

ACL 2012

**50th Annual Meeting of the
Association for Computational Linguistics**

Proceedings of the Student Research Workshop

July 9 - 11, 2012
Jeju Island, Korea

©2012 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

ISBN 978-1-937284-26-8

Introduction

Welcome to the ACL 2012 Student Research Workshop. As was the case last year, this year we have solicited and accepted both Research and Thesis Proposal papers. We accepted 14 out of the 31 submissions that we received from students in a wide variety of countries. Two papers were withdrawn, and 12 will be presented as posters during the main ACL 2012 Poster Session. We have paired senior members of the research community with each student in order to provide feedback and guidance to our student authors.

The overall quality of the submissions was high and we thank our Program Committee for their excellent feedback and reviews. We also thank our Faculty Advisors, Kentaro Inui, Greg Kondrak, and Yang Liu, for their guidance. We were able to provide most students with conference registration and travel stipends thanks to generous support from the U.S. National Science Foundation, the ACL Walker Student Fund, and the Asian Federation of Natural Language Processing. Finally, thank you and congratulations to all of our Student Research Workshop presenters.

Student Co-Chairs

Jackie C. K. Cheung, University of Toronto
Jun Hatori, University of Tokyo
Carlos Henriquez, Technical University of Catalonia
Ann Irvine, Johns Hopkins University

Faculty Advisors

Kentaro Inui, Tohoku University
Greg Kondrak, University of Alberta
Yang Liu, University of Texas at Dallas

Student Program Committee Members:

Shafiq Joty, *University of British Columbia, Canada*
Annie Louis, *University of Pennsylvania, USA*
Courtney Napoles, *Johns Hopkins University, USA*
Jason Naradowsky, *University of Massachusetts at Amherst, USA*
Vahed Qazvinian, *University of Michigan, USA*
Sravana Reddy, *The University of Chicago, USA*
Alan Ritter, *University of Washington, USA*
Nathan Schneider, *Carnegie Mellon University, USA*
Ashish Vaswani, *University of Southern California, USA*
Ainur Yessenalina, *Cornell University, USA*

Non-Student Program Committee Members:

Anabela Barreiro, *INESC-ID, Portugal*
Anja Belz, *University of Brighton, UK*
Steven Bethard, *University of Colorado, USA*
Michael Bloodgood, *University of Maryland, USA*
Giuseppe Carenini, *University of British Columbia, Canada*
Colin Cherry, *National Research Council, Canada*
Marta R. Costa-jussà, *Barcelona Media Research Center, Spain*
Adrià de Gispert, *University of Cambridge, UK*
David Farwell, *Technical University of Catalonia, Spain*
José A. R. Fonollosa, *Technical University of Catalonia, Spain*
Timothy Fowler, *University of Toronto, Canada*
Masato Hagiwara, *Rakuten Institute of Technology, USA*
Maxim Khalilov, *University of Amsterdam, Netherlands*
Alexandre Klementiev, *Saarland University, Germany*
Jonathan May, *SDL Language Weaver, USA*

Yusuke Miyao, *National Institute of Informatics, Japan*
Saif Mohammad, *National Research Council, Canada*
Enric Monte, *Technical University of Catalonia, Spain*
Borja Navarro Colorado, *Universidad de Alicante, Spain*
Daniele Pighin, *Technical University of Catalonia, Spain*
Xian Qian, *University of Texas at Dallas, USA*
Horacio Rodríguez, *Technical University of Catalonia, Spain*
Paolo Rosso, *Technical University of Valencia, Spain*
Kenji Sagae, *University of Southern California, USA*
Helmut Schmid, *University of Stuttgart, Germany*
Lane Schwartz, *Air Force Research Laboratory, USA*
Xu Sun, *Cornell University, USA*
Hiroya Takamura, *Tokyo Institute of Technology, Japan*
Paul Thompson, *Dartmouth College, USA*
Yannick Versley, *Universität Tübingen, Germany*
Theresa Wilson, *Johns Hopkins University, USA*
Yue Zhang, *University of Cambridge, UK*
Geoffrey Zweig, *Microsoft, USA*

Table of Contents

<i>A Broad Evaluation of Techniques for Automatic Acquisition of Multiword Expressions</i> Carlos Ramisch, Vitor De Araujo and Aline Villavicencio	1
<i>Detecting Power Relations from Written Dialog</i> Vinodkumar Prabhakaran	7
<i>Active Learning with Transfer Learning</i> Chunyong Luo, Yangsheng Ji, Xinyu Dai and Jiajun Chen	13
<i>Query classification using topic models and support vector machine</i> Dieu-Thu Le and Raffaella Bernardi	19
<i>Evaluating Unsupervised Ensembles when applied to Word Sense Induction</i> Keith Stevens	25
<i>Topic Extraction based on Prior Knowledge obtained from Target Documents</i> Kayo Tatsukawa	31
<i>TopicTiling: A Text Segmentation Algorithm based on LDA</i> Martin Riedl and Chris Biemann	37
<i>Domain Adaptation of a Dependency Parser with a Class-Class Selectional Preference Model</i> Raphael Cohen, Yoav Goldberg and Michael Elhadad	43
<i>Extracting fine-grained durations for verbs from Twitter</i> Jennifer Williams	49
<i>Discourse Structure in Simultaneous Spoken Turkish</i> Isin Demirsahin	55
<i>A Generic Framework for Multiword Expressions Treatment: from Acquisition to Applications</i> Carlos Ramisch	61
<i>Towards Automatic Construction of Knowledge Bases from Chinese Online Resources</i> Liwei Chen, Yansong Feng, Yidong Chen, Lei Zou and Dongyan Zhao	67

Conference Program

Monday, July 9, 2012, 6:00-8:30pm

A Broad Evaluation of Techniques for Automatic Acquisition of Multiword Expressions

Carlos Ramisch, Vitor De Araujo and Aline Villavicencio

Detecting Power Relations from Written Dialog

Vinodkumar Prabhakaran

Active Learning with Transfer Learning

Chunyong Luo, Yangsheng Ji, Xinyu Dai and Jiajun Chen

Query classification using topic models and support vector machine

Dieu-Thu Le and Raffaella Bernardi

Evaluating Unsupervised Ensembles when applied to Word Sense Induction

Keith Stevens

Topic Extraction based on Prior Knowledge obtained from Target Documents

Kayo Tatsukawa

TopicTiling: A Text Segmentation Algorithm based on LDA

Martin Riedl and Chris Biemann

Domain Adaptation of a Dependency Parser with a Class-Class Selectional Preference Model

Raphael Cohen, Yoav Goldberg and Michael Elhadad

Extracting fine-grained durations for verbs from Twitter

Jennifer Williams

Discourse Structure in Simultaneous Spoken Turkish

Isin Demirsahin

A Generic Framework for Multiword Expressions Treatment: from Acquisition to Applications

Carlos Ramisch

Towards Automatic Construction of Knowledge Bases from Chinese Online Resources

Liwei Chen, Yansong Feng, Yidong Chen, Lei Zou and Dongyan Zhao

