. So he needs an automatic abstracting service. So does a translation user, who needs to know which bits of an incoming text are worth getting translated. We have developed a research prototype to do this - we call it NetSumm.

NETSUMM - A MULTILINGUAL SUMMARISER

NetSumm uses statistical and linguistic methods to decide which sentences in the article are the most important. You can either get it to highlight the key sentences in situ, or have them extracted and presented as an abstract. You can decide what percentage of the article you want it shortened to - anything from the full text right down to a single sentence. The sorts of text which it will summarise most competently are newspaper-type articles, formal reports or learned papers. It doesn't do so well on advertising material or product lists or creative literature. But these are not the sort of thing you would probably wish to summarise anyway. It is not limited to any particular subject area though. We have had good results on articles about such diverse subjects as Robert Maxwell, particle physics, and the behaviour of red squirrels. It doesn't even have to be in English. It can identify the source text as any one of 20 European languages and will then apply the correct rules for the language it is in. Here's a 3% summary of a multiple-page Internet article on cave paintings in France.

A particular relevance of this to a translation service is in helping customers to prune down the amount of translation they need to commission. A customer who might baulk at paying for a full translation of a long document may be perfectly happy with a 10% abstract.

This is only a prototype as yet but you can play with it on the Internet right now - free of charge, at http://www.labs.bt.com

TELEPHONE INTERPRETING: A WORLD "FIRST"

BT has been researching Machine Translation for over a decade. In 1987 we built the world's first telephone interpreting demonstrator, SALT, which enabled the user to phone a hotel and make enquiries or bookings in a language which they did not know at all. The speech recognition module used keyword spotting to match up what the user had said with a phrase serving the same purpose. The machine translation module then sent the equivalent phrase to the speech synthesiser which uttered it to the far-end user. It enabled a both-way conversation to take place in real time - albeit slowly - with the absolute assurance that the translation was correct. A later demonstrator was built for a different domain - share dealing - and was multiway between speakers of English, German, Japanese and French.

MULTILINGUAL AUTHORING: LINITEXT

Some of the same ideas are embodied in the Linitext multilingual authoring tool. Restriction of the domain and interaction between the computer and the user enable a monolingual business person or a secretary to compose reliable outgoing

documentation in a foreign language. Linitext gets around ambiguity problems by asking the user to select from the various possible meanings of what she has written in her own language. It also uses a menu system to offer pretranslated chunks of language, for instance common business expressions such as

Thank you for.....
Further to your. .....
Please find enclosed. .....

Here's an example of how the user can build up a letter in collaboration with the Linitext system.. The user wants to say: *"I am sorry for the delay in writing to you about your case which you left at our office "*

Linitext has pretranslated this amount *:"I am sorry for the delay in writing to you about. ....."* and only therefore has to process this amount "……. *your case which you left at our office ".*

This makes the translation both faster and more reliable. Any ambiguous words in the text which the user has typed in, such as "case" will echo back to the user, who can then pick whichever meaning she intended - not a legal "case" or "in case of fire", but a "suitcase", and thus a correct translation is ensured. Bilingual secretaries will be spared having to remember the rules on preceding direct objects in French. The notorious complications of the many equivalents of "yours faithfully" are no problem to Linitext.

We are currently exploring the possibilities of networking Linitext, and have built a prototype for Internet use. We are also examining domains other than the business letter domain, such as email and Web pages.

A messaging system inspired by the same principles as Linitext and part-sponsored by BT is in use in the Channel Tunnel and at a number of other points on Britain's frontiers. It enables the Kent police force and emergency services to communicate safely with the corresponding forces of France and Belgium.

CONCLUSION

The aim of this paper was to encourage translators and translation managers to be aware of some of the net's current possibilities, and to give them some idea of the sorts of Internet-based services which may be coming in the future. Some of these may be directly useful to the translation industry. Others may enable new kinds of global business to take shape, and thus create more opportunities of work for translators. What is certain is that the Web is here to stay. As more resources are added to it and issues such as security and intellectual property are dealt with, its usefulness to business is likely to grow considerably in the next few years.

BIBLIOGRAPHY

1. Gray, M, Hodson, N, and Gordon, G: *Teleworking Explained.* Wiley, 1993
2. Frost, A and Norris, M: *Exploiting the Internet,* Wiley, (forthcoming)
3. Preston, K.R and Williams, S H: *Managing the Information Overload.* Physics in Business, June 1994
4. *Information Overload.* The Linguist, Vol 34 No 2, 1995 >
5. Gell, M and Cochrane, P: *Education and the Birth of the Experience Industry.* Educational Technology and Learning Conference, Nov 1994