

EACL-2006

**11th Conference
of the European Chapter of the
Association for Computational Linguistics**

Proceedings of the workshop on

**Multilingual Question
Answering – MLQA'06**

Volume Editors

Anselmo Peñas
(UNED, Spain)
anselmo@lsi.uned.es

Richard Sutcliffe
(University of Essex, England)
rsutcl@essex.ac.uk

ISBN: 2-9524532-4-1

The conference, the workshop and the tutorials are sponsored by:



Center for the Evaluation of Language and Communication Technologies

Celct
c/o BIC, Via dei Solteri, 38
38100 Trento, Italy
<http://www.celct.it>

XEROX

Research Centre Europe

Xerox Research Centre Europe
6 Chemin de Maupertuis
38240 Meylan, France
<http://www.xrce.xerox.com>



CELI s.r.l.
Corso Moncalieri, 21
10131 Torino, Italy
<http://www.celi.it>

THALES

Thales
45 rue de Villiers
92526 Neuilly-sur-Seine Cedex, France
<http://www.thalesgroup.com>

EACL-2006 is supported by

Trentino S.p.a.  and Metasistem Group 

© April 2006, Association for Computational Linguistics

Order copies of ACL proceedings from:
Priscilla Rasmussen,
Association for Computational Linguistics (ACL),
3 Landmark Center,
East Stroudsburg, PA 18301 USA

Phone +1-570-476-8006
Fax +1-570-476-0860
E-mail: acl@aclweb.org
On-line order form: <http://www.aclweb.org/>

Preface

This volume contains the papers presented at the Workshop on Multilingual Question Answering MLQA'06, held on 4 April 2006 in conjunction with the European chapter of Association for Computational Linguistics Conference, EACL 2006.

Despite the great deal of attention that Question Answering (QA) has received in recent years due to the landmark Q&A Track at the Text REtrieval Conference (TREC), multilinguality has been outside the mainstream of QA research, which has focused mainly on the English language. Multilingual QA has emerged only in the last few years as a complementary research task, representing a promising direction for at least two reasons. First, it allows users to interact with machines in their native languages, contributing to easier, faster, and more equal information access. Second, cross-lingual capabilities enable QA systems to access information stored only in language-specific text collections.

The success of the QA track in recent evaluation campaigns at CLEF (Cross-Language Evaluation Forum) and NTCIR (Evaluation of Information Access Technologies) has shown increasing interest in both monolingual non-English QA (questions and answers in the same language) and in cross-lingual QA, where the question is posed in a source language and the answer must be found in a target collection of a different language.

The workshop intended to address a number of challenging issues for research in multilingual QA, including searching in multilingual document collections, collecting and merging answers found in documents from different languages, using heterogeneous multilingual data collections (such as the Web, XML data or data in relational databases) for answer generation, and interpreting questions in different languages. Additionally papers related to user studies or case studies where multilingual QA has played a major role were most welcome. Relevant topics for the workshop included:

- Monolingual QA systems for languages other than English;
- Cross-lingual QA systems;
- Multilingual Web-based QA;
- Case studies and user studies of Multilingual QA;
- Specific Translation techniques for Cross-lingual QA;
- The influence of translation on Cross-lingual QA;
- Multilingual QA in restricted vs. open domains;
- QA from heterogeneous multilingual data collections such as unstructured text, semi-structured data and structured data;
- Techniques for porting QA systems among languages;
- The effectiveness of particular techniques in different languages;
- Comparisons between QA in different languages including European and Asian languages;
- Evaluation of Multilingual QA systems;
- Resources for Multilingual QA.

The ten papers finally accepted fall in four main categories: Analysis of QA Multilingual technology, Multilingual and Translingual QA systems, QA systems in other languages than English, and finally, Adaptation of existing QA systems.

