

52nd Annual Meeting of the Association for Computational Linguistics

(ACL 2014)

**Baltimore, Maryland, USA
22-27 June 2014**

**Volume 1 of 2
Part A**

ISBN: 978-1-63439-191-7

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>Learning Ensembles of Structured Prediction Rules</i> Corinna Cortes, Vitaly Kuznetsov and Mehryar Mohri	1
<i>Representation Learning for Text-level Discourse Parsing</i> Yangfeng Ji and Jacob Eisenstein	13
<i>Text-level Discourse Dependency Parsing</i> Sujian Li, Liang Wang, Ziqiang Cao and Wenjie Li	25
<i>Discovering Latent Structure in Task-Oriented Dialogues</i> Ke Zhai and Jason D Williams	36
<i>Learning Structured Perceptrons for Coreference Resolution with Latent Antecedents and Non-local Features</i> Anders Björkelund and Jonas Kuhn	47
<i>Multilingual Models for Compositional Distributed Semantics</i> Karl Moritz Hermann and Phil Blunsom	58
<i>Simple Negation Scope Resolution through Deep Parsing: A Semantic Solution to a Semantic Problem</i> Woodley Packard, Emily M. Bender, Jonathon Read, Stephan Oepen and Rebecca Dridan	69
<i>Logical Inference on Dependency-based Compositional Semantics</i> Ran Tian, Yusuke Miyao and Takuya Matsuzaki	79
<i>A practical and linguistically-motivated approach to compositional distributional semantics</i> Denis Paperno, Nghia The Pham and Marco Baroni	90
<i>Lattice Desegmentation for Statistical Machine Translation</i> Mohammad Salameh, Colin Cherry and Grzegorz Kondrak	100
<i>Bilingually-constrained Phrase Embeddings for Machine Translation</i> Jiajun Zhang, Shujie Liu, Mu Li, Ming Zhou and Chengqing Zong	111
<i>Learning New Semi-Supervised Deep Auto-encoder Features for Statistical Machine Translation</i> Shixiang Lu, Zhenbiao Chen and Bo Xu	122
<i>Learning Topic Representation for SMT with Neural Networks</i> Lei Cui, Dongdong Zhang, Shujie Liu, Qiming Chen, Mu Li, Ming Zhou and Muyun Yang ...	133
<i>Tagging The Web: Building A Robust Web Tagger with Neural Network</i> Ji Ma, Yue Zhang and Jingbo Zhu	144
<i>Unsupervised Solution Post Identification from Discussion Forums</i> Deepak P and Karthik Visweswariah	155
<i>Weakly Supervised User Profile Extraction from Twitter</i> Jiwei Li, Alan Ritter and Eduard Hovy	165
<i>The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter</i> Chenhao Tan, Lillian Lee and Bo Pang	175

<i>Inferring User Political Preferences from Streaming Communications</i>	
Svitlana Volkova, Glen Coppersmith and Benjamin Van Durme	186
<i>Steps to Excellence: Simple Inference with Refined Scoring of Dependency Trees</i>	
Yuan Zhang, Tao Lei, Regina Barzilay, Tommi Jaakkola and Amir Globerson	197
<i>Sparser, Better, Faster GPU Parsing</i>	
David Hall, Taylor Berg-Kirkpatrick and Dan Klein	208
<i>Shift-Reduce CCG Parsing with a Dependency Model</i>	
Wenduan Xu, Stephen Clark and Yue Zhang	218
<i>Less Grammar, More Features</i>	
David Hall, Greg Durrett and Dan Klein	228
<i>Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors</i>	
Marco Baroni, Georgiana Dinu and Germán Kruszewski	238
<i>Metaphor Detection with Cross-Lingual Model Transfer</i>	
Yulia Tsvetkov, Leonid Boytsov, Anatole Gershman, Eric Nyberg and Chris Dyer	248
<i>Learning Word Sense Distributions, Detecting Unattested Senses and Identifying Novel Senses Using Topic Models</i>	
Jey Han Lau, Paul Cook, Diana McCarthy, Spandana Gella and Timothy Baldwin	259
<i>Learning to Automatically Solve Algebra Word Problems</i>	
Nate Kushman, Luke Zettlemoyer, Regina Barzilay and Yoav Artzi	271
<i>Modelling function words improves unsupervised word segmentation</i>	
Mark Johnson, Anne Christophe, Emmanuel Dupoux and Katherine Demuth	282
<i>Max-Margin Tensor Neural Network for Chinese Word Segmentation</i>	
Wenzhe Pei, Tao Ge and Baobao Chang	293
<i>An Empirical Study on the Effect of Negation Words on Sentiment</i>	
Xiaodan Zhu, Hongyu Guo, Saif Mohammad and Svetlana Kiritchenko	304
<i>Extracting Opinion Targets and Opinion Words from Online Reviews with Graph Co-ranking</i>	
Kang Liu, Liheng Xu and Jun Zhao	314
<i>Context-aware Learning for Sentence-level Sentiment Analysis with Posterior Regularization</i>	
Bishan Yang and Claire Cardie	325
<i>Product Feature Mining: Semantic Clues versus Syntactic Constituents</i>	
Liheng Xu, Kang Liu, Siwei Lai and Jun Zhao	336
<i>Aspect Extraction with Automated Prior Knowledge Learning</i>	
Zhiyuan Chen, Arjun Mukherjee and Bing Liu	347
<i>Anchors Regularized: Adding Robustness and Extensibility to Scalable Topic-Modeling Algorithms</i>	
Thang Nguyen, Yuening Hu and Jordan Boyd-Graber	359
<i>A Bayesian Mixed Effects Model of Literary Character</i>	
David Bamman, Ted Underwood and Noah A. Smith	370

