

Special Event  
**LREC 15th Anniversary**

Activities in the area of Language Resources and Evaluation greatly increased over the past 30 years, due to the importance of Language Resources to conduct research investigations in language sciences and to develop language processing systems which are based on automatic Machine Learning.

Some milestones may be identified in this area, such as the launching of the evaluation campaigns of speech recognition systems by NIST for DARPA in 1987, the creation of the Linguistic Data Consortium (LDC) and of the Coordinating Committee on Speech Databases and Speech Input/Output Systems Assessment (Cocosda) in 1991. This was followed by the launching of the European Language Resources Association (ELRA) in 1995, which organized the first Language Resources and Evaluation (LREC) conference in 1998. The oriental branch of Cocosda organized the Oriental-Cocosda conference for the first time on the same year. The Language Resources and Evaluation Journal published by Springer was initiated in 2005.

The idea of adding a scientific dimension to the Language Resources distribution activity provided by ELRA through an international conference specifically devoted to Language Resources and Evaluation was first proposed by Joseph Mariani in 1997. The first conference was held in 1998 in Granada (Spain). It was organized and chaired by Antonio Zampolli. Following its great success, the LREC conference has been organized every two years since then and is now chaired by Nicoletta Calzolari, with the continuous support of Khalid Choukri from the very beginning. Nowadays, LREC conferences gather more than 1,200 participants which form a strong interdisciplinary community.

The LREC 2014 conference gives the opportunity to warmly acknowledge and thank those who contributed to make the whole enterprise successful on the occasion of the LREC 15<sup>th</sup> anniversary.

It was felt useful to reconsider the last 15 years of research in the area of Language Resources and Evaluation through an analysis of the proceedings of the former 8 LREC conferences gathered in the LREC Anthology, with the purpose to illustrate the strong dynamism and the multiple facets of our domain and to inspire us in building up the steps for the forthcoming 15 years.

This analysis is similar to comparable exercises which were conducted in the Computational Linguistics community through an analysis of the ACL Anthology which covered 50 years of the ACL conferences and was presented within a specifically dedicated workshop at the 2012 ACL conference in Jeju (Korea), or in the Spoken Language Processing community through the analysis of the ISCA Archive which covered 25 years of the ECST, Eurospeech, ICSLP and Interspeech conference series and was presented at the 2013 Interspeech conference in Lyon (France). A previous exercise over 15 years of the IEEE ICASSP conference was conducted by the end of the 1980s and served for deciding the launch of the ESCA Eurospeech conference.

A Quiz is also being organized on the content of the analysis during the LREC 2014 conference.

The survey will be presented during the Closing Session on Friday May 30, and the results of the Quiz with the name of the winner(s) will then be reported.

*Joseph Mariani*  
On behalf of the conference Program Committee

## ABSTRACT



### **Rediscovering 15 Years of Discoveries in Language Resources and Evaluation: The LREC Anthology Analysis.**

*Joseph Mariani<sup>1,2</sup>, Patrick Paroubek<sup>1</sup>, Gil Francopoulo<sup>2,3</sup>, Olivier Hamon<sup>4</sup>*

<sup>1</sup>LIMSI-CNRS, <sup>2</sup>IMMI-CNRS, <sup>3</sup>Tagmatica, <sup>4</sup>formerly at ELDA

This paper aims at analyzing the content of the LREC conferences contained in the ELRA Anthology over the past 15 years (1998-2013). It follows similar exercises that have been conducted, such as the survey on the IEEE ICASSP conference series from 1976 to 1990, which served in the launching of the ESCA Eurospeech conference, a survey of the Association of Computational Linguistics (ACL) over 50 years of existence, which was presented at the ACL conference in 2012, or a survey over the 25 years (1987-2012) of the conferences contained in the ISCA Archive, presented at Interspeech 2013. It contains first an analysis of the evolution of the number of papers and authors over time, including the study of their gender, nationality and affiliation, and of the collaboration among authors. It then studies the funding sources of the research investigations that are reported in the papers. It conducts an analysis of the evolution of the research topics within the community over time. It finally looks at reuse and plagiarism in the papers. The survey shows the present trends in the conference series and in the Language Resources and Evaluation scientific community. Conducting this survey also demonstrated the importance of a clear and unique identification of authors, papers and other sources to facilitate the analysis. This survey is preliminary, as many other aspects also deserve attention. But we hope it will help better understanding and forging our community in the global village.