

# COMMISSION OF THE EUROPEAN COMMUNITIES

COM(81) 603 final

Brussels, 4 November 1981

Amended Proposal for a

## COUNCIL DECISION

on the adoption of a European Economic Community research  
and development programme for a machine translation  
system of advanced design

-----

(submitted to the Council by the Commission pursuant to  
the second paragraph of Article 149 of the EEC Treaty)

COM(81) 603 final

Com 603

**Communication from the Commission to the Council**

---

**Subject Creation of a European machine translation system of advanced design (EUROTRA).**

On 23 June 1980 the Commission submitted to the Council a proposal for a Council Decision on the adoption of a European Economic Community programme for a machine translation system of advanced design (DOC. COM(80) 314 final).

Being consulted for their opinion, the Scientific and Technical Committee (CST), the Economic and Social Committee and the European Parliament have successively recommended the EUROTRA project, but suggested a certain number of amendments, which concern the procedures and do not change the substance, nor the orientations.

Subsequently, the Commission has accepted these amendments and submits the text modified accordingly for approval by Council.

The new version of the text is attached in annex.

TEXT SUBMITTED BY THE COMMISSION TO THE COUNCIL ON 23 JUNE 1980 (1)

TEXT AMMENDED (\*) AFTER CONSULTATION OF THE EUROPEAN PARLIAMENT, THE ECONOMIC AND SOCIAL COMMITTEE AND THE SCIENTIFIC AND TECHNICAL RESEARCH COMMITTEE (CREST)

Proposal for a Council Decision on the adoption of a European Economic Community research and development programme for a machine translation system of advanced design

Proposal for a Council Decision on the adoption of a European Economic Community research and development programme for a machine translation system of advanced design

THE COUNCIL OF THE EUROPEAN COMMUNITIES,

THE COUNCIL OF THE EUROPEAN COMMUNITIES,

Having regard to the Treaty establishing the European Economic Community, and in particular Article 235 thereof,

Having regard to the Treaty establishing the European Economic Community, and in particular Article 235 thereof,

Having regard to the proposal from the Commission,

Having regard to the proposal from the Commission,

Having regard to the opinion of the European Parliament,

Having regard to the opinion of the European Parliament,

Having regard to the opinion of the Economic and Social Committee,

Having regard to the opinion of the Economic and Social Committee,

Whereas, under Article 2 of the Treaty, one of the Commission's tasks is to promote throughout the Community a harmonious development of economic activities, a continuous and balanced expansion and an increase in stability,

Whereas, under Article 2 of the Treaty, one of the Commission's tasks is to promote throughout the Community a harmonious development of economic activities, a continuous and balanced expansion and an increase in stability,

(\*) the modifications are underlined.

(1) OJ No C 234, 12,9,1980, p.2.

4

Whereas the multilingual nature of the European Community is of high cultural value, but is also in practice an obstacle to closer ties between the peoples of the Community, to communications and to the development of the internal and external trade of the Community,

Whereas the benefits of teleinformatics and of communication and information networks cannot be fully reaped at Community level unless the language barrier is overcome,

Whereas the development of computational linguistics is likely to contribute to the overcoming of this barrier,

Whereas considerable research has already been carried out on this subject in the Member States,

Whereas this research, by its very nature, can only bear fruit provided that it is supported by a Community action which would, among other things, have a catalytic effect on work already carried out,

Whereas such Community action can, in particular, consist in the creation of a European machine translation system of advanced design,

Whereas the multilingual nature of the European Community is of high cultural value, but is also in practice an obstacle to closer ties between the peoples of the Community, to communications and to the development of the internal and external trade of the Community,

Whereas the benefits of teleinformatics and of communication and information networks cannot be fully reaped at Community level unless the language barrier is overcome,

Whereas the development of computational linguistics is likely to contribute to the overcoming of this barrier,

Whereas considerable research has already been carried out on this subject in the Member States,

Whereas this research, by its very nature, can only bear fruit provided that it is supported by a Community action which would, among other things, have a catalytic effect on work already carried out,

Whereas such Community action can, in particular, consist in the creation of a European machine translation system of advanced design,

Whereas such a system would have numerous industrial applications and be of direct benefit, inter alia, to the various industrial sectors and, in particular, to exporting industries,