<i>Collective Tweet Wikification based on Semi-supervised Graph Regularization</i> Hongzhao Huang, Yunbo Cao, Xiaojiang Huang, Heng Ji and Chin-Yew Lin	380
<i>Zero-shot Entity Extraction from Web Pages</i> Panupong Pasupat and Percy Liang	391
<i>Incremental Joint Extraction of Entity Mentions and Relations</i> Qi Li and Heng Ji	402
<i>That's Not What I Meant! Using Parsers to Avoid Structural Ambiguities in Generated Text</i> Manjuan Duan and Michael White	413
<i>Surface Realisation from Knowledge-Bases</i> Bikash Gyawali and Claire Gardent	424
<i>Hybrid Simplification using Deep Semantics and Machine Translation</i> Shashi Narayan and Claire Gardent	435
<i>Grammatical Relations in Chinese: GB-Ground Extraction and Data-Driven Parsing</i> Weiwei Sun, Yantao Du, Xin Kou, Shuoyang Ding and Xiaojun Wan	446
<i>Ambiguity-aware Ensemble Training for Semi-supervised Dependency Parsing</i> Zhenghua Li, Min Zhang and Wenliang Chen	457
<i>A Robust Approach to Aligning Heterogeneous Lexical Resources</i> Mohammad Taher Pilehvar and Roberto Navigli	468
<i>Predicting the relevance of distributional semantic similarity with contextual information</i> Philippe Muller, Cécile Fabre and Clémentine Adam	479
<i>Interpretable Semantic Vectors from a Joint Model of Brain- and Text- Based Meaning</i> Alona Fyshe, Partha P. Talukdar, Brian Murphy and Tom M. Mitchell	489
<i>Single-Agent vs. Multi-Agent Techniques for Concurrent Reinforcement Learning of Negotiation Dialogue Policies</i> Kallirroi Georgila, Claire Nelson and David Traum	500
<i>A Linear-Time Bottom-Up Discourse Parser with Constraints and Post-Editing</i> Vanessa Wei Feng and Graeme Hirst	511
<i>Negation Focus Identification with Contextual Discourse Information</i> Bowe Zou, Guodong Zhou and Qiaoming Zhu	522
<i>New Word Detection for Sentiment Analysis</i> Minlie Huang, Borui Ye, Yichen Wang, Haiqiang Chen, Junjun Cheng and Xiaoyan Zhu	531
<i>ReNew: A Semi-Supervised Framework for Generating Domain-Specific Lexicons and Sentiment Analysis</i> Zhe Zhang and Munindar P. Singh	542
<i>A Decision-Theoretic Approach to Natural Language Generation</i> Nathan McKinley and Soumya Ray	552
<i>Generating Code-switched Text for Lexical Learning</i> Igor Labutov and Hod Lipson	562

<i>Omni-word Feature and Soft Constraint for Chinese Relation Extraction</i>	
Yanping Chen, Qinghua Zheng and Wei Zhang	572
<i>Bilingual Active Learning for Relation Classification via Pseudo Parallel Corpora</i>	
Longhua Qian, Haotian Hui, Ya'nan Hu, Guodong Zhou and Qiaoming Zhu	582
<i>Learning Soft Linear Constraints with Application to Citation Field Extraction</i>	
Sam Anzaroot, Alexandre Passos, David Belanger and Andrew McCallum	593
<i>A Study of Concept-based Weighting Regularization for Medical Records Search</i>	
Yue Wang, Xitong Liu and Hui Fang	603
<i>Learning to Predict Distributions of Words Across Domains</i>	
Danushka Bollegala, David Weir and John Carroll	613
<i>How to make words with vectors: Phrase generation in distributional semantics</i>	
Georgiana Dinu and Marco Baroni	624
<i>Vector space semantics with frequency-driven motifs</i>	
Shashank Srivastava and Eduard Hovy	634
<i>Lexical Inference over Multi-Word Predicates: A Distributional Approach</i>	
Omri Abend, Shay B. Cohen and Mark Steedman	644
<i>A Convolutional Neural Network for Modelling Sentences</i>	
Nal Kalchbrenner, Edward Grefenstette and Phil Blunsom	655
<i>Online Learning in Tensor Space</i>	
Yuan Cao and Sanjeev Khudanpur	666
<i>Graph-based Semi-Supervised Learning of Translation Models from Monolingual Data</i>	
Avneesh Saluja, Hany Hassan, Kristina Toutanova and Chris Quirk	676
<i>Using Discourse Structure Improves Machine Translation Evaluation</i>	
Francisco Guzmán, Shafiq Joty, Lluís Màrquez and Preslav Nakov	687
<i>Learning Continuous Phrase Representations for Translation Modeling</i>	
Jianfeng Gao, Xiaodong He, Wen-tau Yih and Li Deng	699
<i>Adaptive Quality Estimation for Machine Translation</i>	
Marco Turchi, Antonios Anastasopoulos, José G. C. de Souza and Matteo Negri	710
<i>Learning Grounded Meaning Representations with Autoencoders</i>	
Carina Silberer and Mirella Lapata	721
<i>Joint POS Tagging and Transition-based Constituent Parsing in Chinese with Non-local Features</i>	
Zhiguo Wang and Nianwen Xue	733
<i>Strategies for Contiguous Multiword Expression Analysis and Dependency Parsing</i>	
Marie Candito and Matthieu Constant	743
<i>Correcting Preposition Errors in Learner English Using Error Case Frames and Feedback Messages</i>	
Ryo Nagata, Mikko Vilenius and Edward Whittaker	754
<i>Kneser-Ney Smoothing on Expected Counts</i>	
Hui Zhang and David Chiang	765

