

46th Annual Meeting of the Association for Computational Linguistics

(ACL 2008)

Human Language Technologies

**Columbus, Ohio, USA
15-20 June 2008**

Volume 1, Part A

ISBN: 978-1-63439-829-9

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© (2008) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2015)

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

Preface: General Chair	xi
Preface: Program Chairs	xii
Organizers	xv
Program Committee	xvii
Conference Program	xxi
<i>Mining Wiki Resources for Multilingual Named Entity Recognition</i>	
Alexander E. Richman and Patrick Schone	1
<i>Distributional Identification of Non-Referential Pronouns</i>	
Shane Bergsma, Dekang Lin and Randy Goebel	10
<i>Weakly-Supervised Acquisition of Open-Domain Classes and Class Attributes from Web Documents and Query Logs</i>	
Marius Paşca and Benjamin Van Durme	19
<i>The Tradeoffs Between Open and Traditional Relation Extraction</i>	
Michele Banko and Oren Etzioni	28
<i>PDT 2.0 Requirements on a Query Language</i>	
Jiří Mirovský	37
<i>Task-oriented Evaluation of Syntactic Parsers and Their Representations</i>	
Yusuke Miyao, Rune Sætre, Kenji Sagae, Takuya Matsuzaki and Jun'ichi Tsujii	46
<i>MAXSIM: A Maximum Similarity Metric for Machine Translation Evaluation</i>	
Yee Seng Chan and Hwee Tou Ng	55
<i>Contradictions and Justifications: Extensions to the Textual Entailment Task</i>	
Ellen M. Voorhees	63
<i>Cohesive Phrase-Based Decoding for Statistical Machine Translation</i>	
Colin Cherry	72
<i>Phrase Table Training for Precision and Recall: What Makes a Good Phrase and a Good Phrase Pair?</i>	
Yonggang Deng, Jia Xu and Yuqing Gao	81
<i>Measure Word Generation for English-Chinese SMT Systems</i>	
Dongdong Zhang, Mu Li, Nan Duan, Chi-Ho Li and Ming Zhou	89
<i>Bayesian Learning of Non-Compositional Phrases with Synchronous Parsing</i>	
Hao Zhang, Chris Quirk, Robert C. Moore and Daniel Gildea	97

<i>Applying a Grammar-Based Language Model to a Simplified Broadcast-News Transcription Task</i> Tobias Kaufmann and Beat Pfister	106
<i>Automatic Editing in a Back-End Speech-to-Text System</i> Maximilian Bisani, Paul Vozila, Olivier Divay and Jeff Adams	114
<i>Grounded Language Modeling for Automatic Speech Recognition of Sports Video</i> Michael Fleischman and Deb Roy	121
<i>Lexicalized Phonotactic Word Segmentation</i> Margaret M. Fleck	130
<i>A Re-examination of Query Expansion Using Lexical Resources</i> Hui Fang	139
<i>Selecting Query Term Alternations for Web Search by Exploiting Query Contexts</i> Guihong Cao, Stephen Robertson and Jian-Yun Nie	148
<i>Searching Questions by Identifying Question Topic and Question Focus</i> Huizhong Duan, Yunbo Cao, Chin-Yew Lin and Yong Yu	156
<i>Trainable Generation of Big-Five Personality Styles through Data-Driven Parameter Estimation</i> François Mairesse and Marilyn Walker	165
<i>Correcting Misuse of Verb Forms</i> John Lee and Stephanie Seneff	174
<i>Hypertagging: Supertagging for Surface Realization with CCG</i> Dominic Espinosa, Michael White and Dennis Mehay	183
<i>Forest-Based Translation</i> Haitao Mi, Liang Huang and Qun Liu	192
<i>A Discriminative Latent Variable Model for Statistical Machine Translation</i> Phil Blunsom, Trevor Cohn and Miles Osborne	200
<i>Efficient Multi-Pass Decoding for Synchronous Context Free Grammars</i> Hao Zhang and Daniel Gildea	209
<i>Regular Tree Grammars as a Formalism for Scope Underspecification</i> Alexander Koller, Michaela Regneri and Stefan Thater	218
<i>Classification of Semantic Relationships between Nominals Using Pattern Clusters</i> Dmitry Davidov and Ari Rappoport	227
<i>Vector-based Models of Semantic Composition</i> Jeff Mitchell and Mirella Lapata	236
<i>Exploiting Feature Hierarchy for Transfer Learning in Named Entity Recognition</i> Andrew Arnold, Ramesh Nallapati and William W. Cohen	245

