

# COMMISSION OF THE EUROPEAN COMMUNITIES

SEC(82) 1886 final

Brussels, 26 November 1982

THIRD ACTION PLAN FOR THE IMPROVEMENT  
OF THE TRANSFER OF INFORMATION  
BETWEEN LANGUAGES

---

1. Explanatory memorandum

Broadly speaking, the third plan is a continuation of the projects undertaken under the first and the second plans. However, changes in circumstances have resulted from the appearance of new needs and new techniques, and a number of new projects will be undertaken to take account of these developments.

Additional objectives of the third plan are to make the most of the economic consequences of the projects undertaken, and to bring them into line with the Commission's general policies on innovation.

1.1. Results of the first and the second plans

The first two triennial plans for 'the improvement of the transfer of information between European languages', adopted by the Commission on 23 December 1976 and 10 June 1980, led to the following achievements:

- (a) the Systran system of pre-translation of natural language text was improved considerably. Its development was directed towards initial application in the Commission's own departments and to an application within the scope of the Euronet-Diane network. Already, an efficient electronic office infrastructure allows a number of translators access to the system under optimum conditions.
- (b) EURODICAUTOM, the terminological data bank of the Commission was enlarged and special attention was given to the less well represented languages, especially Greek; it was extended to include other fields of interest to the European institutions; the bank is now available via the Euronet network.
- (c) Two multilingual macrothesauri, covering the fields of agriculture and education, were completed and implemented in the Member States. The development of a macrothesaurus covering Community business was investigated. An infrastructure was set up for the maintenance and updating of the existing thesauri.
- (d) The holding of a large number of conferences and seminars devoted to the major problems of multilingualism drew public attention to the means by which the language barrier can be overcome. These objectives also formed the background for the launching of the review Multilingua, which was made possible with support from the Commission.
- (e) Within the scope of the second action plan, in a vast effort of cooperation ten European university institutes participated in drawing up the specifications for the Eurotra system of machine translation.

1.2. Starting point for the Third Plan

The following factors were taken into account in drawing up the work programme under 2 below:

- (a) The development of new technologies to serve computerized information retrieval systems in the context of telecommunications and business management, provides a new, fast and effective infrastructure for the multilingual tools created during the first action plans.
- (b) The accession of Greece has now become effective, and the forthcoming accession of other states to the European Community will further increase the workload of the translation departments with the incorporation of additional languages.
- (c) The economic crisis has created greater needs in Europe as regards multilingualism. Access to the markets of the Middle and Far East can only be achieved through knowledge of Japanese, Arabic and Chinese.
- (d) Through the efforts of the Commission in its action plans, many professions and organizations in the Member States and elsewhere have become aware of the existence of language barriers in the information market and of the ways to surmount them.
- (e) The Eurotra project, aimed at the creation of a machine translation system of advanced design for the next decade, should be adopted by the Council in 1982. The project will have a separate budget.

2. Work programme

As in the case of the first two plans, the projects will be preceded by studies and surveys and rounded off by evaluations, pilot experiments and operational applications.

The programme of work will be supplemented by an ongoing survey of current research and active and systematic support for those developing new systems in the Member States, so that such systems can be orientated and geared to the needs of the Community.

A large part of the work will be entrusted to firms and specialized agencies in the Member States. As in the case of the first action plan, the Community will be calling for the opinion of the Committee of Experts for the Transfer of Information between Languages (CETIL), which consists of specialists from the Member States and the Commission, chosen for their knowledge of the technical problems and needs on a European scale.

For the applications of multilingual tools to scientific and technical information, the Commission will, as in the past, consult the Committee for Information and Documentation on Science and Technicology (CIDST) and its "Multilingual Aspects" sub-group.