<i>Robust Entity Clustering via Phylogenetic Inference</i>	
Nicholas Andrews, Jason Eisner and Mark Dredze	775
<i>Linguistic Structured Sparsity in Text Categorization</i>	
Dani Yogatama and Noah A. Smith	786
<i>Perplexity on Reduced Corpora</i>	
Hayato Kobayashi	797
<i>Robust Domain Adaptation for Relation Extraction via Clustering Consistency</i>	
Minh Luan Nguyen, Ivor W. Tsang, Kian Ming A. Chai and Hai Leong Chieu	807
<i>Encoding Relation Requirements for Relation Extraction via Joint Inference</i>	
Liwei Chen, Yansong Feng, Songfang Huang, Yong Qin and Dongyan Zhao	818
<i>Medical Relation Extraction with Manifold Models</i>	
Chang Wang and James Fan	828
<i>Distant Supervision for Relation Extraction with Matrix Completion</i>	
Miao Fan, Deli Zhao, Qiang Zhou, Zhiyuan Liu, Thomas Fang Zheng and Edward Y. Chang ..	839
<i>Enhancing Grammatical Cohesion: Generating Transitional Expressions for SMT</i>	
Mei Tu, Yu Zhou and Chengqing Zong	850
<i>Adaptive HTER Estimation for Document-Specific MT Post-Editing</i>	
Fei Huang, Jian-Ming Xu, Abraham Ittycheriah and Salim Roukos	861
<i>Translation Assistance by Translation of L1 Fragments in an L2 Context</i>	
Maarten van Gompel and Antal van den Bosch	871
<i>Response-based Learning for Grounded Machine Translation</i>	
Stefan Riezler, Patrick Simianer and Carolin Haas	881
<i>Modelling Events through Memory-based, Open-IE Patterns for Abstractive Summarization</i>	
Daniele Pighin, Marco Cornolti, Enrique Alfonseca and Katja Filippova	892
<i>Hierarchical Summarization: Scaling Up Multi-Document Summarization</i>	
Janara Christensen, Stephen Soderland, Gagan Bansal and Mausam	902
<i>Query-Chain Focused Summarization</i>	
Tal Baumel, Raphael Cohen and Michael Elhadad	913
<i>Exploiting Timelines to Enhance Multi-document Summarization</i>	
Jun-Ping Ng, Yan Chen, Min-Yen Kan and Zhoujun Li	923
<i>A chance-corrected measure of inter-annotator agreement for syntax</i>	
Arne Skjærholt	934
<i>Two Is Bigger (and Better) Than One: the Wikipedia Bitaxonomy Project</i>	
Tiziano Flati, Daniele Vannella, Tommaso Pasini and Roberto Navigli	945
<i>Information Extraction over Structured Data: Question Answering with Freebase</i>	
Xuchen Yao and Benjamin Van Durme	956
<i>Knowledge-Based Question Answering as Machine Translation</i>	
Junwei Bao, Nan Duan, Ming Zhou and Tiejun Zhao	967

<i>Discourse Complements Lexical Semantics for Non-factoid Answer Reranking</i> Peter Jansen, Mihai Surdeanu and Peter Clark	977
<i>Toward Future Scenario Generation: Extracting Event Causality Exploiting Semantic Relation, Context, and Association Features</i> Chikara Hashimoto, Kentaro Torisawa, Julien Kloetzer, Motoki Sano, István Varga, Jong-Hoon Oh and Yutaka Kidawara	987
<i>Cross-narrative Temporal Ordering of Medical Events</i> Preethi Raghavan, Eric Fosler-Lussier, Noémie Elhadad and Albert M. Lai	998
<i>Language-Aware Truth Assessment of Fact Candidates</i> Ndapandula Nakashole and Tom M. Mitchell	1009
<i>That's sick dude! : Automatic identification of word sense change across different timescales</i> Sunny Mitra, Ritwik Mitra, Martin Riedl, Chris Biemann, Animesh Mukherjee and Pawan Goyal	1020
<i>A Step-wise Usage-based Method for Inducing Polysemy-aware Verb Classes</i> Daisuke Kawahara, Daniel W. Peterson and Martha Palmer	1030
<i>Structured Learning for Taxonomy Induction with Belief Propagation</i> Mohit Bansal, David Burkett, Gerard de Melo and Dan Klein	1041
<i>A Provably Correct Learning Algorithm for Latent-Variable PCFGs</i> Shay B. Cohen and Michael Collins	1052
<i>Spectral Unsupervised Parsing with Additive Tree Metrics</i> Ankur P. Parikh, Shay B. Cohen and Eric P. Xing	1062
<i>Weak semantic context helps phonetic learning in a model of infant language acquisition</i> Stella Frank, Naomi H. Feldman and Sharon Goldwater	1073
<i>Bootstrapping into Filler-Gap: An Acquisition Story</i> Marten van Schijndel and Micha Elsner	1084
<i>Nonparametric Learning of Phonological Constraints in Optimality Theory</i> Gabriel Doyle, Klinton Bicknell and Roger Levy	1094
<i>Active Learning with Efficient Feature Weighting Methods for Improving Data Quality and Classification Accuracy</i> Justin Martineau, Lu Chen, Doreen Cheng and Amit Sheth	1104
<i>Political Ideology Detection Using Recursive Neural Networks</i> Mohit Iyyer, Peter Enns, Jordan Boyd-Graber and Philip Resnik	1113
<i>A Unified Model for Soft Linguistic Reordering Constraints in Statistical Machine Translation</i> Junhui Li, Yuval Marton, Philip Resnik and Hal Daumé III	1123
<i>Are Two Heads Better than One? Crowdsourced Translation via a Two-Step Collaboration of Non-Professional Translators and Editors</i> Rui Yan, Mingkun Gao, Ellie Pavlick and Chris Callison-Burch	1134