<i>Refining Event Extraction through Cross-Document Inference</i>	
Heng Ji and Ralph Grishman	254
<i>Learning Document-Level Semantic Properties from Free-Text Annotations</i>	
S.R.K. Branavan, Harr Chen, Jacob Eisenstein and Regina Barzilay	263
<i>Automatic Image Annotation Using Auxiliary Text Information</i>	
Yansong Feng and Mirella Lapata	272
<i>Hedge Classification in Biomedical Texts with a Weakly Supervised Selection of Keywords</i>	
György Szarvas	281
<i>When Specialists and Generalists Work Together: Overcoming Domain Dependence in Sentiment Tagging</i>	
Alina Andreevskaia and Sabine Bergler	290
<i>A Generic Sentence Trimmer with CRFs</i>	
Tadashi Nomoto	299
<i>A Joint Model of Text and Aspect Ratings for Sentiment Summarization</i>	
Ivan Titov and Ryan McDonald	308
<i>Improving Parsing and PP Attachment Performance with Sense Information</i>	
Eneko Agirre, Timothy Baldwin and David Martinez	317
<i>A Logical Basis for the D Combinator and Normal Form in CCG</i>	
Frederick Hoyt and Jason Baldridge	326
<i>Parsing Noun Phrase Structure with CCG</i>	
David Vadas and James R. Curran	335
<i>Sentence Simplification for Semantic Role Labeling</i>	
David Vickrey and Daphne Koller	344
<i>Summarizing Emails with Conversational Cohesion and Subjectivity</i>	
Giuseppe Carenini, Raymond T. Ng and Xiaodong Zhou	353
<i>Ad Hoc Treebank Structures</i>	
Markus Dickinson	362
<i>A Single Generative Model for Joint Morphological Segmentation and Syntactic Parsing</i>	
Yoav Goldberg and Reut Tsarfaty	371
<i>Which Words Are Hard to Recognize? Prosodic, Lexical, and Disfluency Factors that Increase ASR Error Rates</i>	
Sharon Goldwater, Dan Jurafsky and Christopher D. Manning	380
<i>Name Translation in Statistical Machine Translation - Learning When to Transliterate</i>	
Ulf Hermjakob, Kevin Knight and Hal Daumé III	389

<i>Using Adaptor Grammars to Identify Synergies in the Unsupervised Acquisition of Linguistic Structure</i> Mark Johnson	398
<i>Inducing Gazetteers for Named Entity Recognition by Large-Scale Clustering of Dependency Relations</i> Jun'ichi Kazama and Kentaro Torisawa	407
<i>Evaluating Roget's Thesauri</i> Alistair Kennedy and Stan Szpakowicz	416
<i>Unsupervised Translation Induction for Chinese Abbreviations using Monolingual Corpora</i> Zhifei Li and David Yarowsky	425
<i>Which Are the Best Features for Automatic Verb Classification</i> Jianguo Li and Chris Brew	434
<i>Collecting a Why-Question Corpus for Development and Evaluation of an Automatic QA-System</i> Joanna Mrozinski, Edward Whittaker and Sadaoki Furui	443
<i>Solving Relational Similarity Problems Using the Web as a Corpus</i> Preslav Nakov and Marti A. Hearst	452
<i>Combining Speech Retrieval Results with Generalized Additive Models</i> J. Scott Olsson and Douglas W. Oard	461
<i>A Critical Reassessment of Evaluation Baselines for Speech Summarization</i> Gerald Penn and Xiaodan Zhu	470
<i>Intensional Summaries as Cooperative Responses in Dialogue: Automation and Evaluation</i> Joseph Polifroni and Marilyn Walker	479
<i>Word Clustering and Word Selection Based Feature Reduction for MaxEnt Based Hindi NER</i> Sujan Kumar Saha, Pabitra Mitra and Sudeshna Sarkar	488
<i>Combining EM Training and the MDL Principle for an Automatic Verb Classification Incorporating Selectional Preferences</i> Sabine Schulte im Walde, Christian Hying, Christian Scheible and Helmut Schmid	496
<i>Randomized Language Models via Perfect Hash Functions</i> David Talbot and Thorsten Brants	505
<i>Applying Morphology Generation Models to Machine Translation</i> Kristina Toutanova, Hisami Suzuki and Achim Ruopp	514
<i>Multilingual Harvesting of Cross-Cultural Stereotypes</i> Tony Veale, Yanfen Hao and Guofu Li	523
<i>Semi-Supervised Convex Training for Dependency Parsing</i> Qin Iris Wang, Dale Schuurmans and Dekang Lin	532