The action plan will cover the following activities:

- (a) Development of the output capacity and the quality of the machine translation software to meet the needs of the Community and its institutions.
- (b) Survey and collection of existing terminological resources and the creation of multilingual terminological tools, including thesauri for information management, terminological data banks and other computerized dictionaries, and dictionaries for machine translation.
- (c) Creation of an infrastructure for multilingual activities substantially based upon the latest developments in office automation and telematics.
- (d) Development of methodologies and standards governing multilingual communication. Examination of the implications, in particular as regards organization.
- (e) General investigations relating to multilingualism and study of new problems.
- (f) Promotion of practical applications of existing multilingual tools. Investigation of ways and means to make them available to the European industry.

## 2.1. Machine translation

- 2.1.1. The pilot project under way at the Commission, which is intended to familiarize linguists with the product of machine translation and with word processing equipment, will be continued. Each time a translator judges the quality of a raw translation as acceptable, he will act as post-editor and his productivity will thus be increased. At the end of the action plan the system should be capable of producing raw translation of any Community documents drawn up in the languages covered by the system and presenting no semantic or syntactic peculiarities.

Additional languages will be implemented only when evaluation has shown the utility of the new systems.

- 2.1.2. The Commission will continue its efforts to promote the use of machine translation in the context of Euronet-DIANE, in which the user will be able to request and obtain an immediate translation of the information retrieved into his own language.
- 2.1.3. The Commission will encourage the establishment of service bureaux providing raw or post-edited translations on request, initially for European institutions and government agencies of Member States. A subsequent extension of such services to industry and commerce in the Member States will be left to private initiative.
- 2.1.4. The Commission will ensure that it is kept informed on machine translation systems planned or operational elsewhere, and will be prepared to profit from such projects by way of cooperation agreements. It will continue to carry out studies and surveys with the aim of increasing the efficiency and cost-effectiveness of its own system.

## 2.2. Multilingual terminological tools

### 2.2.1. Scientific and technical information management

Having contributed to the development of the most advanced methods of thesaurus construction, and having set up a mechanism for creating, updating and exploiting multilingual thesauri, the Commission will develop and implement a thesaurus covering all Community languages for the exchange of information between the Community institutions and Member States.

It will continue to provide the computer infrastructure for the management of the thesauri created during the preceding action plans.

2.2.2. Aids to translation and multilingual documentation

For its translators the Commission has set up a highly reliable terminological data bank, the extension of which will be continued through the quest for collaboration with similar projects in Member States. Particular emphasis will be laid on increasing the coverage of the less widely-used languages. For the benefit of the users of Euronet-DIANE, it will encourage the creation of a bank resulting from the merging of a number of multilingual terminological collections covering a vast range of subjects.

2.2.3. Dictionaries for machine translation

The Commission will continue creating dictionaries for the fields which are of interest to the Community. Recourse will be had to the existing terminological data banks to ensure that this is done in the most economical manner.

These dictionaries will be tested and used with the Systran system. Their language coding will be adapted gradually to the requirements of the EUROTRA system.

2.2.4. Computerized dictionaries

The Commission will encourage the publishers of general and specialized dictionaries, who have computerized the maintenance and updating of their works, to pool their resources and exploit them to the benefit of other applications such as terminological data banks and machine translation.

2.3. Data-processing infrastructure

After an extensive market survey the Commission has acquired a word processing infrastructure, which will be of particular benefit to the departments which have to deal with repetitive or frequently altered texts.

The translation services require more specialized equipment which permits the input, indexing, retrieval and printing of documents and link-up with the computers holding the terminological data bank and the translation system, whilst remaining compatible with the less sophisticated equipment of the other services.

After further market research, the translation services will be supplied with suitable equipment, and the staff trained to use it in the best possible way.

#### 2.4 Methodology and standardization

The processing, printing and transmission of texts written in all the languages used by the European community will require a variety of characters larger than that available on current equipment.

The Commission will take part in work towards having character sets and a new transmission code permitting the input, transmission and printing of texts in all these languages adopted by standardization organizations, telecommunications administrations and equipment manufacturers.