<i>A Generalized Language Model as the Combination of Skipped n-grams and Modified Kneser Ney Smoothing</i>	
Rene Pickhardt, Thomas Gottron, Martin Körner, Paul Georg Wagner, Till Speicher and Steffen Staab	1145
<i>A Semiparametric Gaussian Copula Regression Model for Predicting Financial Risks from Earnings Calls</i>	
William Yang Wang and Zhenhao Hua	1155
<i>Polylingual Tree-Based Topic Models for Translation Domain Adaptation</i>	
Yuening Hu, Ke Zhai, Vladimir Eidelman and Jordan Boyd-Graber	1166
<i>Low-Resource Semantic Role Labeling</i>	
Matthew R. Gormley, Margaret Mitchell, Benjamin Van Durme and Mark Dredze	1177
<i>Joint Syntactic and Semantic Parsing with Combinatory Categorical Grammar</i>	
Jayant Krishnamurthy and Tom M. Mitchell	1188
<i>Learning Semantic Hierarchies via Word Embeddings</i>	
Ruiji Fu, Jiang Guo, Bing Qin, Wanxiang Che, Haifeng Wang and Ting Liu	1199
<i>Probabilistic Soft Logic for Semantic Textual Similarity</i>	
Islam Beltagy, Katrin Erk and Raymond Mooney	1210
<i>Abstractive Summarization of Spoken and Written Conversations Based on Phrasal Queries</i>	
Yashar Mehdad, Giuseppe Carenini and Raymond T. Ng	1220
<i>Comparing Multi-label Classification with Reinforcement Learning for Summarisation of Time-series Data</i>	
Dimitra Gkatzia, Helen Hastie and Oliver Lemon	1231
<i>Approximation Strategies for Multi-Structure Sentence Compression</i>	
Kapil Thadani	1241
<i>Opinion Mining on YouTube</i>	
Aliaksei Severyn, Alessandro Moschitti, Olga Uryupina, Barbara Plank and Katja Filippova .	1252
<i>Automatic Keyphrase Extraction: A Survey of the State of the Art</i>	
Kazi Saidul Hasan and Vincent Ng	1262
<i>Pattern Dictionary of English Prepositions</i>	
Ken Litkowski	1274
<i>Looking at Unbalanced Specialized Comparable Corpora for Bilingual Lexicon Extraction</i>	
Emmanuel Morin and Amir Hazem	1284
<i>Validating and Extending Semantic Knowledge Bases using Video Games with a Purpose</i>	
Daniele Vannella, David Jurgens, Daniele Scarfini, Domenico Toscani and Roberto Navigli .	1294
<i>Shallow Analysis Based Assessment of Syntactic Complexity for Automated Speech Scoring</i>	
Suma Bhat, Huichao Xue and Su-Youn Yoon	1305
<i>Can You Repeat That? Using Word Repetition to Improve Spoken Term Detection</i>	
Jonathan Wintrose and Sanjeev Khudanpur	1316

<i>Character-Level Chinese Dependency Parsing</i>	
Meishan Zhang, Yue Zhang, Wanxiang Che and Ting Liu	1326
<i>Unsupervised Dependency Parsing with Transferring Distribution via Parallel Guidance and Entropy Regularization</i>	
Xuezhe Ma and Fei Xia	1337
<i>Unsupervised Morphology-Based Vocabulary Expansion</i>	
Mohammad Sadegh Rasooli, Thomas Lippincott, Nizar Habash and Owen Rambow	1349
<i>Toward Better Chinese Word Segmentation for SMT via Bilingual Constraints</i>	
Xiaodong Zeng, Lidia S. Chao, Derek F. Wong, Isabel Trancoso and Liang Tian	1360
<i>Fast and Robust Neural Network Joint Models for Statistical Machine Translation</i>	
Jacob Devlin, Rabih Zbib, Zhongqiang Huang, Thomas Lamar, Richard Schwartz and John Makhoul	1370
<i>Low-Rank Tensors for Scoring Dependency Structures</i>	
Tao Lei, Yu Xin, Yuan Zhang, Regina Barzilay and Tommi Jaakkola	1381
<i>CoSimRank: A Flexible & Efficient Graph-Theoretic Similarity Measure</i>	
Sascha Rothe and Hinrich Schütze	1392
<i>Is this a wampimuk? Cross-modal mapping between distributional semantics and the visual world</i>	
Angeliki Lazaridou, Elia Bruni and Marco Baroni	1403
<i>Semantic Parsing via Paraphrasing</i>	
Jonathan Berant and Percy Liang	1415
<i>A Discriminative Graph-Based Parser for the Abstract Meaning Representation</i>	
Jeffrey Flanigan, Sam Thomson, Jaime Carbonell, Chris Dyer and Noah A. Smith	1426
<i>Context-dependent Semantic Parsing for Time Expressions</i>	
Kenton Lee, Yoav Artzi, Jesse Dodge and Luke Zettlemoyer	1437
<i>Semantic Frame Identification with Distributed Word Representations</i>	
Karl Moritz Hermann, Dipanjan Das, Jason Weston and Kuzman Ganchev	1448
<i>A Sense-Based Translation Model for Statistical Machine Translation</i>	
Deyi Xiong and Min Zhang	1459
<i>Recurrent Neural Networks for Word Alignment Model</i>	
Akihiro Tamura, Taro Watanabe and Eiichiro Sumita	1470
<i>A Constrained Viterbi Relaxation for Bidirectional Word Alignment</i>	
Yin-Wen Chang, Alexander M. Rush, John DeNero and Michael Collins	1481
<i>A Recursive Recurrent Neural Network for Statistical Machine Translation</i>	
Shujie Liu, Nan Yang, Mu Li and Ming Zhou	1491
<i>Predicting Instructor’s Intervention in MOOC forums</i>	
Snigdha Chaturvedi, Dan Goldwasser and Hal Daumé III	1501
<i>A Joint Graph Model for Pinyin-to-Chinese Conversion with Typo Correction</i>	
Zhongye Jia and Hai Zhao	1512

