

SUMMARY RECORD OF THE WORKSHOP ON WORD PROCESSING

Mr Pigott, chairing the session with Mr Lavorel, began by introducing the many aspects of word processing which might be discussed at the workshop: input, OCR, communications, post-editing, printing, etc.?

The question of input seemed one of the major concerns as it involved a number of problems.

Mr Trabulsi explained the technical problems encountered at Gachot S.A. Optical character readers were not always the best solution as they gave rise to ambiguities and also posed problems for connection to personal computers. Gachot's policy was not to force people to buy prescribed types of equipment but recommended that, as far as possible, they should buy compatible equipment.

Ms Genesio (Orda-B) informed the meeting that conversion problems were the biggest hindrance to the smooth operation of Systran. Input harmonization was difficult to achieve.

Mr Pigott explained that similar problems were being encountered at the Commission and much interface work had to be done on non-compatible machines. Mr Lavorel asserted that Systran highlighted these difficulties and called for a careful study of the problem. Mr Fiamozzi agreed that compatibility was of the utmost importance. He stressed that manufacturers should be encouraged to develop compatible standard equipment. The Commission's EDP department were developing interfaces to make Systran accessible from various types of equipment and were attempting to introduce fully automated networking of the Systran environment.

The question of keyboards was discussed. At the Commission, it was recognized that staff should be able to work on the keyboard of their choice. Mr Lavorel emphasized the importance for inexperienced typists of having a keyboard on which the letters corresponded to what actually appeared on screen. Mr Pahl (Codework) spoke of various new developments in keyboards including the Teletext system which permitted the changing of the actual keyboard and not just the keyboard program.

Various technical difficulties in the translation process were discussed at length. Mr Trabulsi explained the importance of tabulation and the corresponding establishment of translation units for the successful running of a Systran translation. Mr Pahl reiterated the importance of the beginning and end of sentence for Systran. However, he did not think that page or column width posed any particular problems. There was an interface for the detection of abbreviations (with or without full-stops), decimal points, etc. In the Commission environment, end of sentence detection was very successful.

OCR equipment gave rise to some problems, as did badly presented source documents in general, since there was no mechanical way of improving input quality. PCs could be used as input devices for end-users, in connection with an automatic pre-processing facility. Mr Pahl insisted on the need to standardize solutions to input problems.

- 2 -

Mrs Facchin spoke of the problems posed by incorrect recognition of pagination. There were also some format problems: it was often difficult to persuade remote users to work on the 80 character format. Mr Klein explained that the pre-processor and post-processor normally took care of format problems.

Mr Lomax (Mendez) said he had also encountered formatting problems and often had to spend up to half an hour re-formatting a text after translation. Mr Pigott explained that the most universal approach to this problem was achieved by using the non-transparent communications mode from the Wang.

Mr Pahl explained two alternative approaches for solving the problem; first, communication of the print image (obtained via non-transparent TC), which was independent from the word processor on which the text was input, and second, the storing of the text-format codes in telecommunications. The latter method was generally more reliable providing the control characters of the input device could be recognized and supported by the Systran peripherals.

As regards actual type of equipment, it was the opinion of most Commission speakers that the Wang system was the best. It offered much better post-editing possibilities than other makes such as Olivetti or Philips.

It was generally felt that negative (white on black) screen display feature was tiring for the eyes, particularly when one was working in conjunction with a printout placed beside the WP screen. However, many participants had found that use of a split screen (displaying source and target languages) was also tiring and was often difficult to post-edit owing to interference in WP mode between the two languages in question. Other problems such as small print, rapid running of text in, for example, global replace on the CRT screen and changing of characters would have to be solved. It was generally agreed that an A4 screen offered the best display.

A discussion ensued on the various types of errors most frequently encountered by post editors. Mr Lomax, speaking about the French-English system, mentioned nominal groupings, the definite article, and rearrangement problems. The global replace feature was frequently used to correct not-found words or other vocabulary errors.

Various other problems such as prepositions, idioms, articles, homographs, incorrect use of the passive, verb agreement, past participle agreement and the sequence of tenses were enumerated by Messrs Bizzaro, Lavorel and Spoden as being the most frequent source of errors in the English-Italian and English-French systems.

Mr Pigott asked for information on the kinds of error which took the most time to correct in post-editing. Mr Spoden replied that it was very difficult to say but maintained that it took approximately the same amount of time for the translator to post-edit on screen as on paper; however, if corrections were made on paper, they had subsequently to be incorporated on screen by a secretary.

On-screen editing obviously had the advantage of alleviating the work of the already over-burdened typing-pools and preventing the work from a very heavily post-edited text resembling a 'plate of spaghetti'. Mr Foucart added that he did not consider that a translator was demeaning himself or undermining his profession by working on a word processor.

Mr Shinriki (Systran, Tokyo) explained the working methods of the Japanese. They had hitherto relied mainly on pen and paper post-editing because of incompatibility problems posed by word processors. New systems were currently under development, and they would start working with more sophisticated equipment in March 1986.

Mr Pigott wound up the discussion by reading out the recommendations of a post-editor, Mr Roy Green, on possible technical improvements to word processors which would make the post-editor's life easier; they included recommendations on the number of keystrokes required for certain manoeuvres, improvements to programming possibilities, and developments concerning characters and graphics.