

COMMISSION OF THE EUROPEAN COMMUNITIES

COM(89) 603 final - SYN 228

Brussels, 5 December 1989

Proposal for a
COUNCIL DECISION

adopting a specific programme concerning the preparation of
the development of an operational EUROTRA system

(presented by the Commission)

Concl.

Communication from the Commission to Council

Subject : Proposal for a Council Decision adopting a specific programme concerning the preparation of the development of an operational EUROTRA system.

01/02/80

I. BACKGROUND

1. The Framework Programme 1987-1991⁽¹⁾ foresees under action line 8.4. with regard to linguistic problems the following actions :

- completion by 1990 of a first multilingual prototype machine translation system (for nine official Community languages) (EUROTRA)

For this point the Commission has submitted to Council a proposal on 1 June 1988⁽²⁾ and amended it on 7 February 1989⁽³⁾ incorporating a number of recommendations of the European Parliament, and Council has adopted its Decision on 20 June 1989⁽⁴⁾

- support to the industrial development of a machine translation system, development of methods and tools for the reusability of lexical resources in computerized applications and creation of standards for lexical and terminological data.

These points are the subject of the present proposal.

2. Action line 8.4. of the Framework Programme foresees a total expenditure of 55 million ECU. This amount is subdivided among the various specific programmes as follows :

-	VALUE ⁽⁵⁾	38 million ECU
-	Phase 3 of EUROTRA	7 million ECU
-	present specific programme proposal	10 million ECU

TOTAL	55 million ECU
-------	----------------

The present proposal takes into account these factors and the duration of the current EUROTRA programme (ending on 30 June 1990).

(1) OJ n° L 302 of 24.10.1987, p. 1
(2) COM (88) 270 final
(3) COM (89) 20 final/2 - SYN 137
(4) OJ n° L 200 of 13.7.1989, p. 15
(5) OJ n° L 200 of 13.7.1989, p. 23

II. STARTING CONDITIONS OF THE PROGRAMME

3. The proposal builds on the achievements of the current EUROTRA programme.

The objective of the programme is the creation of a machine translation system of advanced design (EUROTRA) 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 serve as a basis for development of a practical system in the period following the current programme.

The practical results likely to be obtained by the end of the third phase include among others :

- for each of the nine languages analysis and synthesis modules with a broad coverage of the linguistic phenomena which occur in the text types chosen (handbooks, manuals, scientific articles in the field of telecommunications);
- for each of the languages the general vocabulary which is used in these texts (approx. 8.000 entries);
- the technical terminology of a number of subfields of telecommunications (approx. 12.000 entries);
- transfer modules with this coverage for a representative number of language pairs (approx. 50);
- a number of linguistic research results achieved during the final year of the programme which should be incorporated in the future implementations.

It is planned to carry out an intermediate evaluation of the progress of the EUROTRA programme during the second half of 1989. As far as possible the same experts as those who carried out the programme assessment in 1987 will be invited to conduct this evaluation. The evaluation report should become available before the end of 1989.

4. In the Council Decisions of 25 July 1988⁽⁶⁾ on the transition of EUROTRA to the third phase and of 20 June 1989⁽⁷⁾ on the completion of the programme, these initial objectives have been maintained. In addition, following the recommendations of the panel of independent experts who assessed the programme in 1987 and of the Energy, Research and Technology Committee of the European Parliament a number of additional actions have been included in the programme of work of the third phase of EUROTRA.
5. These additional actions aim in particular at the preparation for the implementation of a practical EUROTRA system, the involvement of enterprises, setting up cooperation schemes between research institutions and enterprises and training of researchers and engineers.
6. In the course of 1989 the Commission has launched studies covering the following points :
 - outline definition of a practical EUROTRA system including qualitative and quantitative performance, user and system maintenance environments, sub-, by- and spin-off products and systems, development strategies, actors, time scales, costs etc. Among other things this study should provide enterprises with elements which will allow them to make use of the research results of the EUROTRA programme for product development;
 - system analysis of the general architecture of EUROTRA. This study aims at the definition of a basis for further technological development both in view of machine translation and other applications;
 - design studies to define the system development, testing and research environment. They should lead to a number of recommendations and specifications covering the various system components needed for the development of large scale systems, especially in view of the economic feasibility of such development;
 - a project definition study for the re-usability of lexical resources in computerized applications.

