

Summary of the Panel Discussion
on Sessions 1 and 2

Chaired by Loll Rolling

The discussion was opened with a question from the chair to Peter Toma on why he had decided to give up his life's work on Systran to take on new activities in New Zealand. Dr Toma explained that he felt the question of working towards world peace was one that went far beyond his aspirations in machine translation, one indeed which necessitated the kind of positive action he was prepared to take. He would however always be available in an advisory capacity for anyone wishing to benefit from his experience of Systran.

In answer to a question from Mr Sereda on the number of users in Europe, Mr Pigott explained that the Commission had concluded agreements with Kernforschungszentrum Karlsruhe, Aerospatiale and NATO as well as with bureau services such as CSATA (Italy), ECAT (Luxembourg), Gachot (France) and Orda-B (Belgium) which covered rights to use Systran in the public sector. German-English versions of the system had been installed at Dornier (by Systran Institut) and at Festo (by WTC). In addition, the Commission was currently processing a number of new requests for access to its systems from several organizations, most of which were present at the conference.

A number of questions were put to Dr Habermann on the use of Systran at KfK for providing raw machine translations directly to the end users. Dr Habermann explained that the number of not found words was still decreasing but had already reached a level of only about one word per page. The number of input errors depended on the accuracy of the source text and careful checking when input via OCR was undertaken. In general, the users - who were subject field experts - were extremely satisfied with the results obtained and were usually in a position to detect the errors which occurred. In some cases they called on the help of colleagues with some knowledge of French.

On the subject of the suitability of the Commission's dictionaries for bureau service, Mr Mendez commented that for some technical fields to be covered adequately a certain amount of pre-coding was called for. Mr Siebenaler estimated that for fields related to the Esprit project (data processing, telecommunications) around 4,500 terms per language pair would be required.

As regards the volume of translations submitted by the Commission, Mr Houston estimated that Cara was now processing about 200,000 words per month. He had encountered no difficulties with character sets, mainly because the Commission ensured that the correct character conversions were made at the time of job submission. Mr Pigott explained that this entailed conversion work on texts originally input on devices such as Olivetti and Philips.

Mr Sereda considered that one of the more complicated areas of dealing with input/output compatibility was when interfacing with photocomposition. However, EDS had now developed a fairly reliable system of preserving photocomposition codes for reinsertion into the target translations.

- 2 -

Mr Jean Gachot presented his plans regarding the use of Systran in France via the Minitel network. With an overall potential of 900,000 users, the system was already proving very popular. At any one time up to seven or eight people were accessing Systran in parallel. However, many extremely common words such as bonjour, monsieur, madame had been found to be missing from the system and were now being coded up.

Mr Sanford (ILCA), who had been trying for years to find out who could provide him with access to Systran in Ethiopia, was informed that since Gachot S.A. had acquired rights to all the European language combinations (with the exception of English-French and French-English they shared with Systran Corporation, Tokyo), this organization would now be in a position to let ILCA have access to Systran services or systems.

In answer to a question by Mr Paez from Informalux, the Commission's current development and maintenance contractor, on the enhancement in Systran quality over the years, Mr Evans (English translation) explained that there had indeed been considerable improvements. In some cases the dictionaries could even be said to be too accurate - generality often served to maintain an element of ambiguity - while text formatting problems had been solved for quite some time. Absurdities did nonetheless continue to occur in raw MT from time to time.

Mr Rudorfer (Dornier) stressed the necessity of developing technical dictionaries for improving the quality of output for German-English. He had found the topical glossary approach to be particularly attractive. Mr Pigott explained that the Commission had tried to reduce the use of topical glossaries to a strict minimum in an effort to provide the highest possible standards of quality to the greatest number of end users. Meanings in context could usually be ensured by sophisticated contextual coding even in highly technical areas. This approach had however not been implemented in the German-English system which was not yet one of the Commission's Systran language pairs.

Mr Holtz stressed that the Commission had coordinated policies both in the general field of data processing and now, more specifically, in the computerization of the translation services where the Sysling project (which also covered internal Systran use) was to be based on networking via a number of UNIX minicomputers.

In reply to Mr Hafner (DB) who wondered to what extent the Commission supported the extension of the Systran family, Mr Rolling explained that special efforts were being directed towards making Systran available to the other European Community Institutions.

Finally, Mr Fiamozzi stressed the need for introducing operational standards in an effort to achieve compatibility between various types of equipment. The Commission was not, however, in a position to undertake further studies in the area of standardization.