

**2012 Conference of the North
American Chapter of the Association
for Computational Linguistics:
Human Language Technologies**

(NAACL HLT 2012)

**Montreal, Quebec, Canada
3-8 June 2012**

ISBN: 978-1-62276-172-2

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© (2012) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2012)

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

Multiple Narrative Disentanglement: Unraveling <i>Infinite Jest</i>	1
<i>Byron Wallace</i>	
Acoustic-Prosodic Entrainment and Social Behavior	11
<i>Rivka Levitan, Agustin Gravano, Laura Willson, Stefan Benus, Julia Hirschberg, Ani Nenkova</i>	
Identifying High-Level Organizational Elements in Argumentative Discourse	20
<i>Nitin Madhani, Michael Heilman, Joel Tetreault, Martin Chodorow</i>	
Fast Inference in Phrase Extraction Models with Belief Propagation	29
<i>David Burkett, Dan Klein</i>	
Continuous Space Translation Models with Neural Networks	39
<i>Hai-Son Le, Alexandre Allauzen, Francois Yvon</i>	
Machine Translation of Arabic Dialects	49
<i>Rabih Zbib, Erika Malchiodi, Jacob Devlin, David Stallard, Spyros Matsoukas, Richard Schwartz, John Makhoul, Omar F. Zaidan, Chris Callison-Burch</i>	
Entity Clustering Across Languages	60
<i>Spence Green, Nicholas Andrews, Matthew R. Gormley, Mark Dredze, Christopher D. Manning</i>	
Multi-Event Extraction Guided by Global Constraints	70
<i>Roi Reichart, Regina Barzilay</i>	
Reference Scope Identification in Citing Sentences	80
<i>Amjad Abu Jbara, Dragomir Radev</i>	
Intrinsic and Extrinsic Evaluation of an Automatic User Disengagement Detector for an Uncertainty-Adaptive Spoken Dialogue System	91
<i>Kate Forbes-Riley, Diane Litman, Heather Friedberg, Joanna Drummond</i>	
Exploring Content Features for Automated Speech Scoring	103
<i>Shasha Xie, Keelan Evanini, Klaus Zechner</i>	
Hello, Who is Calling?: Can Words Reveal the Social Nature of Conversations?	112
<i>Anthony Stark, Izhak Shafran, Jeffrey Kaye</i>	
Minimum-Risk Training of Approximate CRF-Based NLP Systems	120
<i>Veselin Stoyanov, Jason Eisner</i>	
Unsupervised Learning on an Approximate Corpus	131
<i>Jason Smith, Jason Eisner</i>	
Structured Perceptron with Inexact Search	142
<i>Liang Huang, Suphan Fayong, Yang Guo</i>	
Segmentation Similarity and Agreement	152
<i>Chris Fournier, Diana Inkpen</i>	
HyTER: Meaning-Equivalent Semantics for Translation Evaluation	162
<i>Markus Dreyer, Daniel Marcu</i>	
Apples to Oranges: Evaluating Image Annotations from Natural Language Processing Systems	172
<i>Rebecca Mason, Eugene Charniak</i>	
Re-examining Machine Translation Metrics for Paraphrase Identification	182
<i>Nitin Madhani, Joel Tetreault, Martin Chodorow</i>	
A Dependency Treebank of Classical Chinese Poems	191
<i>John Lee, Yin Hei Kong</i>	
Towards Effective Tutorial Feedback for Explanation Questions: A Dataset and Baselines	200
<i>Myroslava O. Dzikovska, Rodney D. Nielsen, Chris Brew</i>	
Topical Segmentation: a Study of Human Performance and a New Measure of Quality	211
<i>Anna Kazantseva, Stan Szpakowicz</i>	
Structured Ramp Loss Minimization for Machine Translation	221
<i>Kevin Gimpel, Noah A. Smith</i>	
Implicitly Intersecting Weighted Automata using Dual Decomposition	232
<i>Michael J. Paul, Jason Eisner</i>	
Transliteration Mining Using Large Training and Test Sets	243
<i>Ali El-Kahki, Kareem Darwish, Mohamed Abdul-Wahab, Ahmed Taei</i>	
Optimized Online Rank Learning for Machine Translation	253
<i>Taro Watanabe</i>	
Every Sensible Extended Top-Down Tree Transducer is a Multi Bottom-Up Tree Transducer	263
<i>Andreas Maletti</i>	
NOMIT: Automatic Titling by Nominalizing	274
<i>Cedric Lopez, Violaine Prince, Mathieu Roche</i>	