<i>Chinese-English Backward Transliteration Assisted with Mining Monolingual Web Pages</i>	
Fan Yang, Jun Zhao, Bo Zou, Kang Liu and Feifan Liu	541
<i>Robustness and Generalization of Role Sets: PropBank vs. VerbNet</i>	
Beñat Zapirain, Eneko Agirre and Lluís Màrquez	550
<i>A Tree Sequence Alignment-based Tree-to-Tree Translation Model</i>	
Min Zhang, Hongfei Jiang, Aiti Aw, Haizhou Li, Chew Lim Tan and Sheng Li	559
<i>Automatic Syllabification with Structured SVMs for Letter-to-Phoneme Conversion</i>	
Susan Bartlett, Grzegorz Kondrak and Colin Cherry	568
<i>A New String-to-Dependency Machine Translation Algorithm with a Target Dependency Language Model</i>	
Libin Shen, Jinxi Xu and Ralph Weischedel	577
<i>Forest Reranking: Discriminative Parsing with Non-Local Features</i>	
Liang Huang	586
<i>Simple Semi-supervised Dependency Parsing</i>	
Terry Koo, Xavier Carreras and Michael Collins	595
<i>Optimal k-arization of Synchronous Tree-Adjoining Grammar</i>	
Rebecca Nesson, Giorgio Satta and Stuart M. Shieber	604
<i>Enhancing Performance of Lexicalised Grammars</i>	
Rebecca Dridan, Valia Kordoni and Jeremy Nicholson	613
<i>Assessing Dialog System User Simulation Evaluation Measures Using Human Judges</i>	
Hua Ai and Diane J. Litman	622
<i>Robust Dialog Management with N-Best Hypotheses Using Dialog Examples and Agenda</i>	
Cheongjae Lee, Sangkeun Jung and Gary Geunbae Lee	630
<i>Learning Effective Multimodal Dialogue Strategies from Wizard-of-Oz Data: Bootstrapping and Evaluation</i>	
Verena Rieser and Oliver Lemon	638
<i>Phrase Chunking Using Entropy Guided Transformation Learning</i>	
Ruy Luiz Milidiú, Cícero Nogueira dos Santos and Julio C. Duarte	647
<i>Learning Bigrams from Unigrams</i>	
Xiaojin Zhu, Andrew B. Goldberg, Michael Rabbat and Robert Nowak	656
<i>Semi-Supervised Sequential Labeling and Segmentation Using Giga-Word Scale Unlabeled Data</i>	
Jun Suzuki and Hideki Isozaki	665
<i>Large Scale Acquisition of Paraphrases for Learning Surface Patterns</i>	
Rahul Bhagat and Deepak Ravichandran	674