<i>Smart Selection</i>	
Patrick Pantel, Michael Gamon and Ariel Fuxman	1524
<i>Modeling Prompt Adherence in Student Essays</i>	
Isaac Persing and Vincent Ng	1534
<i>ConnotationWordNet: Learning Connotation over the Word+Sense Network</i>	
Jun Seok Kang, Song Feng, Leman Akoglu and Yejin Choi	1544
<i>Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification</i>	
Duyu Tang, Furu Wei, Nan Yang, Ming Zhou, Ting Liu and Bing Qin	1555
<i>Towards a General Rule for Identifying Deceptive Opinion Spam</i>	
Jiwei Li, Myle Ott, Claire Cardie and Eduard Hovy	1566

52nd Annual Meeting of the Association for Computational Linguistics

(ACL 2014)

**Baltimore, Maryland, USA
22-27 June 2014**

Volume 2 of 2

ISBN: 978-1-63439-191-7

Table of Contents

<i>Exploring the Relative Role of Bottom-up and Top-down Information in Phoneme Learning</i> Abdellah Fourtassi, Thomas Schatz, Balakrishnan Varadarajan and Emmanuel Dupoux	1
<i>Biases in Predicting the Human Language Model</i> Alex B. Fine, Austin F. Frank, T. Florian Jaeger and Benjamin Van Durme	7
<i>Probabilistic Labeling for Efficient Referential Grounding based on Collaborative Discourse</i> Changsong Liu, Lanbo She, Rui Fang and Joyce Y. Chai	13
<i>A Composite Kernel Approach for Dialog Topic Tracking with Structured Domain Knowledge from Wikipedia</i> Seokhwan Kim, Rafael E. Banchs and Haizhou Li	19
<i>An Extension of BLANC to System Mentions</i> Xiaoqiang Luo, Sameer Pradhan, Marta Recasens and Eduard Hovy	24
<i>Scoring Coreference Partitions of Predicted Mentions: A Reference Implementation</i> Sameer Pradhan, Xiaoqiang Luo, Marta Recasens, Eduard Hovy, Vincent Ng and Michael Strube 30	
<i>Measuring Sentiment Annotation Complexity of Text</i> Aditya Joshi, Abhijit Mishra, Nivvedan Senthamilselvan and Pushpak Bhattacharyya	36
<i>Improving Citation Polarity Classification with Product Reviews</i> Charles Jochim and Hinrich Schütze	42
<i>Adaptive Recursive Neural Network for Target-dependent Twitter Sentiment Classification</i> Li Dong, Furu Wei, Chuanqi Tan, Duyu Tang, Ming Zhou and Ke Xu	49
<i>Sprinkling Topics for Weakly Supervised Text Classification</i> Swapnil Hingmire and Sutanu Chakraborti	55
<i>A Feature-Enriched Tree Kernel for Relation Extraction</i> Le Sun and Xianpei Han	61
<i>Employing Word Representations and Regularization for Domain Adaptation of Relation Extraction</i> Thien Huu Nguyen and Ralph Grishman	68
<i>Graph Ranking for Collective Named Entity Disambiguation</i> Ayman Alhelbawy and Robert Gaizauskas	75
<i>Descending-Path Convolution Kernel for Syntactic Structures</i> Chen Lin, Timothy Miller, Alvin Kho, Steven Bethard, Dmitriy Dligach, Sameer Pradhan and Guergana Savova	81
<i>Entities' Sentiment Relevance</i> Zvi Ben-Ami, Ronen Feldman and Binyamin Rosenfeld	87
<i>Automatic Detection of Multilingual Dictionaries on the Web</i> Gintare Grigonyte and Timothy Baldwin	93
<i>Automatic Detection of Cognates Using Orthographic Alignment</i> Alina Maria Ciobanu and Liviu P. Dinu	99

<i>Automatically constructing Wordnet Synsets</i>	
Khang Nhut Lam, Feras Al Tarouti and Jugal Kalita	106
<i>Constructing a Turkish-English Parallel TreeBank</i>	
Olcay Taner Yıldız, Ercan Solak, Onur Görgün and Razieh Ehsani	112
<i>Improved Typesetting Models for Historical OCR</i>	
Taylor Berg-Kirkpatrick and Dan Klein	118
<i>Robust Logistic Regression using Shift Parameters</i>	
Julie Tibshirani and Christopher D. Manning	124
<i>Faster Phrase-Based Decoding by Refining Feature State</i>	
Kenneth Heafield, Michael Kayser and Christopher D. Manning	130
<i>Decoder Integration and Expected BLEU Training for Recurrent Neural Network Language Models</i>	
Michael Auli and Jianfeng Gao	136
<i>On the Elements of an Accurate Tree-to-String Machine Translation System</i>	
Graham Neubig and Kevin Duh	143
<i>Simple extensions and POS Tags for a reparameterised IBM Model 2</i>	
Douwe Gelling and Trevor Cohn	150
<i>Dependency-based Pre-ordering for Chinese-English Machine Translation</i>	
Jingsheng Cai, Masao Utiyama, Eiichiro Sumita and Yujie Zhang	155
<i>Generalized Character-Level Spelling Error Correction</i>	
Noura Farra, Nadi Tomeh, Alla Rozovskaya and Nizar Habash	161
<i>Improved Iterative Correction for Distant Spelling Errors</i>	
Sergey Gubanov, Irina Galinskaya and Alexey Baytin	168
<i>Predicting Grammaticality on an Ordinal Scale</i>	
Michael Heilman, Aoife Cahill, Nitin Madnani, Melissa Lopez, Matthew Mulholland and Joel Tetreault	174
<i>I'm a Belieber: Social Roles via Self-identification and Conceptual Attributes</i>	
Charley Beller, Rebecca Knowles, Craig Harman, Shane Bergsma, Margaret Mitchell and Benjamin Van Durme	181
<i>Automatically Detecting Corresponding Edit-Turn-Pairs in Wikipedia</i>	
Johannes Daxenberger and Iryna Gurevych	187
<i>Two Knives Cut Better Than One: Chinese Word Segmentation with Dual Decomposition</i>	
Mengqiu Wang, Rob Voigt and Christopher D. Manning	193
<i>Effective Document-Level Features for Chinese Patent Word Segmentation</i>	
Si Li and Nianwen Xue	199
<i>Word Segmentation of Informal Arabic with Domain Adaptation</i>	
Will Monroe, Spence Green and Christopher D. Manning	206
<i>Resolving Lexical Ambiguity in Tensor Regression Models of Meaning</i>	
Dimitri Kartsaklis, Nal Kalchbrenner and Mehrnoosh Sadrzadeh	212

