

14th Conference of the European Chapter of the Association for Computational Linguistics 2014

(EACL 2014)

**Gothenburg, Sweden
26-30 April 2014**

ISBN: 978-1-63266-396-2

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© (2014) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2014)

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>Improving Word Alignment Using Linguistic Code Switching Data</i> Fei Huang and Alexander Yates	1
<i>Undirected Machine Translation with Discriminative Reinforcement Learning</i> Andrea Gesmundo and James Henderson	10
<i>Minimum Translation Modeling with Recurrent Neural Networks</i> Yuening Hu, Michael Auli, Qin Gao and Jianfeng Gao	20
<i>Maximizing Component Quality in Bilingual Word-Aligned Segmentations</i> Spyros Martzoukos, Christof Monz and Christophe Costa Florencio	30
<i>A Joint Model for Quotation Attribution and Coreference Resolution</i> Mariana S. C. Almeida, Miguel B. Almeida and André F. T. Martins	39
<i>A Hierarchical Bayesian Model for Unsupervised Induction of Script Knowledge</i> Lea Frermann, Ivan Titov and Manfred Pinkal	49
<i>Inducing Example-based Semantic Frames from a Massive Amount of Verb Uses</i> Daisuke Kawahara, Daniel Peterson, Octavian Popescu and Martha Palmer	58
<i>Automated Verb Sense Labelling Based on Linked Lexical Resources</i> Kostadin Cholakov, Judith Eckle-Kohler and Iryna Gurevych	68
<i>Multi-Granular Aspect Aggregation in Aspect-Based Sentiment Analysis</i> John Pavlopoulos and Ion Androutsopoulos	78
<i>Simple, Robust and (almost) Unsupervised Generation of Polarity Lexicons for Multiple Languages</i> Iñaki San Vicente, Rodrigo Agerri and German Rigau	88
<i>Mapping Dialectal Variation by Querying Social Media</i> Gabriel Doyle	98
<i>Modeling the Use of Graffiti Style Features to Signal Social Relations within a Multi-Domain Learning Paradigm</i> Mario Piergallini, Seza Dogruoz, Phani Gadde, David Adamson and Carolyn Rose	107
<i>Modelling the Lexicon in Unsupervised Part of Speech Induction</i> Gregory Dubbin and Phil Blunsom	116
<i>Generalizing a Strongly Lexicalized Parser using Unlabeled Data</i> Tejaswini Deoskar, Christos Christodoulopoulos, Alexandra Birch and Mark Steedman	126
<i>Special Techniques for Constituent Parsing of Morphologically Rich Languages</i> Zsolt Szántó and Richárd Farkas	135
<i>Leveraging Verb-Argument Structures to Infer Semantic Relations</i> Eduardo Blanco and Dan Moldovan	145
<i>Structured and Unstructured Cache Models for SMT Domain Adaptation</i> Annie Louis and Bonnie Webber	155