<i>Contextual Preferences</i>	
Idan Szpektor, Ido Dagan, Roy Bar-Haim and Jacob Goldberger	683
<i>Unsupervised Discovery of Generic Relationships Using Pattern Clusters and its Evaluation by Automatically Generated SAT Analogy Questions</i>	
Dmitry Davidov and Ari Rappoport	692
<i>Improving Search Results Quality by Customizing Summary Lengths</i>	
Michael Kaisser, Marti A. Hearst and John B. Lowe	701
<i>Using Conditional Random Fields to Extract Contexts and Answers of Questions from Online Forums</i>	
Shilin Ding, Gao Cong, Chin-Yew Lin and Xiaoyan Zhu	710
<i>Learning to Rank Answers on Large Online QA Collections</i>	
Mihai Surdeanu, Massimiliano Ciaramita and Hugo Zaragoza	719
<i>Unsupervised Lexicon-Based Resolution of Unknown Words for Full Morphological Analysis</i>	
Meni Adler, Yoav Goldberg, David Gabay and Michael Elhadad	728
<i>Unsupervised Multilingual Learning for Morphological Segmentation</i>	
Benjamin Snyder and Regina Barzilay	737
<i>EM Can Find Pretty Good HMM POS-Taggers (When Given a Good Start)</i>	
Yoav Goldberg, Meni Adler and Michael Elhadad	746
<i>Distributed Word Clustering for Large Scale Class-Based Language Modeling in Machine Translation</i>	
Jakob Uszkoreit and Thorsten Brants	755
<i>Enriching Morphologically Poor Languages for Statistical Machine Translation</i>	
Eleftherios Avramidis and Philipp Koehn	763
<i>Learning Bilingual Lexicons from Monolingual Corpora</i>	
Aria Haghighi, Percy Liang, Taylor Berg-Kirkpatrick and Dan Klein	771
<i>Pivot Approach for Extracting Paraphrase Patterns from Bilingual Corpora</i>	
Shiqi Zhao, Haifeng Wang, Ting Liu and Sheng Li	780
<i>Unsupervised Learning of Narrative Event Chains</i>	
Nathanael Chambers and Dan Jurafsky	789
<i>Semantic Role Labeling Systems for Arabic using Kernel Methods</i>	
Mona Diab, Alessandro Moschitti and Daniele Pighin	798
<i>An Unsupervised Approach to Biography Production Using Wikipedia</i>	
Fadi Biadisy, Julia Hirschberg and Elena Filatova	807
<i>Generating Impact-Based Summaries for Scientific Literature</i>	
Qiaozhu Mei and ChengXiang Zhai	816

<i>Can You Summarize This? Identifying Correlates of Input Difficulty for Multi-Document Summarization</i> Ani Nenkova and Annie Louis	825
<i>You Talking to Me? A Corpus and Algorithm for Conversation Disentanglement</i> Micha Elsner and Eugene Charniak	834
<i>An Entity-Mention Model for Coreference Resolution with Inductive Logic Programming</i> Xiaofeng Yang, Jian Su, Jun Lang, Chew Lim Tan, Ting Liu and Sheng Li	843
<i>Gestural Cohesion for Topic Segmentation</i> Jacob Eisenstein, Regina Barzilay and Randall Davis	852
<i>Multi-Task Active Learning for Linguistic Annotations</i> Roi Reichart, Katrin Tomanek, Udo Hahn and Ari Rappoport	861
<i>Generalized Expectation Criteria for Semi-Supervised Learning of Conditional Random Fields</i> Gideon S. Mann and Andrew McCallum	870
<i>Analyzing the Errors of Unsupervised Learning</i> Percy Liang and Dan Klein	879
<i>Joint Word Segmentation and POS Tagging Using a Single Perceptron</i> Yue Zhang and Stephen Clark	888
<i>A Cascaded Linear Model for Joint Chinese Word Segmentation and Part-of-Speech Tagging</i> Wenbin Jiang, Liang Huang, Qun Liu and Yajuan Lü	897
<i>Joint Processing and Discriminative Training for Letter-to-Phoneme Conversion</i> Sittichai Jiampojarn, Colin Cherry and Grzegorz Kondrak	905
<i>A Probabilistic Model for Fine-Grained Expert Search</i> Shenghua Bao, Huizhong Duan, Qi Zhou, Miao Xiong, Yunbo Cao and Yong Yu	914
<i>Credibility Improves Topical Blog Post Retrieval</i> Wouter Weerkamp and Maarten de Rijke	923
<i>Linguistically Motivated Features for Enhanced Back-of-the-Book Indexing</i> Andras Csomai and Rada Mihalcea	932
<i>Resolving Personal Names in Email Using Context Expansion</i> Tamer Elsayed, Douglas W. Oard and Galileo Namata	941
<i>Integrating Graph-Based and Transition-Based Dependency Parsers</i> Joakim Nivre and Ryan McDonald	950
<i>Efficient, Feature-based, Conditional Random Field Parsing</i> Jenny Rose Finkel, Alex Kleeman and Christopher D. Manning	959
<i>A Deductive Approach to Dependency Parsing</i> Carlos Gómez-Rodríguez, John Carroll and David Weir	968