Whereas such a system would have numerous industrial applications and be of direct benefit, inter alia, to the various industrial sectors and, in particular, to exporting industries,

Whereas the preliminary work already completed has demonstrated the technical feasibility of such a system,

Whereas the preliminary work already completed has demonstrated the technical feasibility of such a system,

Whereas a Community research and development programme on machine translation is therefore likely to make an effective contribution to the achievement of the above objectives of the Treaty,

Whereas a Community research and development programme on machine translation is therefore likely to make an effective contribution to the achievement of the above objectives of the Treaty,

Whereas the Treaty has not provided the necessary powers to this end,

Whereas the Treaty has not provided the necessary powers to this end,

Whereas the Scientific and Technical Research Committee (CREST) has delivered an opinion on the Commission's proposal,

Whereas the Scientific and Technical Research Committee (CREST) has delivered an opinion on the Commission's proposal,

HAS DECIDED AS FOLLOWS

HAS DECIDED AS FOLLOWS

Article 1

Article 1

A research and development programme for the creation of a machine translation system of advanced design is hereby adopted for a period of five years commencing on .....

A research and development programme for the creation of a machine translation system of advanced design is hereby adopted for a period of five and a half years commencing on the day on which this Decision is published in the Official Journal of the European Communities.

The content of the programme is described in the Annex hereto.

### Article 2

The expenditure required for the execution of the programme is estimated at 7.7 million EUA (1979 value), personnel requirements are estimated at 16 persons.

The figures are given for guidance purposes only, the funds needed to execute the programme shall be determined every year by the budgetary procedure.

### Article 3

The Commission shall be responsible for the execution of the programme, in particular by means of research contracts. It shall be assisted by an Advisory Committee on Programme Management, the membership and terms of reference of which are set out in the Council resolution of 18 July 1977 on advisory committees on research programme management (1).

---

(1) OJ No C 192, 11.8.1977, p. 1

Under the terms of Annex 1,

the programme shall be divided into phases, each one concluding with a review which may involve a revision of the programme. The Advisory Committee, referred to in Article 3, CREST and the European Parliament shall be consulted on these revisions

### Article 2

The expenditure required for the execution of the programme is estimated at 9.1 MECU at 1981 values, including expenditure in respect of a staff of 8 temporary agents.

The figures are given for guidance purposes only, the funds needed to execute the programme shall be determined every year by the budgetary procedure.

### Article 3

The Commission shall be responsible for the execution of the programme, in particular by means of research contracts. It shall be assisted by an Advisory Committee on Programme Management (ACPM), the membership and terms of reference of which, in accordance with the Council Resolution of 18 July 1977, are set out in Annex 2.

---

(1) OJ No C 192, 11.8.1977, p.1.

The Commission shall inform the CREST at regular intervals on the progress of the work. Furthermore, the Commission shall submit an annual report to the Council and to the European Parliament concerning the execution of the programme.

#### Article 4

The dissemination of information resulting from the execution of the programme shall be effected in accordance with Council Regulation (EEC) No 2380/74 of 17 September 1974 adopting provisions for the dissemination of knowledge relating to research programmes for the European Economic Community (2).

#### Article 5

In accordance with Article 228 of the Treaty, the Community may conclude agreements with third countries on their participation in the research programme which is the subject of the present Decision.

The Commission is hereby authorized to negotiate such agreements.

---

(2) OJ No L 255, 20.9.1974, p. 1

The Commission shall inform the CREST and the CIDST (2) at regular intervals on the progress of the work. Furthermore, the Commission shall submit an annual report to the Council and to the European Parliament concerning the execution of the programme.

#### Article 4

The dissemination of information resulting from the execution of the programme shall be effected in accordance with Council Regulation (EEC) No 2380/74 of 17 September 1974 adopting provisions for the dissemination of knowledge relating to research programmes for the European Economic Community (3).

#### Article 5

In accordance with Article 228 of the Treaty, the Community may conclude agreements with third countries in particular with those taking part in European co-operation in the field of scientific and technical research (COST), on their participation in the research programme which is the subject of the present Decision.

The Commission is hereby authorized to negotiate such agreements after the ACPM and the CREST have been consulted.