Correcting Comma Errors in Learner Essays, and Restoring Commas in Newswire Text	284
<i>Ross Israel, Joel Tetreault, Martin Chodorow</i>	
The Challenges of Parsing Chinese with Combinatory Categorical Grammar	295
<i>Daniel Tse, James R. Curran</i>	
Using Supertags and Encoded Annotation Principles for Improved Dependency to Phrase Structure Conversion.....	305
<i>Seth Kulick, Ann Bies, Justin Mott</i>	
Getting More from Morphology in Multilingual Dependency Parsing	315
<i>Matt Hohensee, Emily M. Bender</i>	
Stylometric Analysis of Scientific Articles	327
<i>Shane Bergsma, Matt Post, David Yarowsky</i>	
Using Paraphrases for Improving First Story Detection in News and Twitter.....	338
<i>Sasa Petrovic, Miles Osborne, Victor Lavrenko</i>	
Insertion and Deletion Models for Statistical Machine Translation.....	347
<i>Matthias Huck, Hermann Ney</i>	
TransAhead: A Computer-Assisted Translation and Writing Tool	352
<i>Chung-Chi Huang, Ping-Che Yang, Keh-Jiann Chen, Jason S. Chang</i>	
Correction Detection and Error Type Selection as an ESL Educational Aid.....	357
<i>Ben Swanson, Elif Yamangil</i>	
Getting More from Segmentation Evaluation	362
<i>Martin Scaiano, Diana Inkpen</i>	
G2P Conversion of Proper Names Using Word Origin Information	367
<i>Sonjia Waxmonsky, Sravana Reddy</i>	
Evaluating a Morphological Analyser of Inuktitut	372
<i>Jeremy Nicholson, Trevor Cohn, Timothy Baldwin</i>	
Intra-Speaker Topic Modeling for Improved Multi-Party Meeting Summarization with Integrated Random Walk	377
<i>Yun-Nung Chen, Florian Metze</i>	
Towards Using EEG to Improve ASR Accuracy	382
<i>Yun-Nung Chen, Kai-Min Chang, Jack Mostow</i>	
A Comparative Investigation of Morphological Language Modeling for the Languages of the European Union.....	386
<i>Thomas Muller, Hinrich Schutze, Helmut Schmid</i>	
Leveraging Supplemental Representations for Sequential Transduction.....	396
<i>Aditya Bhargava, Grzegorz Kondrak</i>	
A Hierarchical Dirichlet Process Model for Joint Part-of-Speech and Morphology Induction	407
<i>Kairit Sirts, Tanel Alumae</i>	
Encouraging Consistent Translation Choices	417
<i>Ferhan Ture, Douglas W. Oard, Philip Resnik</i>	
Batch Tuning Strategies for Statistical Machine Translation	427
<i>Colin Cherry, George Foster</i>	
Real-Time Incremental Speech-to-Speech Translation of Dialogs	437
<i>Srinivas Bangalore, Vivek K. R. Sridhar, Prakash Kolan, Ladan Golipour, Aura Jimenez</i>	
Parsing Time: Learning to Interpret Time Expressions.....	446
<i>Gabor Angeli, Christopher D. Manning, Daniel Jurafsky</i>	
Fine-Grained Focus for Pinpointing Positive Implicit Meaning from Negated Statements	456
<i>Eduardo Blanco, Dan Moldovan</i>	
Taxonomy Induction Using Hierarchical Random Graphs	466
<i>Trevor Fountain, Mirella Lapata</i>	
Cross-Lingual Word Clusters for Direct Transfer of Linguistic Structure	477
<i>Oscar Tackstrom, Ryan McDonald, Jakob Uszkoreit</i>	
Training Dependency Parser Using Light Feedback	488
<i>Avihai Mejer, Koby Crammer</i>	
Vine Pruning for Efficient Multi-Pass Dependency Parsing.....	498
<i>Alexander M. Rush, Slav Petrov</i>	
Active Learning for Coreference Resolution.....	508
<i>Florian Laws, Florian Heimerl, Hinrich Schutze</i>	
Space Efficiencies in Discourse Modeling via Conditional Random Sampling.....	513
<i>Brian Kjersten, Benjamin Van Durme</i>	
Predicting Overt Display of Power in Written Dialogs.....	518
<i>Vinodkumar Prabhakaran, Owen Rambow, Mona Diab</i>	