We would like to thank all the authors whose research contributions have made this workshop possible. Thanks are also due to Maarten de Rijke and Caroline Sporleder as EACL workshops chairs. Special thanks to Bernardo Magnini who promoted the idea of the workshop and to the Program Committee for their work in all the stages of the organization, from the very proposal to their careful reviewing.

Anselmo Peñas and Richard Sutcliffe

This Workshop has the partial support of the Center for the Evaluation of Language and Communication Technologies (CELCT), the Cross-Language Evaluation Forum (CLEF) and the R2D2 project (Question Answering in Digital Documents, TIC-2003-07158-C04) financed by the Science & Technology Department of Spain 2003-2006.

Workshop Organizers

Anselmo Peñas (UNED, Madrid, Spain)
Richard Sutcliffe (University of Essex¹, England)
(¹) On Sabbatical from University of Limerick

Programme Committee

Lili Aunimo (University of Helsinki, Finland)
Christelle Ayache (ELDA/ELRA, Paris, France)
Johan Bos (University of Rome "La Sapienza", Italy)
Gosse Bouma (University of Groningen, The Netherlands)
Hsin-Hsi Chen (National Taiwan University, Taiwan)
Danilo Giampiccolo (CELCT, Trento, Italy)
Brigitte Grau (LIMSI-CNRS, France)
Donna Harman (NIST, USA)
Michael Hess (University of Zürich, Switzerland)
Valentin Jijkoun (University of Amsterdam, The Netherlands)
Noriko Kando (NII, Tokyo, Japan)
Dominique Laurent (Synapse Développement, Toulouse, France)
Bernardo Magnini (ITC-irst, Trento, Italy)
André Martins (Priberam Informática, Lisbon, Portugal)
Diego Mollá (Macquarie University, Australia)
Günter Neumann (DFKI, Saarbrücken, Germany)
Carol Peters (CNR-ISTI, Pisa, Italy)
Diana Santos (Linguatca, Oslo, Norway)
Yutaka Sasaki (ATR, Kyoto, Japan)
Kiril Simov (Bulgarian Academy of Science, Sofia, Bulgaria)
Bill Teahan (University of Wales, Bangor, Wales)
José Luis Vicedo (University of Alicante, Spain)

Table of Contents

Analysis of Multilingual QA technology

<i>QA better than IR ?</i> Dominique Laurent, Patrick Séguéla, Sophie Nègre	1
<i>The Effect of Machine Translation on the Performance of Arabic-English QA System</i> Azzah Al-Maskari, Mark Sanderson	9
<i>Cross-Cutting Aspects of Cross-Language Question Answering Systems</i> Bogdan Sacaleanu, Günter Neumann	15

Multilingual and Translingual QA systems

<i>Evaluation and Improvement of Cross-Lingual Question Answering Strategies</i> Anne-Laure Ligozat, Brigitte Grau, Isabelle Robba, Anne Vilnat	23
<i>Keyword Translation Accuracy and Cross-Lingual Question Answering in Chinese and Japanese</i> Teruko Mitamura, Mengqiu Wang, Hideki Shima, Frank Lin	31
<i>BRUJA: Question Classification for Spanish. Using Machine Translation and an English Classifier</i> Miguel Ángel García Cumbreiras, L. Alfonso Ureña López, Fernando Martínez Santiago	39

QA systems in other languages than English

<i>Monolingual Web-based Factoid Question Answering in Chinese, Swedish, English and Japanese</i> E.W.D. Whittaker, J. Hamonic, D. Yang, T.Klingberg, S. Furui	45
<i>Dialogue based Question Answering System in Telugu</i> Rami Reddy Nandi Reddy, Sivaji Bandyopadhyay	53

Adaptation of existing QA systems

<i>Adapting a Semantic Question Answering System to the Web</i> Sven Hartrumpf	61
<i>Experiments Adapting an Open-Domain Question Answering System to the Geographical Domain Using Scope-Based Resources</i> Daniel Ferrés, Horacio Rodríguez	69

