

**Introduction of the Conference Chairman:  
Antonio Zampolli  
President of ELRA**

---

The first LREC, held in Granada in May 1998, and the second one, held in Athens in May-June 2000, registered an unexpectedly high number of submissions, acknowledging the need for a common venue for the variety and richness of activities in the field of language resources (LRs) and evaluation.

The further increase in the number of submissions received for this third LREC edition, LREC 2002, is an implicit recognition of the success of the first two and of the contribution of this initiative to the formation of a R&D community in a field of increasing relevance in the context of the blooming Information Society (IS).

Numerous actors are working in different sectors, on different aspects of LRs, focusing on issues of particular relevance for their professional interests: linguists, field linguists, psycholinguists, computational linguists, language engineers, publishers of multimedia products, cultural organisations, software and telecommunication industries, education and language technologists, knowledge engineers, service providers on telecommunication networks, etc. As they belong to different communities, which have their own specific organisations and conferences, they seldom have the opportunity of a common venue to share capabilities, exchange information, and explore possible synergies and co-operation.

LREC aims at providing such a venue, promoting the awareness that all those working in the field of LRs will benefit from considering themselves as members of a well-identified field. As stated in the Conference Announcement, the aim of this Conference is "to provide an overview of the state-of-the-art, discuss problems and opportunities, exchange information regarding language resources, their applications, ongoing and planned activities, industrial use and requirements, discuss evaluation methodologies and demonstrate evaluation tools, explore possibilities and promote initiatives for international co-operation in the areas mentioned above".

In the Introduction to the Granada and the Athens Proceedings, I summarised my perception of the main challenges for the future of our fields, both from the organisational and from the technological point of view. I think that this general framework still holds. In my personal opinion, for some of those topics work is effectively taking place; some topics have not yet received the attention they deserve; strategic choices are still pending for some relevant organisational problems.

In the last years we have witnessed a remarkable acceleration in the development of Internet, mobile and broadcasting networks, digital multimedia, etc.

Natural languages are the main vehicles for digital content products and services in business, education and culture.

The cultural treasures of different nations should be made available to everyone. Citizens of various countries wish — and should have the right — to access the digital contents and services in their own languages and using appropriate modalities.

An increasingly high volume of electronic information is becoming available in languages other than English. This requires a concerted effort in the customisation of content for the needs of different cultures, the extension of technology to different languages, the development of a variety of enabling interfaces: spoken language, text and other modalities.

Individual organisations and countries have been investing in the creation of language resources and the methods for the evaluation of resources, technologies, products and applications. This is evident in the US DARPA HLT programme, the EU HLT programme under FP5-IST, the Francophone AUF programme and others, particularly in the localisation industry (Localisation Industry Standards Association, LISA, and Localisation Research Centre, LRC). The European 6<sup>th</sup> Framework Programme (FP6), planned for a start in 2003, includes multilingual and multisensorial communication as one of the major R&D issue, and the evaluation of technologies appears as a specific item in the Integrated Project instrument presentation. Substantial mutual benefits can be expected from addressing these issues through international co-operation. A key element to international co-operation is sharing knowledge about problems and solutions.

I take the liberty of mentioning a few examples of initiatives aiming at promoting the international co-operation which were launched in the last two years.

With the support of the European Commission, within the 5<sup>th</sup> Framework Programme, we are running the ENABLER project, which proposes a network aiming at improving co-operation among national activities established by national authorities for providing LRs for their languages.

In association with ENABLER, in the framework of the e-Content Programme, the European Commission will also very likely support the INTERA project, which aims at building an integrated European language resource area by connecting international, national and regional data centres and at producing new multilingual language resources. We are also actively involved in the Network-DC project, which aims at co-ordinating the actions of ELRA and the Linguistic Data Consortium (LDC).

ELRA and ELSNET are supporting the foundation and the first steps of an International Federation of Language Resources, an initiative which aims at promoting the co-operation between COCOSDA, which is concerned mainly with spoken language resources, and an analogous structure concerned with written language resources. This federation, which will have representatives in each continent, will also create some working groups for multimodal resources and, in general, for resources needed for technological developments.

Another initiative was taken by the ISO/TC 37/SC 4, which founded the Language Resource Management Group, whose purpose is to develop standards that contain specifications for computer-assisted language

resource management, while focusing on data modelling, mark-up, data exchange and the evaluation of language resources other than terminologies.

Other initiatives will be presented and discussed in various workshops associated to this conference.

### **Acknowledgements**

It is my duty and pleasure, in my capacity of Chairman of LREC, to thank all those who have contributed to the preparation of LREC and the publications of these Proceedings.

The members of the ELRA Board have endorsed with enthusiasm our proposal of making this Conference biennial and have been generous with practical advice and suggestions.

The Board of ELRA has agreed to trust the organisation of LREC 2004, which will be probably held in Lisbon, to the same Program Committee, chaired by myself and with the same structure as long as the Program Committee members wish to continue.

The Local Committee — Angel Martin Municio, Luis Jollanes Aguilar, Antonio Ocon Carreras, Carmen Paz Suarez Araujo and Manuel Gonzales Rodriguez — has spared neither energy nor effort in preparing LREC 2002.