---

(2) Committee for Information and Documentation in Science and Technology

(3) OJ No L 255, 20.9.1974, p.1.



8

Article 6

Insofar as the Community shall promote the programme by means of research contracts, provision shall be made for repayment of part of the contribution made by the Community if the results of the programme are used industrially or commercially and/or lead to industrial property rights and the subsequent award of licences.

The Commission shall negotiate and conclude the necessary contracts. For this purpose it shall draw up, after consultation of the ACPM, a standard contract defining the rights and obligations of each party, including where appropriate conditions and procedures for possible royalty payments and for the repayment of a part of the contributions made by the Community.

## Annex

### 1. Objectives

The objective of the project is the creation of a computer-aided translation system capable of dealing with all official languages of the Community.

### 2. Programme of work

The programme of work includes the following elements:

- The development of an analysis module for each language, which will carry out morphological, syntactic and logico-semantic analysis of source text to the levels specified by the definition of the commonly agreed interface structure.

- The development of a generation module for each language, which will be able to accept an interface structure of the type commonly agreed and generate from it a surface representation in that language.

## Annex 1

### 1. Objectives

The objective of the programme is the creation of a computer-aided translation system capable of dealing with all official languages of the Community.

On completion of the programme an operational system prototype should be available in a limited field and for limited categories of text, which would provide the basis for development on an industrial scale in the period following the current programme.

### 2. Programme of work

The programme is divided into three phases:

(a) Preparatory phase (18 months, 1.0 MECU)

In this phase the following work would be carried out in parallel:

1. First

- The development of a transfer module for each language couple, which will take as input an interface structure of the type commonly agreed, derived from an analyses of the source text, and produce from it an interface structure of the same type, from which the target text may be generated.

- The compilation of monolingual dictionaries in two steps:

a) an initial test dictionary for each language, to be used both in analysis and in generation, comprising approximately 2 500 entries in a specific subject field,

b) an expanded monolingual dictionary, whose final size will be approximately 20 000 entries.

- The compilation of a multilingual lexical data base for use in the transfer phase, in two stages as for the monolingual dictionaries.

- setting up of the ACPM,  
definition of the project and its organization and of the responsibilities of the participating countries and centres,

- definition of the methodology of the work and preparation of a detailed programme of linguistic work to be carried out by the participating centres, and of the sectors and categories of texts covered by the research,

- allocation of intellectual and industrial property rights,

- identification of specific problems related to the inclusion of Greek, and assessment of the related costs.

2. Second

- definition of the arrangements for disseminating the results of the work,

- preparation of more detailed specifications of the linguistic models and strategies for the various components of the process (analysis, transfer, generation),

- The development of a basic software package, including support software for dictionary handling and updating as well as software capable of dealing with the analysis, transfer and generation modules described above. This will include interpreter compiler programmes for the specially defined high-level definition language in which grammars and dictionaries are to be written.

- The development of the software for the user-system interface, by which the different modules may be combined into integrated systems corresponding to the various user options defined.

- The Commission will ensure that the objectives relating to portability and compliance with the international standards are accurately reflected in the specifications.

- preparation of detailed and binding specifications for the EUROTRA basic software and the data processing programmes capable of carrying out the various processes: analysis, transfer, generation, monitoring functions and text management.

The Commission will ensure that the objectives of portability and compliance with international standards are correctly reflected in the specifications.

- preparation of the contracts of association.

At the end of this phase the opinion of the ACPM must be obtained on the above specifications in order that the linguistic work can progress quickly and so that the widest possible call for tenders for construction of the software can be put out as soon as possible (see (b) (2) below).

12

b) Phase of basic and applied linguistic research (2 years, 4.5 MECU)

On successful completion of the first phase, and after a favourable opinion has been received from CREST, the second phase will be divided into two parts:

1. Basic linguistic research (2 MECU)

This part will consist of the following work (1)

- the development of initial linguistic models for the analysis and generation of each of the official Community languages covered by this Decision and for transfer between these languages. This work will be based on a corpus and a vocabulary in a limited field, estimated at around 2 500 entries,

- preparation of the lexical data base, for the abovementioned vocabulary, which will serve both for the analysis and for the generation of each of the languages and for the transfer between these languages,

- a study of the linguistic strategies best suited to machine-execution of the various processes.

