

Introduction of the Conference Chair

Nicoletta Calzolari

Let me first express to His Majesty Mohammed VI, King of Morocco, the Program Committee gratitude for his Royal Patronage of this 6th edition of LREC.

This year LREC celebrates its 10th birthday! It started in 1998 in Granada, from a great and visionary idea of Antonio Zampolli, and it was then really an adventure and a challenge. I still remember Antonio saying “we won’t even have 50 participants” ... and we had about 500. When LREC was established, Language Resources (LRs) – and with them Evaluation – were only just starting to receive by larger sections of the HLT (Human Language Technology) community the attention that for many years was given to other aspects of language technology (LT). After only 10 years, LREC has become one of the big and established conferences in the sector.

This year LREC has established a new record: we received more than 900 submissions, we have accepted 645 papers, and we may have more than 1,000 participants. These figures are quite impressive to me, and are a sign of the great vitality of the broad area of LR and Evaluation.

LREC is judged – I believe – a great observation post for feeling the pulse of today’s initiatives in the field, with the possibility not just of listening to the “best or most innovative” method or approach, but of examining the large variety of approaches, of analysing the many angles, of exploring the variety of resources for many languages, the new emerging trends, the large projects and initiatives, the infrastructures. I believe that this broadness of themes, topics, perspectives, is an essential contribution to form a better global vision of our field and thus to stimulate new ideas.

We do not have any longer to defend or promote the ‘data-driven’ approaches: they are pervasive and have a well-deserved and ample recognition. Large-scale LR are unanimously recognised as the necessary infrastructure underlying LT. It is the merit (mostly) of LR that LT acquires the maturity and attains the robustness needed to become truly usable in real world applications. This is probably the major result of LR, with an impact on the transformation of LT from ‘just’ an R&D sector to a technology with a great impact in the society.

We can say today that a ‘LR community’ exists: we simply have to observe the large LREC attendance. The achievement of a worldwide linguistic infrastructure, however, requires the coverage not only of a range of technical aspects, but also – and maybe most critically – of a number of organisational and coordination aspects. The stable growth of the field brings in itself some sort of revolution, and a need now to converge. An essential element for ensuring an integrated basis is to enhance the cooperation among many communities now acting separately, such as LR and LT developers, written, speech and multimodality specialists, terminology, semantic web and ontology experts, content providers, linguists and so on. This is one of the challenges for the next years, for a usable and useful ‘language’ scenario in the global and multilingual network. I believe that LREC may play an important role in the integration of these various communities.

This does not mean that there are no infrastructural issues still to be discussed and solved. Let us only think about the problem of maintenance of LR, or the big issues of interoperability and sharing. The big growth of the field should be complemented by a reflection – within the same community – on priorities and future strategies. It is a big achievement, and a great opportunity for our field, that recently a number of strategic-infrastructure initiatives have started, or are going to start, in all the continents. This is also a sign that funding agencies recognise the strategic value of our work and the

importance of helping a coherent growth also through a number of coordinated actions. LREC, together with its workshops, is the place where these – and future – initiatives will be presented, discussed, and promoted.

LREC, both the main conference and its workshops/tutorials, deals with the topics of LRs and evaluation within all the modalities: written and spoken language, and multimodality. Which are the main themes and issues at LREC 2008? Some very sketchy observations delineating some trends:

- There is a lot of semantics, knowledge, ontologies, content, analysed from every possible angle, from representation and annotation to acquisition, from spatio-temporal information, coreference and discourse, to emotions, affect, opinions.
- Complementary to this, even if with less emphasis, also syntactically annotated resources and related tools continue to receive attention.
- Among the LR-related tools and systems, information extraction is still a major issue, not yet solved, and more and more ambitious; well represented are also named entity recognition, document classification, question answering and summarisation.
- In speech: broadcast news processing, speaker identification, pronunciation data, speech synthesis, recognition, dialogue, large varieties of speech corpora for many languages, but also the topics of affect and emotion.
- Multimodality and multimedia span from TV and video processing to sign language, communication and dialogue in multimodal environment, emotion and subjective content again.
- As usual, large infrastructures, architectures, big initiatives and large projects, together with the issues of interoperability and standardisation, receive wide attention at LREC.
- Many papers, as expected, on lexicons, corpora and many different types of related tools.
- Evaluation of systems and validation of resources, both for written and spoken language, are broadly represented, as well as evaluation methodologies and evaluation campaigns.
- Many papers are on multilingual resources and systems and machine translation applications, but also dialects and language varieties.
- Terminology is an important issue, and the most represented terminological domain is the bio-medical one.