#### 2.5. Studies and surveys

A wide range of studies will undoubtedly be required if the programme is to be carried out effectively, even if it is not yet possible to give precise specifications for such studies. Those which follow are given as examples, and may well be replaced during the programme by others which appear at that time to be more appropriate.

Studies will be undertaken on language comprehension and the mechanism of information transfer between individuals and between man and machine.

Study of the translation market may be supplemented by specific studies of the industrial and institutional market for machine translation and terminological data banks.

The consequences of the accession of new Member States to the Community for the data-processing aspects of multilingual communication will be analysed. It may be necessary to pay particular attention to incorporating the Greek alphabet in the character set of automated office equipment.

Economic and technical studies may be devoted to the possible extension of Community multilingual tools to non-Community languages, such as Arabic, Japanese and Chinese.

#### 2.6. Promotion plans

Such a wide-ranging action plan requires the participation of all those who are involved in multilingualism, whether at first hand or less directly, including teachers, publishers, documentalists, computer specialists, translators, interpreters and terminologists and those responsible for industrial marketing.

To be sure of their assistance, it is necessary first to identify them, contact them, inform them and interest them. That will be done through the organization of courses and seminars and a demonstration campaign. The publication of review articles, abstracts of published works and reports of projects undertaken during the period under review is ensured through the review Multilingua.

In addition to the technical projects, all the work carried out under this action plan will aim at improving multilingual communication within the Communities.



3. Budget

The budgetary expenditure requested for the action plan is estimated at 3 750 000 EUA, allocated as follows:

	1983	1984	1985
Machine translation	550 000	500 000	450 000
Terminological tools	250 000	250 000	250 000
Data-processing infrastructure	150 000	200 000	250 000
Other activities	250 000	300 000	350 000
	-----	-----	-----
TOTAL	1 200 000	1 250 000	1 300 000

These figures and the distribution thereof are approximate and may be altered in the light of the budgetary appropriations entered each year under heading 751 "Complementary activities to the Three Year Action Plans".

9

3rd Action Plan for the improvement of the transfer of information  
between European languages

FINANCIAL STATEMENT

1 - Budgetary heading concerned: 7510

2 - Wording of the budgetary heading:

Complementary activities to the three-year Action Plans

3 - Legal basis and references:

General authority of the Commission on the organization of its services.

References:

Commission Decision of 23 December 1976 adopting the first Action Plan for the improvement of the transfer of information between European languages (COM(76) 705 final).

Commission Decision of 10 June 1980 adopting the second Action Plan (COM(80) PV 314 final).

4 - Description, objective and justification of the action:

4.1. - Description:

The third three-year Action Plan has, broadly, a twofold thrust: on the one hand, the further development of what the Commission has already achieved relative to multilingual thesauri, terminological data banks and the machine translation system and the technical infrastructure for it; and on the other, activities not yet covered by the 1st or 2nd Action Plans and meeting a Community need, such as the implementation of new multilingual tools and a promotion and information campaign directed towards the circles concerned with multilingualism.

4.2. - Objectives:

To improve the work of the Community institutions by contributing to the progressive elimination of the linguistic barriers to information transfer.

4.3. - Justification:

The existence to date of seven official languages, with the prospect of additional ones consequent upon the accession of new Member States to the Community; the wish to preserve the cultural variety of the nations of Europe, which finds its expression in linguistic diversity; and the need to eliminate the obstacles to communication which result from this diversity, have led the European institutions to encourage the development of multilingual tools which may provide assistance with these problems.

5 - Financial impact of the action in EUA:

The total cost of the action is about 3,750,000 EUA, distributed over three years approximately as follows:

	(1983)	(1984)	(1985)
Machine translation	550,000	500,000	450,000
Terminological tools	250,000	250,000	250,000
Data-processing infrastructure	150,000	200,000	250,000
Other activities	250,000	300,000	350,000
<u>TOTAL</u>	<u>1.200.000(*)</u> =====	<u>1.250.000</u> =====	<u>1.300.000</u> =====

(\*) Partial or full completion of this programme will depend on the budgetary amounts available for 1983.