<i>A Novel Content Enriching Model for Microblog Using News Corpus</i> Yunlun Yang, Zhihong Deng and Hongliang Yu	218
<i>Learning Bilingual Word Representations by Marginalizing Alignments</i> Tomáš Kočiský, Karl Moritz Hermann and Phil Blunsom	224
<i>Detecting Retries of Voice Search Queries</i> Rivka Levitan and David Elson	230
<i>Sliding Alignment Windows for Real-Time Crowd Captioning</i> Mohammad Kazemi, Rahman Lavaee, Iftekhar Naim and Daniel Gildea	236
<i>Detection of Topic and its Extrinsic Evaluation Through Multi-Document Summarization</i> Yoshimi Suzuki and Fumiyo Fukumoto	241
<i>Content Importance Models for Scoring Writing From Sources</i> Beata Beigman Klebanov, Nitin Madnani, Jill Burstein and Swapna Somasundaran	247
<i>Chinese Morphological Analysis with Character-level POS Tagging</i> Mo Shen, Hongxiao Liu, Daisuke Kawahara and Sadao Kurohashi	253
<i>Part-of-Speech Tagging using Conditional Random Fields: Exploiting Sub-Label Dependencies for Improved Accuracy</i> Miikka Silfverberg, Teemu Ruokolainen, Krister Lindén and Mikko Kurimo	259
<i>POS induction with distributional and morphological information using a distance-dependent Chinese restaurant process</i> Kairit Sirts, Jacob Eisenstein, Micha Elsner and Sharon Goldwater	265
<i>Improving the Recognizability of Syntactic Relations Using Contextualized Examples</i> Aditi Muralidharan and Marti A. Hearst	272
<i>How to Speak a Language without Knowing It</i> Xing Shi, Kevin Knight and Heng Ji	278
<i>Assessing the Discourse Factors that Influence the Quality of Machine Translation</i> Junyi Jessy Li, Marine Carpuat and Ani Nenkova	283
<i>Automatic Detection of Machine Translated Text and Translation Quality Estimation</i> Roei Aharoni, Moshe Koppel and Yoav Goldberg	289
<i>Improving sparse word similarity models with asymmetric measures</i> Jean Mark Gawron	296
<i>Dependency-Based Word Embeddings</i> Omer Levy and Yoav Goldberg	302
<i>Vector spaces for historical linguistics: Using distributional semantics to study syntactic productivity in diachrony</i> Florent Perex	309
<i>Single Document Summarization based on Nested Tree Structure</i> Yuta Kikuchi, Tsutomu Hirao, Hiroya Takamura, Manabu Okumura and Masaaki Nagata	315
<i>Linguistic Considerations in Automatic Question Generation</i> Karen Mazidi and Rodney D. Nielsen	321

<i>Polynomial Time Joint Structural Inference for Sentence Compression</i> Xian Qian and Yang Liu	327
<i>A Bayesian Method to Incorporate Background Knowledge during Automatic Text Summarization</i> Annie Louis	333
<i>Predicting Power Relations between Participants in Written Dialog from a Single Thread</i> Vinodkumar Prabhakaran and Owen Rambow	339
<i>Tri-Training for Authorship Attribution with Limited Training Data</i> Tieyun Qian, Bing Liu, Li Chen and Zhiyong Peng	345
<i>Automation and Evaluation of the Keyword Method for Second Language Learning</i> Gözde Özbal, Daniele Pighin and Carlo Strapparava	352
<i>Citation Resolution: A method for evaluating context-based citation recommendation systems</i> Daniel Duma and Ewan Klein	358
<i>Hippocratic Abbreviation Expansion</i> Brian Roark and Richard Sproat	364
<i>Unsupervised Feature Learning for Visual Sign Language Identification</i> Binyam Gebrekidan Gebre, Onno Crasborn, Peter Wittenburg, Sebastian Drude and Tom Heskes 370	
<i>Experiments with crowdsourced re-annotation of a POS tagging data set</i> Dirk Hovy, Barbara Plank and Anders Søgaard	377
<i>Building Sentiment Lexicons for All Major Languages</i> Yanqing Chen and Steven Skiena	383
<i>Difficult Cases: From Data to Learning, and Back</i> Beata Beigman Klebanov and Eyal Beigman	390
<i>The VerbCorner Project: Findings from Phase 1 of crowd-sourcing a semantic decomposition of verbs</i> Joshua K. Hartshorne, Claire Bonial and Martha Palmer	397
<i>A Corpus of Sentence-level Revisions in Academic Writing: A Step towards Understanding Statement Strength in Communication</i> Chenhao Tan and Lillian Lee	403
<i>Determiner-Established Deixis to Communicative Artifacts in Pedagogical Text</i> Shomir Wilson and Jon Oberlander	409
<i>Modeling Factuality Judgments in Social Media Text</i> Sandeep Soni, Tanushree Mitra, Eric Gilbert and Jacob Eisenstein	415
<i>A Topic Model for Building Fine-grained Domain-specific Emotion Lexicon</i> Min Yang, Dingju Zhu and Kam-Pui Chow	421
<i>Depeche Mood: a Lexicon for Emotion Analysis from Crowd Annotated News</i> Jacopo Staiano and Marco Guerini	427
<i>Improving Twitter Sentiment Analysis with Topic-Based Mixture Modeling and Semi-Supervised Training</i> Bing Xiang and Liang Zhou	434