2. Construction of the basic software for EUROTRA (2.5 MECU)

This part comprises the following work:

---

(1) Some parts of this work may continue during the subsequent phase.

- scrutiny by the Commission of the replies to the call for tenders and selection, after consultation of the ACPM, of a body to construct the EUROTRA basic software, within as short a time as possible,

- development of the basic software by the body selected, including:

= the high level language for describing the linguistic data and strategies,

= the high-level language for interaction between the user and the system, which will make it possible to combine the various modules into integrated systems corresponding to the different utilization options,

= the set of utility software systems for compiling the high-level languages, for tests and for management of the data bases.

This initial version of the software is intended to enable the development and machine-testing of the linguistic models defined by the participating centres when they are sufficiently advanced. Its development is consequently a prerequisite for validating the linguistic work under this programme.

The industrial development of the EUROTRA system, including adaptation of the software to the performance and reliability requirements for producing translations under commercial conditions, will not be put in hand until the present programme has been completed.

141

c) Phase of stabilization of the linguistic models and evaluation of results (2 years, 3.6 MECU)

After opinions have been received from the ACPM and the CREST at the end of the second phase, i.e. when it is possible to carry out systematic testing of the initial linguistic models, comprising complete language pairs and consisting of analysis, transfer and generation, the objective of the work will be concentrated on the following aspects:

- adapting the linguistic models, in order to produce linguistic modules which are as reliable as possible. The modules will then be fit for pre-operational use,

- progressively extending the basis of the text corpus, the linguistic models and the vocabulary for a specific field, and on texts of increasing complexity,

- revising and progressively extending the lexical bases to cover the chosen field as exhaustively as possible (about 20 000 entries in all the languages),

- evaluating the technical and economic performance of the system,

- preparing a proposal for the development of an operational system on an industrial scale and proceeding to the stage of commercial exploitation.

Annex 2

TERMS OF REFERENCE

for an Advisory Committee on Programme Management for a European Advanced Machine Translation System (EUROTRA)

1. Without prejudice to the Commission's responsibility in the implementation of the programme, it shall be the task of the Advisory Committee on Programme Management for an Advanced Machine Translation System (EUROTRA), hereinafter referred to as "the Committee", to contribute, in its advisory capacity as defined by the Council Resolution of 18 July 1977 (1), to the best possible implementation of the programme for which it is responsible, and in particular to:

- the general organization of the programme, specifying in particular the main deadlines and the mechanisms envisaged for following the progress of the work,

- the preparation of contracts of association setting out the commitments of the institutions participating in the project at national level,

- the definition and solution of problems concerning intellectual and industrial property rights, connected with the different system components, and of the procedures for disseminating the results of the work,

(1) OJ No C 192, E1.8.1977, p.1.



16

- the drawing up of binding specifications for the software to be jointly developed, to serve as a basis for the widest possible invitation to tender,

- the clarification of user requirements, in particular in the field of information and documentation,

- the drawing up of detailed financing plans relating to the various fields of work and components of the system.

2. In carrying out these tasks the Committee may if necessary call on the services of independent experts.

3. The Committee will ensure that there is the necessary coherence between this programme, the Community's plans of action for the improvement of the transfer of information between European languages, and the plans of action in the field of scientific and technical information and documentation. For this purpose it will maintain close contact with the CIDST and the other committees set up under the auspices of the Community Institutions with responsibility in the field of multilingualism.

17

4. The Committee will ensure the evaluation of the results of the programme and, in particular, the dissemination of the knowledge acquired to Community users in accordance with the third indent of (1).

5. The Committee is also requested to give opinions on:

- the annual preparation of budgets and the allocation of appropriations accorded,

- any proposals for reviewing the programme and proposals for future research programmes within its field of competence,

- where appropriate, the conduct of negotiations with institutes in third countries with a view to their participation in the research programme,

- the multilingual aspects of the plans of action in the field of IDST.

6. The Committee shall submit to the Commission and to the Council a detailed report at the end of each phase of implementation of the programme. It shall also submit a final report which will be forwarded to the European Parliament.