LANGUAGE PROCESSING R&D PROGRAMMES DIRECTORATE XIII E OF THE EUROPEAN COMMISSION

Roberto Cencioni, Giovanni Battista Varile

{Roberto.Cencioni,Nino.Varile}@eurokom.ie

European Commission
DG XIII E 4 - Jean Monnet Building B4/002
L-2920 Luxembourg

THE PRESENT: LINGUISTIC RESEARCH AND ENGINEERING

The Linguistic Research and Engineering (LRE) programme was launched in 1992 with the aims of (a) preparing a more substantial follow-up action in language engineering and (b) prototyping an RTD scheme for language processing characterized by a substantial *involvement of the private sector*.

The actions covered by LRE include *application projects, language resources* and *tools, standardization, assessment* and *evaluation,* and *applied research,* for a total of 26 projects. A number of *accompanying measures* have also been put in motion. Work started in late 1992 and first results are expected in early 1995.

A substantial part of the LRE funds are devoted to *application projects* ranging from full text retrieval through message extraction, categorization and routing, software localization, foreign language reading aids, to automated translation of simplified texts and support tools for human translation.

LRE actions on *language resources* include methods and tools for corpora, lexica and terminology collections. The development of a common software platform, ALEP, is currently reaching beta its testing phase.

In order to favor the reuse of expensive language resources, LRE is promoting through the EAGLES group consensus building amongst major European projects on a common resource encoding scheme.

Other action concern funding of *industrial interest groups*, namely in the area of controlled languages (e.g. for product and maintenance documentation) and full text retrieval.

Finally, two small-scale *international collaboration* projects are currently being supported by LRE, namely a state-of-the-art survey in Natural Language Processing and Speech in collaboration with the National Science Foundation and a project on multi-lingual corpora.

THE FUTURE: LANGUAGE ENGINEERING

The Language Engineering (LE) programme will be more important than LRE and give a major new focus to initiatives in the area of natural language processing, speech technology, and machine translation, previously dispersed over different programmes.

Pilot application projects will be the major means for stimulating collaboration between *users, suppliers and researchers*. Building on the successful LRE scheme, these projects will provide a more *user-centered* context and encourage a coordinated and market-driven approach to RTD tasks, and at the same time demonstrate and validate the integration of language components within information and communications systems and services.

While substantial resources will be devoted to selected short- to medium-term objectives, support will be provided for longer-term endeavors, leading to more advanced and effective technologies.

The focus of the LE programme will also be on *application projects* aiming at providing pilot systems and services covering areas relating to advanced document creation, translation and management, interactive information services, inter-personal communication and language acquisition and use.

Support actions will be carried on in the area of language resources such as corpora, lexica and terminology, software tools, and general research focused on the integration of Natural Language and Speech processing.

Continued support for *standardization and validation* actions will be ensured, and include assessment, evaluation and bench marking.

Finally, *accompanying measures* will cover valorization and dissemination of R&D work, training, support for user and industrial interest groups, and international collaboration.