<i>Cross-cultural Deception Detection</i>	
Verónica Pérez-Rosas and Rada Mihalcea	440
<i>Particle Filter Rejuvenation and Latent Dirichlet Allocation</i>	
Chandler May, Alex Clemmer and Benjamin Van Durme	446
<i>Comparing Automatic Evaluation Measures for Image Description</i>	
Desmond Elliott and Frank Keller	452
<i>Learning a Lexical Simplifier Using Wikipedia</i>	
Colby Horn, Cathryn Manduca and David Kauchak	458
<i>Cheap and easy entity evaluation</i>	
Ben Hachey, Joel Nothman and Will Radford	464
<i>Identifying Real-Life Complex Task Names with Task-Intrinsic Entities from Microblogs</i>	
Ting-Xuan Wang, Kun-Yu Tsai and Wen-Hsiang Lu	470
<i>Mutual Disambiguation for Entity Linking</i>	
Eric Charton, Marie-Jean Meurs, Ludovic Jean-Louis and Michel Gagnon	476
<i>How Well can We Learn Interpretable Entity Types from Text?</i>	
Dirk Hovy	482
<i>Learning Translational and Knowledge-based Similarities from Relevance Rankings for Cross-Language Retrieval</i>	
Shigehiko Schamoni, Felix Hieber, Artem Sokolov and Stefan Riezler	488
<i>Two-Stage Hashing for Fast Document Retrieval</i>	
Hao Li, Wei Liu and Heng Ji	495
<i>An Annotation Framework for Dense Event Ordering</i>	
Taylor Cassidy, Bill McDowell, Nathanael Chambers and Steven Bethard	501
<i>Linguistically debatable or just plain wrong?</i>	
Barbara Plank, Dirk Hovy and Anders Søgaard	507
<i>Humans Require Context to Infer Ironic Intent (so Computers Probably do, too)</i>	
Byron C. Wallace, Do Kook Choe, Laura Kertz and Eugene Charniak	512
<i>Automatic prediction of aspectual class of verbs in context</i>	
Annemarie Friedrich and Alexis Palmer	517
<i>Combining Word Patterns and Discourse Markers for Paradigmatic Relation Classification</i>	
Michael Roth and Sabine Schulte im Walde	524
<i>Applying a Naive Bayes Similarity Measure to Word Sense Disambiguation</i>	
Tong Wang and Graeme Hirst	531
<i>Fast Easy Unsupervised Domain Adaptation with Marginalized Structured Dropout</i>	
Yi Yang and Jacob Eisenstein	538
<i>Improving Lexical Embeddings with Semantic Knowledge</i>	
Mo Yu and Mark Dredze	545

<i>Optimizing Segmentation Strategies for Simultaneous Speech Translation</i>	
Yusuke Oda, Graham Neubig, Sakriani Sakti, Tomoki Toda and Satoshi Nakamura	551
<i>A joint inference of deep case analysis and zero subject generation for Japanese-to-English statistical machine translation</i>	
Taku Kudo, Hiroshi Ichikawa and Hideto Kazawa	557
<i>A Hybrid Approach to Skeleton-based Translation</i>	
Tong Xiao, Jingbo Zhu and Chunliang Zhang	563
<i>Effective Selection of Translation Model Training Data</i>	
Le Liu, Yu Hong, Hao Liu, Xing Wang and Jianmin Yao	569
<i>Refinements to Interactive Translation Prediction Based on Search Graphs</i>	
Philipp Koehn, Chara Tsoukala and Herve Saint-Amand	574
<i>Cross-lingual Model Transfer Using Feature Representation Projection</i>	
Mikhail Kozhevnikov and Ivan Titov	579
<i>Cross-language and Cross-encyclopedia Article Linking Using Mixed-language Topic Model and Hypernym Translation</i>	
Yu-Chun Wang, Chun-Kai Wu and Richard Tzong-Han Tsai	586
<i>Nonparametric Method for Data-driven Image Captioning</i>	
Rebecca Mason and Eugene Charniak	592
<i>Improved Correction Detection in Revised ESL Sentences</i>	
Huichao Xue and Rebecca Hwa	599
<i>Unsupervised Alignment of Privacy Policies using Hidden Markov Models</i>	
Rohan Ramanath, Fei Liu, Norman Sadeh and Noah A. Smith	605
<i>Enriching Cold Start Personalized Language Model Using Social Network Information</i>	
Yu-Yang Huang, Rui Yan, Tsung-Ting Kuo and Shou-De Lin	611
<i>Automatic Labelling of Topic Models Learned from Twitter by Summarisation</i>	
Amparo Elizabeth Cano Basave, Yulan He and Ruifeng Xu	618
<i>Stochastic Contextual Edit Distance and Probabilistic FSTs</i>	
Ryan Cotterell, Nanyun Peng and Jason Eisner	625
<i>Labelling Topics using Unsupervised Graph-based Methods</i>	
Nikolaos Aletras and Mark Stevenson	631
<i>Training a Korean SRL System with Rich Morphological Features</i>	
Young-Bum Kim, Heemoon Chae, Benjamin Snyder and Yu-Seop Kim	637
<i>Semantic Parsing for Single-Relation Question Answering</i>	
Wen-tau Yih, Xiaodong He and Christopher Meek	643
<i>On WordNet Semantic Classes and Dependency Parsing</i>	
Kepa Bengoetxea, Eneko Agirre, Joakim Nivre, Yue Zhang and Koldo Gojenola	649
<i>Enforcing Structural Diversity in Cube-pruned Dependency Parsing</i>	
Hao Zhang and Ryan McDonald	656