<i>Regularized Structured Perceptron: A Case Study on Chinese Word Segmentation, POS Tagging and Parsing</i>	
Kaixu Zhang, Jinsong Su and Changle Zhou	164
<i>About Inferences in a Crowdsourced Lexical-Semantic Network</i>	
Mathieu Lafourcade, Manel Zarrouk and Alain Joubert	174
<i>Incremental Query Generation</i>	
Laura Perez-Beltrachini, Claire Gardent and Enrico Franconi	183
<i>PARADIGM: Paraphrase Diagnostics through Grammar Matching</i>	
Jonathan Weese, Juri Ganitkevitch and Chris Callison-Burch	192
<i>Translation memory retrieval methods</i>	
Michael Bloodgood and Benjamin Strauss	202
<i>Frame Semantic Tree Kernels for Social Network Extraction from Text</i>	
Apoorv Agarwal, Sriramkumar Balasubramanian, Anup Kotalwar, Jiehan Zheng and Owen Ram- bow	211
<i>Statistical Script Learning with Multi-Argument Events</i>	
Karl Pichotta and Raymond Mooney	220
<i>Improving Distributional Semantic Vectors through Context Selection and Normalisation</i>	
Tamara Polajnar and Stephen Clark	230
<i>Source-side Preordering for Translation using Logistic Regression and Depth-first Branch-and-Bound Search</i>	
Laura Jehl, Adrià de Gispert, Mark Hopkins and Bill Byrne	239
<i>Incremental Bayesian Learning of Semantic Categories</i>	
Lea Frermann and Mirella Lapata	249
<i>Word Ordering with Phrase-Based Grammars</i>	
Adrià de Gispert, Marcus Tomalin and Bill Byrne	259
<i>Iterative Constrained Clustering for Subjectivity Word Sense Disambiguation</i>	
Cem Akkaya, Janyce Wiebe and Rada Mihalcea	269
<i>Identifying fake Amazon reviews as learning from crowds</i>	
Tommaso Fornaciari and Massimo Poesio	279
<i>Assessing the relative reading level of sentence pairs for text simplification</i>	
Sowmya Vajjala and Detmar Meurers	288
<i>Subcategorisation Acquisition from Raw Text for a Free Word-Order Language</i>	
Will Roberts, Markus Egg and Valia Kordoni	298
<i>Classifying Temporal Relations with Simple Features</i>	
Paramita Mirza and Sara Tonelli	308
<i>Using idiolects and sociolects to improve word prediction</i>	
Wessel Stoop and Antal van den Bosch	318
<i>Dynamic Topic Adaptation for Phrase-based MT</i>	
Eva Hasler, Phil Blunsom, Philipp Koehn and Barry Haddow	328

<i>Deterministic Parsing using PCFGs</i>	
Mark-Jan Nederhof and Martin McCaffery	338
<i>Empirically-motivated Generalizations of CCG Semantic Parsing Learning Algorithms</i>	
Jesse Glass and Alexander Yates	348
<i>Correcting Grammatical Verb Errors</i>	
Alla Rozovskaya, Dan Roth and Vivek Srikumar	358
<i>Fast Statistical Parsing with Parallel Multiple Context-Free Grammars</i>	
Krasimir Angelov and Peter Ljunglöf	368
<i>Sentiment Propagation via Implicature Constraints</i>	
Lingjia Deng and Janyce Wiebe	377
<i>Acquisition of Noncontiguous Class Attributes from Web Search Queries</i>	
Marius Pasca	386
<i>Learning from Post-Editing: Online Model Adaptation for Statistical Machine Translation</i>	
Michael Denkowski, Chris Dyer and Alon Lavie	395
<i>Predicting and Characterising User Impact on Twitter</i>	
Vasileios Lampos, Nikolaos Aletras, Daniel Preoțiuc-Pietro and Trevor Cohn	405
<i>A Knowledge-based Representation for Cross-Language Document Retrieval and Categorization</i>	
Marc Franco-Salvador, Paolo Rosso and Roberto Navigli	414
<i>Dependency Tree Abstraction for Long-Distance Reordering in Statistical Machine Translation</i>	
Chenchen Ding and Yuki Arase	424
<i>Improving the Lexical Function Composition Model with Pathwise Optimized Elastic-Net Regression</i>	
Jiming Li, Marco Baroni and Georgiana Dinu	434
<i>Is Machine Translation Getting Better over Time?</i>	
Yvette Graham, Timothy Baldwin, Alistair Moffat and Justin Zobel	443
<i>Learning Dictionaries for Named Entity Recognition using Minimal Supervision</i>	
Arvind Neelakantan and Michael Collins	452
<i>Improving Vector Space Word Representations Using Multilingual Correlation</i>	
Manaal Faruqui and Chris Dyer	462
<i>Using Distributional Similarity of Multi-way Translations to Predict Multiword Expression Compositionality</i>	
Bahar Salehi, Paul Cook and Timothy Baldwin	472
<i>Word Embeddings through Hellinger PCA</i>	
Rémi Lebreton and Ronan Collobert	482
<i>A Latent Variable Model for Discourse-aware Concept and Entity Disambiguation</i>	
Angela Fahrni and Michael Strube	491
<i>Topical PageRank: A Model of Scientific Expertise for Bibliographic Search</i>	
James Jardine and Simone Teufel	501

