

## **Eurotra Denmark: Diversifying**

**Another in a series of occasional articles on Eurotra, the European Community's Machine Translation project.**

The beginning of 1991—for some the end of the previous Eurotra budget—saw Eurotra Denmark rebaptized as the Centre for Language Technology (Sprogteknologi) and cast as an independent (non-profit) institute within the University of Copenhagen. Currently, the Centre is funded by the Danish government to the amount of DKR 3 million per year (US\$ 500,000), initially for a period of five years. This sum provides for the necessary administrative infrastructure as well as for funding some basic research. Seventeen people are now employed at the spacious and well-equipped Centre on the University campus.

The group is involved in a variety of activities, including consultancy work in the areas of computer-aided translation, terminology management, and document processing, explains Bente Maegaard, the Centre's affable director. There are also two ongoing research projects funded by the Danish research council, which concern translation theory and dialogue systems. The group recently completed a feasibility study for a Danish company which resulted in the signing of a contract at the end of December for what Maegaard describes as "a customized and application-oriented MT system which exploits Eurotra know-how." More details on this noteworthy development should be forthcoming in early 1992.

The group also gives courses and lectures on MT and language technology. The Centre's Annelise Bech has spoken to groups around Denmark at the behest of LAK/HK, the highly active and progressive division of translation-related professionals within Denmark's largest trade union, HK. Her message to them is that Machine Translation will not put anyone out of a job but that one day soon it may be one translation tool of many which are available. Currently, the Centre and LAK are arranging the first MT and language technology forum, to be held in Copenhagen, on February 4, 1992.

Also at the Centre, Uffe Sonne Svendsen is developing a grammar checker together with Eurotra colleagues in Italy. He is working on one for English while his colleagues, naturally enough, are developing an Italian version. The first step for the Danish/Italian team was a survey of the existing English language grammar checking software for the PC. Having evaluated their capabilities, the team has concluded that they could probably produce a better one.

Not surprisingly, the grammar checker Svendsen is developing draws heavily on Eurotra techniques, right down to the terminology used. As the prototype system parses a sentence, it returns terse error messages such as "verb does not agree with second argument." (In Eurotra "case" lingo the subject, object, and indirect object of a verb are, respectively, its first, second, and third arguments.) Svendsen apologizes for the leisurely tempo of the prototype he demonstrates on a workstation, saying that it is currently written in Prolog. A production version for the PC, he adds, is being written in C to bring it up to speed. It would also receive an elegant user interface.

### **Performance anxiety**

In interactive software such as a grammar checker, performance is not a purely technical matter. Svendsen says that while they previously discounted the idea of tailoring versions for specific groups of non-native speakers, having to check all of the potential errors made by all non-native English speakers considerably slows things down. "In certain languages, such as Italian, you can leave the subject out—it's implied in their language but must be explicit in English.

It takes a lot of CPU time to determine the subject.” The best solution for the Danish, French, German, etc., who are writing in English may be to produce different versions of the grammar checker for each. If Svendsen and his colleagues can exploit Eurotra research to produce robust and efficient end-products and succeed in attracting development and marketing partnerships, the suppliers of similar products could face some surprising competition.

Alongside all of these other activities, work on the Eurotra system continues (“co-financed” by the EC), albeit in a more modest fashion than previously. Maegaard explains that, in theory, each group should be working all the languages at once. “We’ve had people working here on nearly all of the languages at one time or another. But now, we are just concentrating on English, French, and Italian.”

Reflecting on Eurotra as a whole, Bente Maegaard says, “We’ve had problems with the software, problems implementing the formalisms. The system is adequate for research purposes, but it will have to be made more efficient. “Eurotra has been a great stimulus,” says Maegaard. “It’s dispersed linguistic knowledge throughout all of the countries involved. We have shared ideas and become more scientific, and what we’ve created is state-of-the-art. Moreover, we are now beginning to see applications for the research results. This is certainly true for Danish, English, German, and Italian—and maybe other languages as well.”