I would like to express our deepest gratitude to the eminent members of the International Advisory Committee — Sture Allen, Georges Carayannis, Roberto Cencioni, Zhiwei Feng, Hiroya Fujisaki, Mark Maybury, Bernard Quemada, Ann Souguil, Gary Strong, Piet G. J. Van Sterkenburg, Giovanni Battista Varile and Charles Wayne — coming from Academia, Industry, International Organisations, whose prestigious co-operation clearly underlines the high cultural, scientific, social and strategic role and implications of the topics of the Conference.

Each of the about 500 submissions has been reviewed by at last three reviewers: they deserve particular acknowledgement for having accepted to review a number of submissions far greater than that originally foreseen. In fact, given the quantity of submissions received, the Programme Committee was forced to assign, on average, to the members of the Scientific Committee, nearly twice the number of submissions they had accepted to review, while keeping unchanged the time span allowed for the review process.

As for the first two LREC, the Programme Committee has been at the heart of all the scientific, organisational and financial measures and decisions. The members have practically followed day-by-day the preparation of the Conference, dedicating an impressive quantity of time and effort to monitoring progress and solving problems. The Committee has worked along the lines already proven for Granada and Athens, collectively sharing responsibility for the overall scientific and organisational structure of the Conference, which has been established on the basis of the results of careful preparatory work, which was distributed among the Committee members for three major areas: NLP; speech multimodality and speech evaluation; terminology and evaluation for NLP.

Let me now express my deep and sincere gratitude to the individual members of the Programme Committee, who have been of invaluable help for me in a task which grew continuously with the increasing number of submissions and the differentiation of the initiatives.

Joseph Mariani, for sharing day-by-day with me the responsibility, for sometimes, difficult scientific and organisational decisions and for providing invaluable scientific competence and administrative experience.

Daniel Tapias and Bente Maegaard, for contributing greatly to the scientific programme of the conference, and Daniel, for being instrumental in keeping the organisational situation under control.

Angel Martin Municio, for being instrumental in ensuring the liaison with our Golden sponsor Telefónica Foundation of Spain, and the Spanish Government

Khalid Choukri, for inspiring everyone with his enthusiasm, energy and efficiency, and the entire ELDA staff, for their support.

Nicoletta Calzolari, for her dedication and invaluable contribution in designing the structure and finalising the organisation of the scientific program. In particular, she deserves our gratitude and recognition for supporting largely the burden of my tasks, which my not very good health condition prevented me to carry out completely during the last months.

The workshops, whose Proceedings are published in separate volumes, have contributed greatly to the scientific relevance of the Conference, both for the quality of the work and the choice of topics. Their organisers deserve the gratitude of all the participants to LREC.

The active participation of outstanding representatives of the major Funding Agencies has given a particular significance to LREC and has confirmed, together with the participation of national and international authorities who will honour LREC with their presence, the high strategic value of LRs and Evaluation for the development of a user-friendly Information Society.

I would like to personally thank all the authors of the papers published in the Proceedings: the number, variety and quality of the contributions will provide one of the most complete overviews and reliable definitions of the field of LRs and Evaluation, of its state of the art, of the progress accomplished in the last two years and of the main trends and research directions.

This large participation would not have been possible without the precious co-operation of the associations and consortia who have joined ELRA in promoting LREC (ACL, AFNLPA, ALLC, CLASS, COCOSDA, ORIENTAL COCOSDA, EAFL, EAFT, EAGLES/ISLE, ELSNET, ENABLER, EURALEX, FRANCIL, ISCA, ISCP Permanent Committee, LDC, ONTOWEB, PAROLE, TEI etc.) and of the major national and international organisations, including the Information Society DG of the European Commission, DARPA, NSF and the Japanese Project for International Co-ordination of East-Asian Spoken Language Resources and Evaluation.

In addition, I would like to thank all of the numerous authorities, companies and organisations that have given time and resources in support of LREC.

In particular, many thanks are due to the Telefónica Foundation of Spain and to the Spanish Government, which have generously contributed to the organisation of LREC 2002.

I would like, in particular, to express our thanks and gratitude to the staff of the ELRA distribution agency (ELDA), which contributed to the organisation of LRECs and provided the necessary support during the preparatory phases as well as during the conference: Valérie Raymond, Valérie Mapelli, Magali Jeanmaire and Alexandre Perez.

Sara Goggi, of the Consorzio Pisa Ricerche, and Sergio Rossi, of the Institute of Computational Linguistics of the National Research Council of Pisa, have taken the full responsibility of creating and managing the database of submissions and have been instrumental, through the timely and work-intensive updating of the information, for making it possible to keep the review process within the very short time allowed.

Paola Baroni, of the Consorzio Pisa Ricerche, has been instrumental for the drawing up of my introduction.

Mabel Vargas, supported by the Department of Linguistics of the University of Pisa, has been instrumental for the relationships with the Local Committee.

Antonio Zampolli

Chair

Third International Conference on Language Resources and Evaluation