In addition the Commission supports, and participates actively in, an international project, called "Text Encoding Initiative" which was launched by professional associations and aims at the definition of a proposal for international standards concerning textual and lexical data.

(6) OJ N°. L 222 of 12.08.1988, p.1

(7) OJ N°. L 200 of 13.07.1989, p. 15

III. OBJECTIVES AND CONTENT OF THE PROGRAMME

7. Considering the time-scales and the levels of financing foreseen for the proposed programme, the objectives are necessarily limited.

The programme is foreseen to last for a period of 2 years from 1 July 1990 to 30 June 1992 with an average Community funding of 5 million ECU per annum.

8. Following the actions foreseen in the Framework Programme the objectives of this proposal are articulated under three main headings :

- (1) Preparation for the development of an operational EUROTRA system.

The main objective of the actions under this heading is to create, starting from the EUROTRA prototype and building on the results of a number of ESPRIT projects, appropriate conditions for its further development. This includes further research aiming at the improvement of the linguistic performance of the system. The main action lines include :

- (a) Implementation of a stable and efficient system development and testing environment; it should include :

- a streamlined rule formalism with special emphasis on expressive power, reversibility, declarativity, modularity and control mechanisms;
- an efficient rule interpreter capable of handling large dictionaries, grammars and texts at an acceptable speed;
- a data base system for the storage and management of large dictionaries, grammars and text corpora;
- a user environment for the creation and maintenance of large dictionaries and grammars;
- a testing environment with special emphasis on interactive testing .

- (b) Language specific linguistic research and development work aiming at the extension of the prototype system and at the improvement of its linguistic performance. This action line should include :

- **Review of the existing linguistic implementations**

The existing linguistic implementations will have to be adapted to the revised framework and environment, so that they can be used as a stable basis for further extensions foreseen within the present programme and beyond.

- **Extension of the linguistic coverage**

The grammatical coverage is to be gradually extended so as to deal with phenomena occurring in new text types. Special attention should be given to the optimal use of the modularity of the system architecture to capture text type and subject field specific phenomena.

- **Experimentation with relayed transfer**

The basic principle consists in using one interface structure as a pivot through which transfer from and into all other languages transits. The economic stake for a practical multilingual system (awaiting an interlingual solution) is the reduction of the geometric progression of individual language pairs (72 for nine languages) to exactly one transfer from and into each language. The success of this experimentation would greatly enhance the economic feasibility of a truly multilingual system with equal coverage of all directions of translation.

- **Incorporation and testing of research results**

EUROTRA is designed as an evolutionary system able to incorporate new research results. When such results become available, they must be integrated in the relevant modules and extensively tested with regard to their effectiveness for the solution of the problems which they are intended to solve, mostly in three areas :

- the resolution of ambiguities;
- the reduction of complex transfer;
- the extension of the interlinguality of the interface structure.

Ultimately, they aim at the improvement of the quality of translation.

(o) Linguistic research (basic and applied) of general interest

To ensure the evolution of EUROTRA, a continuing effort of linguistic research is necessary.

An initial set of priority areas is expected to be defined as an outcome of the system review undertaken during the third phase and the evaluation at the end of the EUROTRA programme.

Strategically the systematic and exhaustive use of general linguistic knowledge is to be given priority. It is, however, well-known that certain classes of problems cannot be solved on the basis of linguistic knowledge alone, and that "extra-linguistic" knowledge is needed for their solution (common sense world knowledge, domain-specific knowledge, discourse-specific knowledge).

Since considerable research efforts are being made outside the present programme in relation to knowledge bases, expert systems, inferencing, advanced information systems etc., the main thrust will be in the use of the research results obtained in such projects and in experimentation in cooperation with such projects.

(d) Research in system architectures

EUROTRA is designed as an intelligent knowledge based system (IKBS) which at present incorporates linguistic specialist knowledge and eventually will incorporate, or interface with, extra-linguistic knowledge. Considerable effort is being put elsewhere in knowledge representation/engineering, expert systems and advanced hardware architectures (parallel, associative, connectionist machines). In order to give EUROTRA and language technology in general a sufficient potential for innovation, a permanent effort is necessary to explore the possibilities offered by the new developments and eventually to experiment their use for natural languages.

- (ii) Development of methods and tools for the reusability of lexical resources in computarized applications.