As another sign of the great significance and the success of the field, LREC is now somehow complemented by the *Language Resources and Evaluation* journal, endorsed by ELRA and co-edited by Nancy Ide and me. Springer offers LREC 2008 participants a subscription to the journal at a special rate. Moreover, we will ask the authors of the papers mostly recommended by reviewers as appropriate for the journal if they want to submit a longer version to the journal.

I am also proud to announce that we have decided to make all the LREC Proceedings, together with the proceedings of accompanying workshops, available on the web as a service to the community. It will happen soon after this conference.

Few last notes, on which we would like to get your opinions. You see in the program that we experiment this year a new conference structure, with Posters in parallel with Oral sessions. This also allowed us to accommodate more posters. We have also accepted more Tutorials and Workshops than usual, because of the large number of good submissions. We invite your feedback on these changes. We are also asking ourselves, given the so high – and ever increasing – number of submissions, if we should experiment with a 4-days conference: we will ask you about this too.

Acknowledgments

First of all my thanks go to the Program Committee, that at every LREC has a harder job, having to deal with an ever increasing number of submissions, of a better quality.

Also on behalf of all the members of the Program Committee, I warmly thank all the various Committees that have made this LREC possible, and hopefully successful.

I thank ELRA and the ELRA Board, for their continuous commitment to LREC.

I thank our impressively large Scientific Committee, composed of about 500 colleagues from all over the world. They did a wonderful job, succeeding to complete their reviews in time for so many papers. We are indebted to the International Advisory Board and the Local Advisory Board, that have provided moral support to our Conference, and to the Local Organising Committee, and in particular Abdelhak Mouradi, for helping in finding solutions to local issues.

I am grateful to authorities, associations, organisations, committees, agencies, companies that have supported LREC in various ways, for their important cooperation.

I express my big gratitude to all the sponsors that have believed in the importance of our conference, and have helped with economic support.

I thank the workshop, tutorial, and panel organisers, who surround LREC of so many interesting events. A big thanks goes to all the authors, who provide the 'substance' to LREC, and give us such a broad picture of the field.

I wish finally to thank the two institutions that have provided economic support and dedicated so much effort, in terms of manpower, to this LREC, as to the previous ones, i.e. ELDA in Paris and my institute, ILC-CNR in Pisa. Without their commitment LREC would not have been possible.

The last, but not least thanks, are thus devoted, with all my sympathy, to the people of these institutions who have worked so intensely to make this LREC possible in all its details. Despite the distance (Paris-Pisa) they have worked together as a unique team, with enthusiasm and dedication. My biggest thanks go to them: Roberto Bartolini, Olivier Hamon, Vincenzo Parrinelli, Valeria Quochi, Mathieu Robin-Vinet, Sergio Rossi, and in particular Sara Goggi and H el ene Mazo who have become over the years one of the pillars of LREC. I cannot list the many tasks they carried out, but I can say for sure that without their daily work and real commitment since many months, LREC would not happen. We have solved together many big and small problems of such a large conference. They, together with other young researchers of these two institutions, will assist you during the days of the tutorials/workshops and the conference.

Now LREC is in your hands, the participants. You are the protagonist of LREC, you will make this LREC great (I am sure). So at the very end my greatest thanks go to you all. I may not be able to speak with each of you during the Conference (I'll try). I hope that you learn something, that you perceive and touch the excitement, fervour and liveliness of the field, that you have fruitful conversations (conferences are useful also for this), most of all that you profit of so many contacts to organise new exciting work and programmes in the field of LRs and evaluation, which you will show at the next LREC.

I particularly hope that funding agencies all over the world are impressed by the quality and quantity of initiatives in our sector that LREC displays, and by the fact that the field attracts practically all the best groups of R&D from all continents. This is a sign they must take into account in their programmes and funding strategies. The success of LREC means to us in reality the success of the field of LRs and Evaluation.

With all the Programme Committee, I welcome you at LREC 2008 in such a wonderful country as Morocco and wish you a fruitful Conference.

Enjoy LREC in Marrakech!

Nicoletta Calzolari
Chair

6th International Conference on Language Resources and Evaluation