

15th Conference on Computational Natural Language Learning 2011

(CoNLL 2011)

Co-Located with ACL HLT 2011

**Portland, Oregon, USA
23-24 June 2011**

ISBN: 978-1-61839-235-0

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© (2011) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2011)

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

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
Web: www.proceedings.com

Table of Contents

<i>Modeling Syntactic Context Improves Morphological Segmentation</i>	
Yoong Keok Lee, Aria Haghighi and Regina Barzilay	1
<i>The Effect of Automatic Tokenization, Vocalization, Stemming, and POS Tagging on Arabic Dependency Parsing</i>	
Emad Mohamed	10
<i>Punctuation: Making a Point in Unsupervised Dependency Parsing</i>	
Valentin I. Spitzkovsky, Hiyen Alshawhi and Daniel Jurafsky	19
<i>Modeling Infant Word Segmentation</i>	
Constantine Lignos	29
<i>Word Segmentation as General Chunking</i>	
Daniel Hewlett and Paul Cohen	39
<i>(Invited talk) Computational Linguistics for Studying Language in People: Principles, Applications and Research Problems</i>	
Bruce Hayes	48
<i>Search-based Structured Prediction applied to Biomedical Event Extraction</i>	
Andreas Vlachos and Mark Craven	49
<i>Using Sequence Kernels to Identify Opinion Entities in Urdu</i>	
Smruthi Mukund, Debanjan Ghosh and Rohini Srihari	58
<i>Subword and Spatiotemporal Models for Identifying Actionable Information in Haitian Kreyol</i>	
Robert Munro	68
<i>Gender Attribution: Tracing Stylometric Evidence Beyond Topic and Genre</i>	
Ruchita Sarawgi, Kailash Gajulapalli and Yejin Choi	78
<i>Improving the Impact of Subjectivity Word Sense Disambiguation on Contextual Opinion Analysis</i>	
Cem Akkaya, Janyce Wiebe, Alexander Conrad and Rada Mihalcea	87
<i>Effects of Meaning-Preserving Corrections on Language Learning</i>	
Dana Angluin and Leonor Becerra-Bonache	97
<i>Assessing Benefit from Feature Feedback in Active Learning for Text Classification</i>	
Shilpa Arora and Eric Nyberg	106
<i>ULISSE: an Unsupervised Algorithm for Detecting Reliable Dependency Parses</i>	
Felice Dell'Orletta, Giulia Venturi and Simonetta Montemagni	115
<i>Language Models as Representations for Weakly Supervised NLP Tasks</i>	
Fei Huang, Alexander Yates, Arun Ahuja and Doug Downey	125

<i>Automatic Keyphrase Extraction by Bridging Vocabulary Gap</i>	
Zhiyuan Liu, Xinxiong Chen, Yabin Zheng and Maosong Sun	135
<i>Using Second-order Vectors in a Knowledge-based Method for Acronym Disambiguation</i>	
Bridget T. McInnes, Ted Pedersen, Ying Liu, Serguei V. Pakhomov and Genevieve B. Melton	145
<i>Using the Mutual k-Nearest Neighbor Graphs for Semi-supervised Classification on Natural Language Data</i>	
Kohei Ozaki, Masashi Shimbo, Mamoru Komachi and Yuji Matsumoto	154
<i>Automatically Building Training Examples for Entity Extraction</i>	
Marco Pennacchiotti and Patrick Pantel	163
<i>Probabilistic Word Alignment under the L_0-norm</i>	
Thomas Schoenemann	172
<i>Authorship Attribution with Latent Dirichlet Allocation</i>	
Yanir Seroussi, Ingrid Zukerman and Fabian Bohnert	181
<i>Evaluating a Semantic Network Automatically Constructed from Lexical Co-occurrence on a Word Sense Disambiguation Task</i>	
Sean Szumlanski and Fernando Gomez	190
<i>Filling the Gap: Semi-Supervised Learning for Opinion Detection Across Domains</i>	
Ning Yu and Sandra Kübler	200
<i>A Normalized-Cut Alignment Model for Mapping Hierarchical Semantic Structures onto Spoken Documents</i>	
Xiaodan Zhu	210
<i>(Invited talk) Bayesian Tools for Natural Language Learning</i>	
Yee Whye Teh	219
<i>Composing Simple Image Descriptions using Web-scale N-grams</i>	
Siming Li, Girish Kulkarni, Tamara L. Berg, Alexander C. Berg and Yejin Choi	220
<i>Adapting Text instead of the Model: An Open Domain Approach</i>	
Gourab Kundu and Dan Roth	229
<i>Learning with Lookahead: Can History-Based Models Rival Globally Optimized Models?</i>	
Yoshimasa Tsuruoka, Yusuke Miyao and Jun'ichi Kazama	238
<i>Learning Discriminative Projections for Text Similarity Measures</i>	
Wen-tau Yih, Kristina Toutanova, John C. Platt and Christopher Meek	247