The objectives of this action line are twofold:

(e) Use of existing lexical and terminological resources for building EUROTRA dictionaries

The availability of large-scale high-quality dictionaries and terminological resources is crucial for a practical MT system and any other application involving natural language. This action line is to provide methods and tools for this purpose. The project definition study under way will provide the necessary details.

(f) Re-usability of lexical and grammatical resources

The purpose of this action line is to demonstrate that modules developed for EUROTRA can actually be used in a variety of applications with some natural language component. It involves pilot and demonstrator projects in cooperation with those responsible for relevant application projects.

(iii) Creation of standards for lexical and terminological data:

In every field of technology the availability of international standards is a necessity. For linguistic technologies the Commission considers the creation of such standards as a long-term objective. Within the present programme the only realistic actions to be envisaged must be of a preparatory nature including prenormative research and stimulatory actions, such as :

(g) Active participation in initiatives aiming at the creation of such standards (dictionaries, terminologies, corpora).

(h) Stimulation of activities aiming at providing adequate terminological coverage of the official languages in key areas of science and technology.

9. The results expected from this programme are :

- a set of stable, efficient and user-friendly software tools and a tested methodology for large-scale system development;
- a set of tools and methods for building large-scale dictionaries by the extensive use of existing lexical resources;
- a set of tools and methods for the incorporation of specialized terminological collections;
- for all nine Community languages well-tested implementations of analysis and synthesis modules with broad grammatical coverage and a representative lexical coverage (no large-scale lexicographic work is foreseen in this programme, pending the outcome of the work on the reusability of lexical and terminological resources);
- should the experimentation with relayed transfer be successful, the translation from each language into all other languages, otherwise the translation between a representative number of language pairs;

- experience of cooperation models aiming at the development of a practical multilingual translation system, and of a number of intermediate sub-, by- and spin-off products;

IV. ORGANIZATION AND ACTORS

10. EUROTRA involves the active participation of all twelve Member states. This participation is implemented through contracts of association with the Member States which determine also the ratio of Community and national contributions to the programme.
11. In general the responsibility of each national research team is the analysis and synthesis of its own national language and the transfer from the other languages into its own language. The work on the Dutch language is done jointly by the Netherlands and Belgium and on the French language by France and Belgium. Ireland is responsible for the coordination of the terminological work, while Luxembourg provides general infrastructural support for documentation and testing and distribution of the software and Belgium is also responsible for electronic lexicographic work of general interest.
12. All national teams participate actively in basic and applied linguistic research of general interest. This work is coordinated by the project team in the Commission with the support of a number of researchers from the national teams whose work is financed 100 % from Community funds.
13. The design of the general system framework and of the rule formalism and the initial prototype implementation of the software were carried out under a similar schema as the general linguistic research. In the course of the third phase, software development and maintenance have progressively concentrated in the Commission project team with the assistance of experts from the European software industry. This work is also fully financed from Community funds.

14. In the course of the third phase, the virtually closed schema of the contracts of association has been opened to additional research institutes and enterprises, who have been entrusted with a number of system analysis, feasibility and design studies.
15. In the light of the experience gained with the organizational structure of EUROTRA, the following schema is proposed for this programme :
 - (a) The design and specifications of the system development and testing environment are expected to be ready for the start of this programme. It is estimated that some 20 % of the overall budget is necessary for its implementation, which should be fully financed from Community funds.
 - (b) The language-specific R&D work described under point 8 letter b should continue and be co-financed by the Community and the Member States based on the same arrangements as the current contracts of association. It is proposed that about 40 % of the Community funds should be destined for this purpose.
 - (c) The actions of a more exploratory nature, especially those described under point 8, letters c, d, e and f should be carried out as cooperative ventures between industry, research institutes and existing EUROTRA teams and be co-financed on a 50-50 basis by the Community and the participants. The projects will be selected through calls for proposals/tenders. 30 % of the overall budget should be used for this purpose.
 - (d) The Commission is aware of the scarcity of expertise and talent in the field of machine translation and computational linguistics. It is therefore proposed to reserve 10 % of the Community budget to award grants to graduate students to enable them to participate in the work of the EUROTRA teams or of the other participants in the programme.
 - (e) The Commission will maintain the programme management and coordination function. It is estimated that 5 additional temporary staff are required for the implementation of this specific programme.

(f) The Commission proposes that in the execution of this Programme it should be assisted by an Advisory Committee, which should among others, take over the role of the present CGC-12.

