

## Machines take over the drudgery

"The only major project of its kind in the world, the Eurotra project is at the frontier of current human knowledge."

This description of a proposed European system for computerized or machine translation comes not from its promoters, but from the European Parliament's Patterson report on the "problems arising from the multilingualism of the European Community".

To finance Eurotra the European Commission is seeking £6m from EEC institutions and national governments — a small sum compared to the cost of the Community's unique multilingualism and the world-wide investment in machine translation.

The EEC has seven official languages, two more than the United Nations. All are equal. Under the Treaty of Rome, legislation and major documents appear in all seven languages and sometimes also in Irish. Language-related costs — translating, interpreting and back-up services — account for more than 40 per cent of the cost of administering the Community (60 per cent in the case of the Council and Parliament). They totalled £170m in 1979 and are rising constantly, with a 10 per cent annual increase in papers translated, and a greater jump with every new official language.

Last year each of six languages was translated and interpreted into the other five, making 30 language pairs. Now Greek gives seven times six, or 42 pairs, and soon Spain and Portugal are likely to make that 72. For many pairs (Danish to Greek, for instance) enough competent linguists do not even exist.

The reason for this high ideal of Community multilingualism has been well expressed by the Dutch head of the Commission's translators in Luxembourg, Albert Bachrach. Recognize a country's language, and you will perhaps respect that country, rather than invade it; but impose your language on it, and you invade without even noticing.

The cost and inconvenience are appalling. However, almost all alternatives would favour two or more of the big four (French, English, German, Italian) at the expense of the other languages. (An exception is a suggestion that all speak French or English, justice being ensured by making the French speak English and the English speak French.) The Commission was therefore forced to seek more acceptable ways of mitigating the difficulty and have been investigating practical aids, including machine translation.

Machine translation (MT) is a notorious area. The Russian for "we demand peace" was once translated "we require world".

An adverse government report in 1966 almost killed United States research into fully automatic high-quality translation. Yet pockets of research persisted, usually on tiny budgets, kept alive by the need for scientific and technical information, Quebec's demand for French translations, or the translation of Mormon texts into hundreds of languages.

One of the resulting systems, Systran, was bought by the European Commission in 1976. Alone of Systran's customers, the Commission does development work on the system, which has therefore much improved.

In 1977, however, the European Commission decided to use the MT skill in European universities (Grenoble, Saarbrücken, Essex, Pisa and others) in an advanced project, now known as Eurotra.

Some 60 academics from eight EEC countries have been working on Eurotra since 1978. Inside the Commission, Serge Perschke is head of MT; outside, in neutral Geneva, Maggie King coordinates. The group has made some progress on very limited funds, and if the £6m budget is approved soon, hopes to have a pilot scheme ready in 1983 and a full-scale prototype in 1986. The pilot is to translate Commission texts of 10,000 words between a limited number of languages in one subject area. Later Eurotra, like Systran, should be available on the Euronet-Diane information network and to EEC government bodies and universities.

The basic strategy of Eurotra is to have a transfer module for each language pair, but only one input (analysis) and output (synthesis) module for each language (to be plugged into the transfer module required).

Work to date has concentrated on agreeing structures for the interfaces between the modules.

Systran, though not good by human standards, is operational and improving. Why, then, is Eurotra wanted? Maggie King recently gave the following reasons: Systran is not good enough, and may never be: its static and dynamic parts — data and the actions performed on it — are inextricable, faults are hard to locate and to eliminate without unfortunate side effects.

Eurotra, by contrast, will be modular, that is broken up into well-defined sections, each one with its task clearly known. Easier to de-bug, it should also allow new modules (for EEC languages now undreamt of) to be "plugged in" quite easily.

Each Systran system translates only from one source language to one target language, so that the EEC's present 42 language pairs demand a number of these one-to-one Systrans, but only one many-to-many Eurotra. The

projected system will be more readily portable from computer to computer. It will incorporate the advances of the decade since Systran appeared. Better still, its greater modularity will enable it to absorb the results of future research, and this, the Commission believes, will in turn stimulate such research and help Europe to retain a lead in this.

Competition is keen, particularly in Asia, the Soviet Union and North America. And not all research is done on the systems listed in the tables. There is much other investment in MT and related subjects (linguistics, artificial intelligence and computing). It is probably significant that Japan, which is now working very hard on information technology, is

sending 10 industrialists to view European MT this month.

Finally, what of human translation? "Computers threaten ten translators' jobs", a scientific weekly wrote last year — of a system which promptly collapsed. Those of us who get to grips with the computer soon see that good translators have little to fear. Machine translation, such as Eurotra in due, course, will supply fast cheap translations, often to people who would otherwise have no translations at all. By eliminating drudgery and routine, he would free translators to do the difficult, interesting work which they alone can do.

**Veronica Lawson**

translator and consultant,  
London

### Machine translation systems

These or their translations are commercially available.

#### GEORGETOWN (US)

First MT (Washington DC 1954), the 1969 version used still by Atomic Energy commission and till 1976 by Euratom in Ispra, Italy.

#### SYSTRAN (US)

Developed after Georgetown for big IBM 360/70 computer. Relative linguistic advancement. Bases in North and South America, Munich, Japan. \$10,000 a month to lease.

#### LOGOS (US)

Premising. Unfortunately began on Vietnamese and Farsi. Now French, working on German.

#### CULT (Hongkong)

Machine-translates China's maths journals and sells the

printout to libraries world wide, VINITI/VCN — TBD (Soviet Union)

Translates patents. One of various systems in Soviet Union.

#### WEIDNER(US)

Marketed as aid for translators. For sale (from £80,000), lease or rent in North America, Europe and Japan. Just gone public. ALPS (US)

New commercial offshoot of Brigham Young University's MT. Some similarity to Weidner, but with preediting of text.

#### TITUS (France)

Abstracts drafted in limited syntax are translated into several languages simultaneously. Developed by Institut Textile de France, used by various countries

#### METEO (Canada)

Translates 85% of sentences in weather forecasts.

### Who has them

Present Systran and Weidner installations, including some still experimental but excluding those undergoing only preliminary tests.

#### LUXEMBOURG

European Commission (see text)  
Informalux (new bureau service, also for Belgium)

#### GERMANY

BOTS (translation agency, Munich)

#### UNITED STATES

Government agencies (Army, Air Force)

Xerox Corp (Webster)  
Inter-American Development Bank (Washington)  
Princeton International (translation agency, Boston)  
Siemens (Florida)  
Computer Science Corp (time sharing, Los Angeles)

#### CANADA

General Motors  
Bell-Northeastern Research Computrans (offering bureau Service)  
Simpson Sears (stores)  
Mitel (telecommunications, Ottawa, in electronic publishing system)

#### JAPAN

Bravice International (translation agency, Tokyo)