

EACL 2009

**Proceedings of the
Student Research Workshop**

2 April 2009

Megaron Athens International Conference Centre
Athens, Greece

Production and Manufacturing by
TEHNOGRAFIA DIGITAL PRESS
7 Ektoros Street
152 35 Vrilissia
Athens, Greece

©2009 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

Preface

On behalf of the Programme Committee, we are pleased to present the proceedings of the Student Research Workshop held at the 12th Conference of the European Chapter of the Association for Computational Linguistics. Following the tradition of providing a forum for student researchers and the success of the previous workshops held in Bergen (1999), Toulouse (2001), Budapest (2003) and Trento (2006), a panel of senior researchers will take part in the presentation of the papers, providing detailed comments on the work of the authors.

The Student Workshop will run as four parallel sessions, during which 11 papers will be presented. These high standard papers were carefully chosen from a total of 38 submissions coming from 18 countries.

We would like to take this opportunity to thank the many people that have contributed in various ways to the success of the Student Workshop: the members of the Programme Committee for their evaluation of the submissions and for taking the time to provide useful detailed comments and suggestions for the improvement of papers; the panelists for providing detailed feedback directly; and the students for their hard work in preparing their submissions.

We are also very grateful to the EACL for providing sponsorship for students who would otherwise be unable to attend the workshop and present their work. And finally, thanks to those who have given us advice and assistance in planning this workshop (especially Nuria Bertomeu, Alex Lascarides, Joakim Nivre, Konstantinos Stamatakis).

We hope you enjoy the Student Research Workshop.

Vera Demberg, University of Edinburgh
Yanjun Ma, Dublin City University
Nils Reiter, Heidelberg University
EACL 2009 Student Research Workshop Chairs

Program Committee

Program Chairs:

Vera Demberg, University of Edinburgh (UK)
Yanjun Ma, Dublin City University (Ireland)
Nils Reiter, Heidelberg University (Germany)

Program Committee Members:

Eneko Agirre, Basque Country University (Spain)
Timothy Baldwin, University of Melbourne (Australia)
Srinivas Bangalore, AT&T Research (USA)
Yassine Benajiba, Polytechnic University of Valencia (Spain)
Alexandra Birch, University of Edinburgh (UK)
Tamás Biró, University of Groningen (The Netherlands)
Thorsten Brants, Google Inc. (USA)
João Cabral, University of Edinburgh (UK)
Aoife Cahill, Stuttgart University (Germany)
Marine Carpuat, Hong Kong University of Science & Technology (Hong Kong)
Ben Carterette, University of Delaware (USA)
Pi-Chuan Chang, Stanford University (USA)
Ciprian Chelba, Google Inc. (USA)
Trevor Cohn, University of Edinburgh (UK)
Irene Cramer, Dortmund University (Germany)
Ido Dagan, Bar Ilan University (Israel)
Hal Daumé III, University of Utah (USA)
Güneş Erkan, Google Inc. (USA)
Raquel Fernández, University of Amsterdam (The Netherlands)
Sharon Goldwater, University of Edinburgh (UK)
Hany Hassan, Dublin City University (Ireland)
Julia Hockenmaier, University of Illinois (USA)
Akshay Java, Microsoft Live Labs (USA)
Sittichai Jiampojamarn, University of Alberta (Canada)
Gareth Jones, Dublin City University (Ireland)
Alexander Koller, Saarland University (Germany)
Jochen Leidner, Thomson Reuters (UK)
Vanessa Murdock, Yahoo! Research Barcelona (Spain)
Malvina Nissim, University of Bologna (Italy)
Stefan Oepen, University of Oslo (Norway)
Ulrike Padó, Pearson Knowledge Technologies (USA)
Simone Ponzetto, Heidelberg University (Germany)
David Reitter, Carnegie Mellon University (USA)
Bogdan Sacaleanu, DFKI GmbH (Germany)
Richard Sproat, Oregon Health & Science University (USA)
Mark Stevenson, University of Sheffield (UK)
Simone Teufel, University of Cambridge (UK)
Sebastian Varges, University of Trento (Italy)
Rui Wang, Saarland University (Germany)
Haifeng Wang, Toshiba Research & Development Centre (PRC)
Andy Way, Dublin City University (Ireland)
Nick Webb, University at Albany, State University of New York (USA)
Feiyu Xu, DFKI GmbH (Germany)
Yi Zhang, Saarland University (Germany)
Thomas Fang Zheng, Tsinghua University (PRC)

Table of Contents

<i>Modelling Early Language Acquisition Skills: Towards a General Statistical Learning Mechanism</i> Guillaume Aimetti	1
<i>A Memory-Based Approach to the Treatment of Serial Verb Construction in Combinatory Categorical Grammar</i> Prachya Boonkwan	10
<i>Combining a Statistical Language Model with Logistic Regression to Predict the Lexical and Syntactic Difficulty of Texts for FFL</i> Thomas François	19
<i>Finding Word Substitutions Using a Distributional Similarity Baseline and Immediate Context Overlap</i> Aurelie Herbelot	28
<i>Structural Correspondence Learning for Parse Disambiguation</i> Barbara Plank	37
<i>A Chain-starting Classifier of Definite NPs in Spanish</i> Marta Recasens	46
<i>Speech Emotion Recognition With TGI+.2 Classifier</i> Julia Sidorova	54
<i>A Comparison of Merging Strategies for Translation of German Compounds</i> Sara Stymne	61
<i>A Generalized Vector Space Model for Text Retrieval Based on Semantic Relatedness</i> George Tsatsaronis and Vicky Panagiotopoulou	70
<i>Aligning Medical Domain Ontologies for Clinical Query Extraction</i> Pinar Wennerberg	79
<i>Extraction of Definitions Using Grammar-Enhanced Machine Learning</i> Eline Westerhout	88