V. FINANCIAL ESTIMATIONS

16. The overall budget of 10 million ECU should be subdivided as follows (in 1,000 ECU) :

	EC Contributions	National Contributions	TOTAL
. SYSTEM DEVELOPMENT AND TESTING ENVIRONMENT	2.000		2.000
. NATIONAL RESEARCH TEAMS			8.000
- Community contribution	4.000		
- National contribution		4.000	
. COOPERATIVE PROJECTS			6.000
- EC contribution	3.000		
- participants		3.000	
. TRAINING	1.000		1.000
. TOTAL	10.000	7.000	17.000

VI. DECISION REQUESTED FROM COUNCIL

17. The Commission requests the Council to adopt the proposal for a decision concerning a specific programme for development of an operational EUROTRA system which is annexed to this communication.

VII. ANNEXES

- I. Proposal for a Council Decision adopting a specific programme for the preparation of the development of an operational EUROTRA system .
- II. Opinion of CGC-12 of 26 June 1989.
- III. Financial Statement.

Proposal for a
COUNCIL DECISION

adopting a specific programme concerning the preparation
of the development of an operational EUROTRA system

THE COUNCIL OF THE EUROPEAN COMMUNITIES,

Having regard to the Treaty establishing the European Economic Community, and in particular Article 130q, (2) thereof,

Having regard to the proposal from the Commission⁽¹⁾,

In cooperation with the European Parliament⁽²⁾,

Having regard to the opinion of the Economic and Social Committee⁽³⁾,

Whereas Article 130k of the Treaty provides that the Framework Programme is to be implemented through specific programmes developed within each activity;

Whereas by Decision 87/516/Euratom, EEC⁽⁴⁾ the Council adopted a Framework programme of Community research and technological development (1987 to 1991) providing inter alia for activities on linguistic problems (action line 8.4), including the preparation of the development of an operational machine translation system;

Whereas by Decision 82/752/EEC⁽⁵⁾, the Council adopted a Community research and development programme for a machine translation system of advanced design (EUROTRA);

Whereas Decision 86/591/EEC⁽⁶⁾ amended Decision 82/752/EEC on account of the accession of Spain and Portugal;

Whereas by Decision 88/445/EEC⁽⁷⁾, the Council approved the transition of the EUROTRA programme to its third phase;

Whereas by Decision 89/410/EEC⁽⁸⁾, the Council adopted a specific programme for the completion of EUROTRA;

Whereas Decision 82/752/EEC provided that on completion of the EUROTRA programme an operational system prototype should be available in a limited field and for limited categories of text;

(1) OJ No

(2) OJ No

(3) OJ No

(4) OJ No L 302, 24.10.1987, p. 1.

(5) OJ No L 317, 13.11.1982, p. 19.

(6) OJ No L 341, 04.12.1986, p. 39.

(7) OJ No L 222, 12.08.1988, p. 1.

(8) OJ No L 200 13.07.1989, p. 15.

Whereas Decision 82/752/EEC included in the programme of work for the third phase of the EUROTRA programme the preparation of a proposal for the further development of the system;

Whereas the objectives of the EUROTRA programme are expected to be achieved by 30 June 1990;

Whereas the results of the programme are expected to constitute a solid basis for the preparation of the development of an operational multilingual machine translation system for the nine official Community languages;

Whereas the programme is expected to make an important contribution to the economic, social and cultural cohesion of the Community, and to reconcile the benefits of its linguistic plurality with the attainment of the internal market;

Whereas the programme constitutes the first step towards the development of a practical machine translation system;

Whereas additional efforts will have to be made by the Community, the Member States and European industry in order to reach this objective;

Whereas the Scientific and Technical Research Committee (CREST) has been consulted;

Whereas the Management and Coordination Advisory Committee "Linguistic Problems" (CGC-12) has delivered a favorable opinion;

Whereas an estimate of the amount of the financial contribution of the Community deemed necessary for the execution of this programme should be made;

Whereas this amount is included in the financial forecast annexed to the interinstitutional agreement of 29 June 1988⁽⁹⁾;

Whereas the appropriations effectively available will be determined by the budgetary procedure according to the provisions of the aforementioned interinstitutional agreement,

HAS ADOPTED THIS DECISION:

Article 1

A specific research and technological development programme concerning the preparation of the development of an operational EUROTRA system as defined in the Annex is hereby adopted for a period of two years from 1 July 1990.