<i>Distributional Lexical Entailment by Topic Coherence</i>	
Laura Rimell	511
<i>Information Structure Prediction for Visual-world Referring Expressions</i>	
Micha Elsner, Hannah Rohde and Alasdair Clarke	520
<i>Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality</i>	
Jey Han Lau, David Newman and Timothy Baldwin	530
<i>What Substitutes Tell Us - Analysis of an "All-Words" Lexical Substitution Corpus</i>	
Gerhard Kremer, Katrin Erk, Sebastian Padó and Stefan Thater	540
<i>Weighted Krippendorff's alpha is a more reliable metrics for multi-coders ordinal annotations: experimental studies on emotion, opinion and coreference annotation</i>	
Jean-Yves Antoine, Jeanne Villaneau and Anaïs Lefeuvre	550
<i>Discriminating Rhetorical Analogies in Social Media</i>	
Christoph Lofi, Christian Nieke and Nigel Collier	560
<i>Semi-supervised learning of morphological paradigms and lexicons</i>	
Mans Hulden, Markus Forsberg and Malin Ahlberg	569
<i>How to Produce Unseen Teddy Bears: Improved Morphological Processing of Compounds in SMT</i>	
Fabienne Cap, Alexander Fraser, Marion Weller and Aoife Cahill	579
<i>Type-Supervised Domain Adaptation for Joint Segmentation and POS-Tagging</i>	
Meishan Zhang, Yue Zhang, Wanxiang Che and Ting Liu	588
<i>Applying the semantics of negation to SMT through n-best list re-ranking</i>	
Federico Fancellu and Bonnie Webber	598
<i>Bilingual Sentiment Consistency for Statistical Machine Translation</i>	
Boxing Chen and Xiaodan Zhu	607
<i>Augmenting Translation Models with Simulated Acoustic Confusions for Improved Spoken Language Translation</i>	
Yulia Tsvetkov, Florian Metze and Chris Dyer	616
<i>A Generative Model for User Simulation in a Spatial Navigation Domain</i>	
Aciel Eshky, Ben Allison, Subramanian Ramamoorthy and Mark Steedman	626
<i>Verbose, Laconic or Just Right: A Simple Computational Model of Content Appropriateness under Length Constraints</i>	
Annie Louis and Ani Nenkova	636
<i>Discovering Implicit Discourse Relations Through Brown Cluster Pair Representation and Coreference Patterns</i>	
Attapol Rutherford and Nianwen Xue	645
<i>"I Object!" Modeling Latent Pragmatic Effects in Courtroom Dialogues</i>	
Dan Goldwasser and Hal Daumé III	655
<i>Encoding Semantic Resources in Syntactic Structures for Passage Reranking</i>	
Kateryna Tymoshenko, Alessandro Moschitti and Aliaksei Severyn	664

<i>Automatic Food Categorization from Large Unlabeled Corpora and Its Impact on Relation Extraction</i> Michael Wiegand, Benjamin Roth and Dietrich Klakow	673
<i>Redundancy Detection in ESL Writings</i> Huichao Xue and Rebecca Hwa	683
<i>Fast Recursive Multi-class Classification of Pairs of Text Entities for Biomedical Event Extraction</i> Xiao Liu, Antoine Bordes and Yves Grandvalet	692
<i>Cluster-based Prediction of User Ratings for Stylistic Surface Realisation</i> Nina Dethlefs, Heriberto Cuayáhuatl, Helen Hastie, Verena Rieser and Oliver Lemon	702
<i>Improving the Estimation of Word Importance for News Multi-Document Summarization</i> Kai Hong and Ani Nenkova	712
<i>Hybrid text simplification using synchronous dependency grammars with hand-written and automatically harvested rules</i> Advaith Siddharthan and Angrosh Mandya	722
<i>A Summariser based on Human Memory Limitations and Lexical Competition</i> Yimai Fang and Simone Teufel	732
<i>Learning part-of-speech taggers with inter-annotator agreement loss</i> Barbara Plank, Dirk Hovy and Anders Søgaard	742