European Commission

A Translation Service with a View

A special feature of the MT Summit V held in Luxembourg last July 10-13, was the EC's invitation to all participants to attend a presentation of their Translation Service, its organisation, computer environment and multilingual translation aids.

**report by
Thierry van
Steenberghe**

In spite of the burning sun inundating the Kirchberg heights, many crossed the deserted lanes that separate the Parliament house from the Jean Monnet building. In the cool conference room, *Dimitri Theologitis* was in preparation; going over the final details of the slide-show with his team. Theologitis, a civil engineer, once in charge of the "Rationalisation" sector, has now become head of the "Development of Multilingual Computer Aids" unit of the Translation Service.

He treated the participants to a brilliant presentation in the form of a computer-based slide-show giving an overview of the SdT (the much used French acronym for Service de Traduction), and of the document life-cycle.

The SdT is probably the largest civilian Translation Service in the world, with its present staff of some 1720, of which 1020 are translators, and the remainder divide equally between support staff (terminologists, documentalists, developers and computer people) and secretaries. This figure has been steadily growing with successive enlargements of the Union. In addition, a large pool of freelance translators allows for the flexibility required to cope with peak loads. These translators are selected through Call for Tenders procedures and work with three year contracts.

The staff is organised in "thematic groups" in charge of the various EC activity domains: five are based in Brussels, two in Luxembourg. Orthogonally to these groups, divisions corresponding to the working languages create a classical matrix organisation scheme. Besides, specific support groups take care of planning resources, information technology, and modernisation. Other groups called AGL (Affaires Générales et Linguistiques) are in charge of what the name implies: language coordination, training, terminology, computer aids, freelance management.

The "clients" of the SdT are the potential authors among the 15 000 EC staff and, indirectly, the 360 million potential readers of the remarkably abundant EC-produced literature. With 15 countries, the EC now deals with 11 working languages in order to meet "internal" requirements. In addition, and out of politeness towards external governments, the EC also tries to handle a number of other languages as

well, although to a much lesser extent. This explains the 1994 "production" figures: 250 000 pages "in" and about 1 000 000 pages "out".

That same year, for the first time, English just overtook French as the single most common source language, with German a solid third. Target languages are, of course, more equally distributed, due to the politics of equal treatment of national languages.

Dimitri Theologitis then went on to discuss the document life-cycle. After its creation (hopefully on a word processor), the document is sent to the SdT using POETRY, a program for the Processing of Electronic Translation Requests.

An electronic form gathers administrative data, language(s) and deadline requirements and is then sent to the SdT by e-mail (X400) together with the source text. Even though most source texts still arrive on paper, the percentage of electronic versions has grown from 5% two years ago up to one third now.

The SdT-preferred word processor is WordPerfect for Windows, although different EC services can still choose their own alternative, leading to much conversion work. Actually, the texts arriving at the SdT are about half in Word Perfect for Windows and half in Word for Windows. Macros are widely used, eg. for the selection of specific keyboards, including Greek and Cyrillic ones.

The translator's workstation interface and the organisation of the standard file structure was then detailed by *Tim Cooper*. He also described the networking of the various computer resources available to the SdT.

As for translation aids, Theologitis said three modes were available. The fully automatic mode (MT). i.e. Systran, is used for raw translation from one of four currently available source languages, often directly by the users themselves: this will be made even easier thanks to the new ERAMIS front-end interface in development. Translators also use Systran to get Celex (the law terminology database) or Eurodicautom to check standard terms. In batch mode, the repetitive phrases appearing in many fairly standard documents are marked to appear in red on the screen, so that only the variable part is translated by overtyping the black text: this is said to eliminate up to one third of the translation work. The interactive mode was illustrated by *Charles Coperland*, who explained how the source-to-target mapping could be done as a function of the input text type, possibly using Celex or Eurodicautom preprocessing. He also discussed the relative merits of term analysis versus full text search techniques.

Dimitri Theologitis went on to detail the SdT support services tasks, before concluding with some views of the future. Specifically, he mentioned the local terminology management software due for completion by the end of 1995, as well as a full text search facility for both translated and original text to be made available soon after. Full EURAMIS services are expected to be operational by mid-96, whereas a full Translator's Work Bench, including advanced tools such as the Canadian TransCheck, is forecasted by the end of the year.

Finally, our host invited further questions. However, the lively and fascinating presentation and demonstrations were so self-explanatory and complete that, beyond applause, only a few questions arose. Or were there too many?

Thierry van Steenberghe works at RIL - University of Louvain, Louvain-la-Neuve, Belgium