(9) OJ No L 185, 15.7.1988, p. 33.

Article 2

1. The amount of Community funds deemed necessary for the execution of the programme adopted by this Decision is ECU 10 million, including expenditure on 5 temporary staff.
2. The appropriations to be made available each year shall be determined in the budget.

Article 3

Detailed rules for the implementation of the programme and the rate of the Community's financial participation are set out in the Annex.

Article 4

The Commission shall conduct an evaluation of the results achieved at the beginning of 1992 and shall transmit a report on this subject to the Council and the European Parliament. This report shall be established having regard to the objectives set out in Annex and in accordance with Article 2, (2) of Decision 87/516/Euratom, EEC.

Article 5

1. The Commission shall be responsible for the execution of the programme.
2. The contracts concluded by the Commission shall regulate the rights and obligations of each party including the methods of disseminating, protecting and exploiting the research results.

Article 6

1. The Commission shall be assisted by a committee of an advisory nature composed of the representatives of the Member States and chaired by the representative of the Commission.
2. The representative of the Commission shall submit to the committee a draft of the measures to be taken. The committee shall deliver its opinion on the draft, within a time limit which the chairman may lay down according to the urgency of the matter, if necessary by taking a vote.
3. The opinion shall be recorded in the minutes; in addition, each Member State shall have the right to ask to have its position recorded in the minutes.
4. The Commission shall take the utmost account of the opinion delivered by the committee. It shall inform the committee of the manner in which its opinion has been taken into account.

Article 7

In accordance with Article 130n of the Treaty, the Commission is hereby authorised to negotiate agreements with international organisations, third countries participating in European cooperation in the field of Scientific and Technological Research (COST) and European countries having concluded framework agreements for scientific and technical cooperation with the Community, with a view to associating them wholly or partly with the programme.

Where framework agreements for scientific and technical cooperation between European countries and the European Communities have been concluded, organisations and enterprises established in those countries may participate in a project undertaken within this programme.

Article 8

This Decision is addressed to the Member States.

Done at Brussels,

For the Council

The President

ANNEX

1. Objectives

This programme constitutes the first step towards the development of an operational machine translation system of advanced design, capable of dealing with all official Community languages. The specific objectives of the programme are :

(a) Creation of the conditions for the transition to an operational system

- implementation of a development, testing and research environment capable of supporting large-scale systems;
- extension of the linguistic coverage and large-scale testing of the analysis and synthesis modules for all languages covered by EUROTRA;
- the definition of common methods for large-scale development for machine translation and other applications involving natural language;
- experimentation and evaluation of relayed transfer, using an interface structure as pivot;
- research, prototype implementation and evaluation of new linguistic models aimed at the improvement of the interlinguality of the interface structure and control of overgeneration;
- research, prototypical implementation and evaluation of methods for the use of subject-field and text-type specific knowledge for translation and other applications.

(b) Advancement of work on lexicography and terminology

- definition of common methods and tools for the integration of existing lexical and terminological collections;
- participation in the definition of international standards for textual, lexical and terminological data.

(c) Training and cooperative projects

- training of researchers and engineers through a grant scheme;
- setting-up, experimentation and evaluation of cooperation schemes between research institutes and industry.

2. Priority action lines and scientific and technical content

2.1. System development, testing and research environment

On the basis of the critical review of the prototype implementations and the specifications produced by the end of June 1990 the implementation of a system development, testing and research environment will be commissioned to the European software industry on a turn-key basis.

This environment should have the following characteristics :

- a powerful and user-friendly formalism for describing linguistic facts (to encode dictionaries and grammars);
- a special-purpose data management system for the creation and maintenance of large-scale dictionaries and grammars with special user services for the addition, inspection and modification of the linguistic data;
- an efficient rule interpreter capable of dealing with large dictionaries, grammars and texts;
- a set of testing tools with special attention to interactive run-time testing and correction of dictionaries and grammars.

The system development environment will pay special attention to modularity to ensure the re-usability of the linguistic resources created through the possibility of combining the various modules in different ways to fulfil special tasks and of interfacing them with external applications.

2.2. Language-specific Research and Development work

2.2.1. Review and extension of the existing implementations

For each of the nine languages covered, the analysis and synthesis modules produced by the EUROTRA programme will be thoroughly reviewed to ensure their generality and adjusted to the features of the revised formalism.

