Automatic gift of tongues: Technology .

More computers are being used to translate documents, and they save money

By JULIAN PERKIN

The galactic hitch-hikers in the world of science fiction author Douglas Adams had a simple way to understand any language they came across - they simply popped a Babel fish into their ears.

Back on earth, things are not so easy. The translation of spoken and written languages was one of the first non-numeric applications envisaged for computers, but initial high hopes were dashed. Quality translation, particularly of general texts, remains to this day an elusive goal.

The past decade, however, has seen a resurgence of research into machine translation (MT). Funding for academic and commercial research projects in Japan, Europe and the US has been stepped up.

The globalisation of business has been a driving force. Complex technical products traded around the world require volumes of supporting documentation, much of which needs to be translated to facilitate sales.

A big problem in translation is ambiguity both in the structure of sentences and in vocabulary. Words often have multiple meanings and functions and rarely correspond directly to words in other languages.

While people naturally make the most plausible interpretation of the words they read, computers calculate every possible rendering of a sentence and find it hard to work out which is the most reasonable.

In many instances, the same sentence may require alternative translations when repeated in different contexts. People draw on a wealth of general and contextual knowledge in understanding and translating a sentence. This is hard to model in a computer.

More problems lie in the fact that, for cultural reasons, the approach to writing varies according to the language used. For example, in English the passive form is frequently used, but seems inelegant if translated directly into any of the Latinate languages. Minutes of meetings are reported in the past tense in English, but are normally written in the present tense in French.

Translation involves more than producing equivalent sentences in another language. Paragraphs or whole sections may need to be replaced. How a computer may understand a text is a matter for philosophical debate. But a model of understanding can be simulated, drawing on areas such as artificial intelligence and linguistics.

Progress in research is being made for specific translation problems within limited and specialised texts. But a higher order of complexity is faced when combining and scaling up prototype solutions to address a broader range of texts.

Commercial MT systems are actively being used already. Surveys by the International Association for Machine Translation indicate that the number of words translated annually runs to hundreds of millions and is growing. Companies in manufacturing, electronics, software and other sectors are using automated translation to support their products internationally.

MT is also widely used in the European Commission and in the US defence department.

Corel, the Canadian software maker, and Oce-van der Grinten, the Dutch manufacturers of copiers and plotters, have both invested in MT systems in the past year. Each has managed to recoup the investment in the first important project.

SAP, a big German software company, has been using MT productively since 1991 and has translated more than 10m words. Daniel Grasmick, head of the company's MT unit, says: 'In urgent cases we can achieve two to three times the productivity of human translators. Over the past two years, we would not have been able to deliver without machine translation technology.'

Technical documentation is an apposite target for translation by computer. Often dry and dull, and requiring scientific or technical competence, texts may be long, difficult and unrewarding for human translators. Terminological consistency is a serious problem, particularly if the job is split between a number of translators.

Most sentences in technical documents are straightforward enough not to catch computers out, and computers are good at using consistent terminology. They are also fast and inexpensive, so long as the text they generate is good enough to avoid excessive reworking in the revision process.

Most of all, technical documentation is voluminous. This offers considerable economies of scale. As John Hutchins of the University of East Anglia notes: 'The market for the translation of technical documentation already outstrips the availability of human translators.'

The need for the assistance of computers is clear.

The leading established MT systems, such as Metal, Systran and Logos, cannot simply be bought off the shelf and plugged in. They need to be attuned to the user's application and a team of linguists is required to maintain the system's dictionaries and database of translation rules.

A service sector is emerging. Translation agencies are taking on the work of polishing up the output from MT systems. This work is often outsourced, and MT specialists or developers may offer to operate and maintain systems for corporate users.

Translation agencies such as Mendes in Brussels and Gecap in Munich offer entire translation services based on MT and geared to specific subject matter. At the other end of the scale, inexpensive translation software packages for desktop computers are readily available.

While these products do not offer the highest quality, they have attracted a sizeable user base due to their low cost - prices range from less than Pounds 100 - and the fact that they are easy to install and use on standard office or home computers.

Colin Brace, editor of Language Industry Monitor newsletter, explains: 'Casual business and home use is where we are seeing the most growth'.

The quality of translation offered is adequate for the required purposes. Rough translations can be used for information gathering, judging whether a

document is suitable for more accurate translation and simple communication via E-mail or letter.

These newly available tools seem to be filling a latent demand for translation. Dan Rootham, who runs a small software distribution company in London, started using translation software three years ago for business correspondence with Russia, Kazakhstan and Ukraine.

'The majority of users are individuals in small to medium-sized companies using MT software packages to translate faxes and E-mails. The heaviest users (of Russian MT) are in the oil and gas industry.' He adds: 'The past 12 months have seen a marked improvement in quality coinciding with driven down prices.'

In line with other parts of the computer software industry, big MT applications are gradually being moved on to smaller, less expensive and more standard computers.

Meanwhile, the software packages are beginning to be founded on stronger methods and to offer more sophisticated tools. Commercial online access to MT systems is now available through services such as CompuServe or via the Internet or by direct dial access. For around 3 cents per word, letters, messages or business documents can be translated. For 10 cents per word, translations can be verified and polished by a translator.

Netscape, the leading Internet software group, has joined forces with Globalink, commercially the most successful translation software developer, and intends to offer instant translations with the next release of its Internet browser.

Computers which translate general texts well, or specialised texts perfectly, may still be some way off. But the potential is evident in the many millions of words already being translated each year by computers, and in the benefits perceived by individuals in businesses using translation software running on desktop computers.