

# **51st Annual Meeting of the Association for Computational Linguistics (ACL 2013)**

## **Student Research Workshop**

**Sofia, Bulgaria  
5-7 August 2013**

ISBN: 978-1-62748-976-8

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2013) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2013)

For permission requests, please contact the Association for Computational Linguistics  
at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006

Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2634  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Categorization of Turkish News Documents with Morphological Analysis</i> Burak Kerim Akkuş and Ruket Cakici .....	1
<i>Crawling microblogging services to gather language-classified URLs. Workflow and case study</i> Adrien Barbaresi .....	9
<i>Patient Experience in Online Support Forums: Modeling Interpersonal Interactions and Medication Use</i> Annie Chen .....	16
<i>Detecting Metaphor by Contextual Analogy</i> Eirini Florou .....	23
<i>Survey on parsing three dependency representations for English</i> Angelina Ivanova, Stephan Oepen and Lilja Øvrelid .....	31
<i>What causes a causal relation? Detecting Causal Triggers in Biomedical Scientific Discourse</i> Claudiu Mihăilă and Sophia Ananiadou .....	38
<i>Text Classification based on the Latent Topics of Important Sentences extracted by the PageRank Algorithm</i> Yukari Ogura and Ichiro Kobayashi .....	46
<i>Automated Collocation Suggestion for Japanese Second Language Learners</i> Lis Pereira, Eryln Manguilimotan and Yuji Matsumoto .....	52
<i>Understanding Verbs based on Overlapping Verbs Senses</i> Kavitha Rajan .....	59
<i>Topic Modeling Based Classification of Clinical Reports</i> Efsun Sarioglu, Kabir Yadav and Hyeong-Ah Choi .....	67
<i>Annotating named entities in clinical text by combining pre-annotation and active learning</i> Maria Skeppstedt .....	74
<i>Multigraph Clustering for Unsupervised Coreference Resolution</i> Sebastian Martschat .....	81
<i>Computational considerations of comparisons and similes</i> Vlad Niculae and Victoria Yaneva .....	89
<i>Question Analysis for Polish Question Answering</i> Piotr Przybyła .....	96
<i>A Comparison of Techniques to Automatically Identify Complex Words.</i> Matthew Shardlow .....	103
<i>Detecting Chronic Critics Based on Sentiment Polarity and User's Behavior in Social Media</i> Sho Takase, Akiko Murakami, Miki Enoki, Naoaki Okazaki and Kentaro Inui .....	110
<i>Addressing Ambiguity in Unsupervised Part-of-Speech Induction with Substitute Vectors</i> Volkan Cirik .....	117

<i>Psycholinguistically Motivated Computational Models on the Organization and Processing of Morphologically Complex Words</i>	
Tirthankar Dasgupta .....	123
<i>A New Syntactic Metric for Evaluation of Machine Translation</i>	
Melania Duma, Cristina Vertan and Wolfgang Menzel.....	130
<i>High-quality Training Data Selection using Latent Topics for Graph-based Semi-supervised Learning</i>	
Akiko Eriguchi and Ichiro Kobayashi.....	136
<i>Simple, readable sub-sentences</i>	
Sigrid Klerke and Anders Søgaard.....	142
<i>Exploring Word Order Universals: a Probabilistic Graphical Model Approach</i>	
Xia Lu.....	150
<i>Robust multilingual statistical morphological generation models</i>	
Ondřej Dušek and Filip Jurčiček .....	158
<i>A corpus-based evaluation method for Distributional Semantic Models</i>	
Abdellah Fourtassi and Emmanuel Dupoux.....	165
<i>Deepfix: Statistical Post-editing of Statistical Machine Translation Using Deep Syntactic Analysis</i>	
Rudolf Rosa, David Mareček and Aleš Tamchyna.....	172