On the basis of these revised implementations the grammatical coverage will be gradually extended to include additional text and discourse types. No large-scale lexical and terminological development work is planned for this phase pending the outcome of the research on the re-usability of lexical and terminological resources (see 2.3 below).

2.2.2. Relayed transfer

The research and experimentation are to determine the feasibility of this approach and the optimal strategy for its implementation.

2.3. Linguistic research of general interest

This action line is intended to gradually improve the linguistic performance of the system and the quality of translation. It will be organized along three main directions :

- general linguistic research to increase the interlinguality of the interface structure and to reduce overgeneration;
- use of subject-field specific knowledge (terminologies, classification schemes, paradigmatic relations, domain models, knowledge bases etc.);
- use of text and discourse type-specific constraints to reduce overgeneration.

It can be predicted that some progress will be made in the course of this programme, but additional efforts must be foreseen for the future.

2.4. Research into advanced system architectures

To create a potential for innovation and keep pace with the fast advancement of hardware and software technologies a continuous research into new formalisms, software and hardware architectures is to be foreseen, which will lead in selected cases to experimentation and prototype implementations (e.g. parallel system architectures).

2.5. Re-usability of lexical and terminological resources

The details of this action line will be defined through project definition studies to be carried out in 1989 and 1990.

It is expected to have two main components :

- development of methods and tools for the conversion of the formalized parts of existing dictionaries which cover mostly orthographic, phonological, morphological and syntactic information;
- research into the utilization of non-formalized portions of dictionaries which concern mostly subject classification, discourse types, definitions and examples or citations. This is an advanced research topic whose outcome cannot be predicted now.

2.6. Standards for textual, lexical and terminological data

This activity is closely related to the re-usability of linguistic resources in the future. The Commission will support and stimulate international activities in this field in close cooperation with professional associations and national and international standards organizations.

2.7. Education and training

A number of research grants will be awarded to post-graduate students to participate in the research and development work in the projects outlined hereabove.

2.8. Evaluation

At the end of the programme the results will be evaluated by independent experts.

3. Modalities of Execution

The various action lines pursue different objectives which demand different forms of organizational, contractual and financial schemes.

3.1. Services Contracts

The implementation of the system development, testing and research environment (action line 2.1) which will provide all participating parties with a common set of tools, will be entrusted to industrial contractors on the basis of calls for tenders. It will be financed fully from the Community budget.

3.2. National research teams

The work concerning the different languages (action line 2.2) will be carried out by national research teams in the member States, and co-financed by the Community and the member States.

3.3. Shared cost projects

The linguistic research of general interest (action line 2.3), research and development into advanced system architectures (action line 2.4) and re-usability of lexical and terminological resources (action line 2.5) will be carried out as cooperative ventures between industries, research centres and EUROTRA teams. The specific projects and contractors will be selected on a call for tenders/proposals basis. The Community contribution to these projects will be in general 50%.

3.4. Grants

Grants will be awarded to qualified post-graduate students.

3.5. Subsidies

The Commission will award subsidies to professional associations and standards organisations for action line 2.6.

4. Indicative break-down of the expenses

The indicative break-down of the amount of 10 million ECU deemed necessary for the execution of the programme is as follows (in 1000 ECU) :

(a) System development environment	2,000
(b) EC contribution to the national research teams	4,000
(c) Shared-cost research projects	3,000
(d) Training, subsidies, evaluation	1,000
	<hr/>
Total	10,000

OPINION OF THE MANAGEMENT AND COORDINATION ADVISORY COMMITTEE

LINGUISTIC PROBLEMS (CGC-12)

of 26 JUNE 1989

Subject : Draft proposal for a specific programme for the preparation of the industrial development of a practical EUROTRA system

At its meeting of 26 June 1989 the Committee has examined the revised draft of the Communication to Council and the proposal for a Council Decision annexed to it concerning a specific programme for the preparation of the development of a practical EUROTRA system (doc CGC-12/193/89 rev.1).

The Committee notes that the programme proposed covers the actions foreseen under line 8.4 of the Framework Programme 1987-1991 which includes with regard to linguistic problems in addition to the completion of the EUROTRA programme decided by Council on 20 June 1989 :

- support to the industrial development of a machine translation system;
- development of methods and tools for the reusability of lexical resources in computerized applications and;
- creation of standards for lexical and terminological data.