<i>The Penn Parsed Corpus of Modern British English: First Parsing Results and Analysis</i>	
Seth Kulick, Anthony Kroch and Beatrice Santorini	662
<i>Parser Evaluation Using Derivation Trees: A Complement to evalb</i>	
Seth Kulick, Ann Bies, Justin Mott, Anthony Kroch, Beatrice Santorini and Mark Liberman ..	668
<i>Learning Polylingual Topic Models from Code-Switched Social Media Documents</i>	
Nanyun Peng, Yiming Wang and Mark Dredze	674
<i>Normalizing tweets with edit scripts and recurrent neural embeddings</i>	
Grzegorz Chrupała	680
<i>Exponential Reservoir Sampling for Streaming Language Models</i>	
Miles Osborne, Ashwin Lall and Benjamin Van Durme	687
<i>A Piece of My Mind: A Sentiment Analysis Approach for Online Dispute Detection</i>	
Lu Wang and Claire Cardie	693
<i>A Simple Bayesian Modelling Approach to Event Extraction from Twitter</i>	
Deyu Zhou, Liangyu Chen and Yulan He	700
<i>Be Appropriate and Funny: Automatic Entity Morph Encoding</i>	
Boliang Zhang, Hongzhao Huang, Xiaoman Pan, Heng Ji, Kevin Knight, Zhen Wen, Yizhou Sun, Jiawei Han and Bulent Yener	706
<i>Applying Grammar Induction to Text Mining</i>	
Andrew Salway and Samia Touileb	712
<i>Semantic Consistency: A Local Subspace Based Method for Distant Supervised Relation Extraction</i>	
Xianpei Han and Le Sun	718
<i>Concreteness and Subjectivity as Dimensions of Lexical Meaning</i>	
Felix Hill and Anna Korhonen	725
<i>Infusion of Labeled Data into Distant Supervision for Relation Extraction</i>	
Maria Pershina, Bonan Min, Wei Xu and Ralph Grishman	732
<i>Recognizing Implied Predicate-Argument Relationships in Textual Inference</i>	
Asher Stern and Ido Dagan	739
<i>Measuring metaphoricity</i>	
Jonathan Dunn	745
<i>Empirical Study of Unsupervised Chinese Word Segmentation Methods for SMT on Large-scale Corpora</i>	
Xiaolin Wang, Masao Utiyama, Andrew Finch and Eiichiro Sumita	752
<i>EM Decipherment for Large Vocabularies</i>	
Malte Nuhn and Hermann Ney	759
<i>XMEANT: Better semantic MT evaluation without reference translations</i>	
Chi-kiu Lo, Meriem Beloucif, Markus Saers and Dekai Wu	765
<i>Sentence Level Dialect Identification for Machine Translation System Selection</i>	
Wael Salloum, Heba Elfardy, Linda Alamir-Salloum, Nizar Habash and Mona Diab	772

<i>RNN-based Derivation Structure Prediction for SMT</i>	
Feifei Zhai, Jiajun Zhang, Yu Zhou and Chengqing Zong	779
<i>Hierarchical MT Training using Max-Violation Perceptron</i>	
Kai Zhao, Liang Huang, Haitao Mi and Abe Ittycheriah	785
<i>Punctuation Processing for Projective Dependency Parsing</i>	
Ji Ma, Yue Zhang and Jingbo Zhu	791
<i>Transforming trees into hedges and parsing with "hedgebank" grammars</i>	
Mahsa Yarmohammadi, Aaron Dunlop and Brian Roark	797
<i>Incremental Predictive Parsing with TurboParser</i>	
Arne Köhn and Wolfgang Menzel	803
<i>Tailoring Continuous Word Representations for Dependency Parsing</i>	
Mohit Bansal, Kevin Gimpel and Karen Livescu	809
<i>Observational Initialization of Type-Supervised Taggers</i>	
Hui Zhang and John DeNero	816
<i>How much do word embeddings encode about syntax?</i>	
Jacob Andreas and Dan Klein	822
<i>Distributed Representations of Geographically Situated Language</i>	
David Bamman, Chris Dyer and Noah A. Smith	828
<i>Improving Multi-Modal Representations Using Image Dispersion: Why Less is Sometimes More</i>	
Douwe Kiela, Felix Hill, Anna Korhonen and Stephen Clark	835
<i>Bilingual Event Extraction: a Case Study on Trigger Type Determination</i>	
Zhu Zhu, Shoushan Li, Guodong Zhou and Rui Xia	842
<i>Understanding Relation Temporality of Entities</i>	
Taesung Lee and Seung-won Hwang	848
<i>Does the Phonology of L1 Show Up in L2 Texts?</i>	
Garrett Nicolai and Grzegorz Kondrak	854
<i>Cross-lingual Opinion Analysis via Negative Transfer Detection</i>	
Lin Gui, Ruifeng Xu, Qin Lu, Jun Xu, Jian Xu, Bin Liu and Xiaolong Wang	860