Co-reference via Pointing and Haptics in Multi-Modal Dialogues	523
<i>Lin Chen, Barbara Di Eugenio</i>	
Trait-Based Hypothesis Selection For Machine Translation	528
<i>Jacob Devlin, Spyros Matsoukas</i>	
Improved Reordering for Shallow-<i>n</i> Grammar based Hierarchical Phrase-based Translation.....	533
<i>Baskaran Sankaran, Anoop Sarkar</i>	
Automatic Parallel Fragment Extraction from Noisy Data	538
<i>Jason Riesa, Daniel Marcu</i>	
Tuning as Linear Regression	543
<i>Marzieh Bazrafshan, Tagyoung Chung, Daniel Gildea</i>	
Ranking-Based Readability Assessment for Early Primary Children’s Literature	548
<i>Yi Ma, Eric Fosler-Lussier, Robert Lofthus</i>	
How Text Segmentation Algorithms Gain from Topic Models.....	553
<i>Martin Riedl, Chris Biemann</i>	
Identifying Comparable Corpora Using LDA.....	558
<i>Judita Preiss</i>	
Behavioral Factors in Interactive Training of Text Classifiers.....	563
<i>Burr Settles, Xiaojin Zhu</i>	
Better Evaluation for Grammatical Error Correction	568
<i>Daniel Dahlmeier, Hwee Tou Ng</i>	
Are You Sure? Confidence in Prediction of Dependency Tree Edges	573
<i>Avihai Mejer, Koby Crammer</i>	
Concavity and Initialization for Unsupervised Dependency Parsing	577
<i>Kevin Gimpel, Noah A. Smith</i>	
Multimodal Grammar Implementation.....	582
<i>Katya Alahverdzhieva, Dan Flickinger, Alex Lascarides</i>	
Portable Features for Classifying Emotional Text.....	587
<i>Saif Mohammad</i>	
Stance Classification using Dialogic Properties of Persuasion.....	592
<i>Marilyn A. Walker, Pranav Anand, Robert Abbott, Ricky Grant</i>	
Context-Enhanced Citation Sentiment Detection	597
<i>Awais Athar, Simone Teufel</i>	
Predicting Responses to Microblog Posts	602
<i>Yoav Artzi, Patrick Pantel, Michael Gamon</i>	
The Intelius Nickname Collection: Quantitative Analyses from Billions of Public Records	607
<i>Vitor R. Carvalho, Yigit Kiran, Andrew Borthwick</i>	
A Comparison of Models of Word Meaning in Context.....	611
<i>Georgiana Dinu, Stefan Thater, Sören Laue</i>	
Measuring Word Relatedness Using Heterogeneous Vector Space Models.....	616
<i>Wen-Tau Yih, Vahed Qazvinian</i>	
Expectations of Word Sense in Parallel Corpora.....	621
<i>Xuchen Yao, Benjamin Van Durme, Chris Callison-Burch</i>	
Why Not Grab a Free Lunch? Mining Large Corpora for Parallel Sentences to Improve Translation Modeling.....	626
<i>Ferhan Ture, Jimmy Lin</i>	
Summarization of Historical Articles Using Temporal Event Clustering.....	631
<i>James Gung, Jugal Kalita</i>	
Comparing HMMs and Bayesian Networks for Surface Realisation.....	636
<i>Nina Dethlefs, Heriberto Cuayahuitl</i>	
On The Feasibility of Open Domain Referring Expression Generation Using Large Scale Folksonomies.....	641
<i>Fabian Pacheco, Pablo Ariel Duboue, Martin Ariel Dominguez</i>	
Structured Event Retrieval over Microblog Archives.....	646
<i>Donald Metzler, Congxing Cai, Eduard Hovy</i>	
Learning from Bullying Traces in Social Media	656
<i>Jun-Ming Xu, Kwang-Sung Jun, Xiaojin Zhu, Amy Bellmore</i>	
Grammatical Structures for Word-Level Sentiment Detection	667
<i>Asad B. Sayeed, Jordan Boyd-Graber, Bryan Rusk, Amy Weinberg</i>	
Graph-Based Lexicon Expansion with Sparsity-Inducing Penalties	677
<i>Dipanjan Das, Noah A. Smith</i>	
Unified Expectation Maximization.....	688
<i>Rajhans Samdani, Ming-Wei Chang, Dan Roth</i>	

Low-Dimensional Discriminative Reranking	699
<i>Jagadeesh Jagarlamudi, Hal Daume III</i>	
Autonomous Self-Assessment of Autocorrections: Exploring Text Message Dialogues	710
<i>Tyler Baldwin, Joyce Y. Chai</i>	
Translation-Based Projection for Multilingual Coreference Resolution.....	720
<i>Altaf Rahman, Vincent Ng</i>	
Exploring Semi-Supervised Coreference Resolution of Medical Concepts using Semantic and Temporal Features	731
<i>Preethi Raghavan, Eric Fosler-Lussier, Albert M. Lai</i>	
Mind the Gap: Learning to Choose Gaps for Question Generation.....	742
<i>Lee Becker, Sumit Basu, Lucy Vanderwende</i>	
Unsupervised Concept-to-Text Generation with Hypergraphs	752
<i>Ioannis Konstas, Mirella Lapata</i>	
Detecting Visual Text	762
<i>Jesse Dodge, Amit Goyal, Xufeng Han, Alyssa Mensch, Margaret Mitchell, Karl Stratos, Kota Yamaguchi, Yejin Choi, Hal Daume III, Alexander C. Berg, Tamara L. Berg</i>	
Unsupervised Translation Sense Clustering.....	773
<i>Mohit Bansal, John DeNero, Dekang Lin</i>	
Shared Components Topic Models.....	783
<i>Matthew R. Gormley, Mark Dredze, Benjamin Van Durme, Jason Eisner</i>	
Textual Predictors of Bill Survival in Congressional Committees.....	793
<i>Tae Yano, Noah A. Smith, John D. Wilkerson</i>	
Author Index	