The Committee notes that for this specific programme 10 million ECU are earmarked, of which 5 will become available only on 1992 as the contribution to the overhang of the Framework Programme, and that therefore the objectives of this programme must be necessarily limited.

The Committee endorses the objectives and priorities set out in the proposal and considers them realistic with respect to the resources and time scales of this programme. It stresses the importance of a Community policy in the domain of lexicography.

The Committee welcomes the proposed mode of execution of the programme by which in addition to R&D carried out by national research teams under contracts of association, cooperative projects are foreseen which are to stimulate the involvement of industry through cooperation with universities. It considers also that the use of 10% of the resources earmarked for training through grants and exchange of researchers is an effective measure to improve the situation with regard to the scarcity of qualified researchers and engineers in this subject field.

The Committee reiterates the urgent need for a global Community policy and a strategic R&D programme in the field of linguistic technologies, which it expressed in its opinion of 27 February 1989, and expects that its recommendations will be taken into account in the Commission's proposal for the revision of the Framework Programme.

SPECIFIC PROGRAMME FOR THE PREPARATION OF THE DEVELOPMENT
OF AN OPERATIONAL EUROTRA SYSTEM

FINANCIAL STATEMENT

1. Budget item concerned : Item B7389 (EUROTRA)

2. Legal basis

Article 130 Q, (2) of the EEC Treaty
Council Decision 87/516/Euratom, CEE of 28 September 1987⁽¹⁾
Council Decision / / CEE of ⁽²⁾

3. Description, aims and justification

3.1. Description

The objective of the programme is to support the development of a machine translation system by turning the achievements of the current EUROTRA programme, namely, an operational system prototype in a limited field and for limited categories of text, into the basis for further development .

The programme spans nine action lines, from the implementation of a system development, testing and research environment, through activities leading to the reusability of lexical and terminological resources, to a final evaluation by independent experts. The various lines will have different contractual and financial schemes.

3.2 Aims

On completion of the programme, the following results are expected :

- a set of software tools and a tested methodology for large scale system development;
- a set of tools and methods for building large-scale dictionaries from existing lexical resources;
- a set of tools and methods for the incorporation of terminological collections;
- well tested implementations of analysis and synthesis modules in all nine Community languages with broad grammatical coverage and a representative lexical coverage;

(1) OJ N°. L 302 of 24.10.1987, p. 1

(2) OJ N°. L of

- the implementation of relayed transfer if the corresponding experimentation is successful;
- a strategy for the industrial development of a practical multilingual translation system and of a number of related products;
- a strategy for the development of a generalized linguistic technology reusable in a large variety of applications.

3.3. Justification

The main objective of the programme is to create, starting from the EUROTRA prototype, the appropriate conditions for a large-scale industrial development. This includes the development of methods and tools for the reusability of lexical resources in computer applications as well as the creation of standards for lexical and terminological data.

4. Financial implications

The overall costs of the duration of the programme, 1.7.1990 - 30.6.1992, break down as follows ('000 ECU) :

(a) System development and testing environment	2 000
(b) Community contribution to the national research teams	4 000
(c) Cooperative projects	3 000
(d) Training, subsidies, evaluation	1 000
	<hr/>
Total	10 000

To this must be added 7 million ECU of national contributions (4 million to the national research teams and 3 million ECU of contributions from participants to cooperative projects) bringing the total cost of the programme to 17 million ECU.

The planned financing of the programme, 1990-1992, is the following :

Item B7389

Budgeted research appropriations :

<u>Year</u>	<u>Commitments</u>	<u>Payments</u>
1990	2 000*	100
1991	3 000	2 800
1992	5 000**	5 000
1993	-----	2 100

* earmarked for the specific programme in the 1990 preliminary draft budget

** contribution to the overhang of the Framework Programme, 1987-1991

Budgeted operating appropriations :

The preliminary draft budget for 1990 foresees under item B7389 five temporary posts (2A + 1B + 2C) for this specific programme. The administrative expenditure related to these posts has been included in the appropriations requested for chapter A29 of the 1990 preliminary draft budget.

DOCUMENTS

EN

16

Catalogue number : CB-CO-89-579-EN-C

ISBN 92-77-55706-0

Office for Official Publications of the European Communities
L-2985 Luxembourg