<i>Evaluating a Crosslinguistic Grammar Resource: A Case Study of Wambaya</i>	
Emily M. Bender	977
<i>Better Alignments = Better Translations?</i>	
Kuzman Ganchev, João V. Graça and Ben Taskar	986
<i>Mining Parenthetical Translations from the Web by Word Alignment</i>	
Dekang Lin, Shaojun Zhao, Benjamin Van Durme and Marius Paşca	994
<i>Soft Syntactic Constraints for Hierarchical Phrased-Based Translation</i>	
Yuval Marton and Philip Resnik	1003
<i>Generalizing Word Lattice Translation</i>	
Christopher Dyer, Smaranda Muresan and Philip Resnik	1012
<i>Combining Multiple Resources to Improve SMT-based Paraphrasing Model</i>	
Shiqi Zhao, Cheng Niu, Ming Zhou, Ting Liu and Sheng Li	1021
<i>Extraction of Entailed Semantic Relations Through Syntax-Based Comma Resolution</i>	
Vivek Srikumar, Roi Reichart, Mark Sammons, Ari Rappoport and Dan Roth	1030
<i>Finding Contradictions in Text</i>	
Marie-Catherine de Marneffe, Anna N. Rafferty and Christopher D. Manning	1039
<i>Semantic Class Learning from the Web with Hyponym Pattern Linkage Graphs</i>	
Zornitsa Kozareva, Ellen Riloff and Eduard Hovy	1048

46th Annual Meeting of the Association for Computational Linguistics

(ACL 2008)

Human Language Technologies

**Columbus, Ohio, USA
15-20 June 2008**

Volume 2

ISBN: 978-1-63439-829-9

Table of Contents

<i>Language Dynamics and Capitalization using Maximum Entropy</i>	
Fernando Batista, Nuno Mamede and Isabel Trancoso	1
<i>Surprising Parser Actions and Reading Difficulty</i>	
Marisa Ferrara Boston, John T. Hale, Reinhold Kliegl and Shravan Vasishth	5
<i>Improving the Performance of the Random Walk Model for Answering Complex Questions</i>	
Yllias Chali and Shafiq Joty	9
<i>Dimensions of Subjectivity in Natural Language</i>	
Wei Chen	13
<i>Extractive Summaries for Educational Science Content</i>	
Sebastian de la Chica, Faisal Ahmad, James H. Martin and Tamara Sumner	17
<i>Dialect Classification for Online Podcasts Fusing Acoustic and Language Based Structural and Semantic Information</i>	
Rahul Chitturi and John Hansen	21
<i>The Complexity of Phrase Alignment Problems</i>	
John DeNero and Dan Klein	25
<i>Novel Semantic Features for Verb Sense Disambiguation</i>	
Dmitriy Dligach and Martha Palmer	29
<i>Icelandic Data Driven Part of Speech Tagging</i>	
Mark Dredze and Joel Wallenberg	33
<i>Beyond Log-Linear Models: Boosted Minimum Error Rate Training for N-best Re-ranking</i>	
Kevin Duh and Katrin Kirchhoff	37
<i>Coreference-inspired Coherence Modeling</i>	
Micha Elsner and Eugene Charniak	41
<i>Enforcing Transitivity in Coreference Resolution</i>	
Jenny Rose Finkel and Christopher D. Manning	45
<i>Simulating the Behaviour of Older versus Younger Users when Interacting with Spoken Dialogue Systems</i>	
Kallirroi Georgila, Maria Wolters and Johanna Moore	49
<i>Active Sample Selection for Named Entity Transliteration</i>	
Dan Goldwasser and Dan Roth	53

<i>Four Techniques for Online Handling of Out-of-Vocabulary Words in Arabic-English Statistical Machine Translation</i>	
Nizar Habash	57
<i>Combined One Sense Disambiguation of Abbreviations</i>	
Yaakov HaCohen-Kerner, Ariel Kass and Ariel Peretz	61
<i>Assessing the Costs of Sampling Methods in Active Learning for Annotation</i>	
Robbie Haertel, Eric Ringger, Kevin Seppi, James Carroll and McClanahan Peter	65
<i>Blog Categorization Exploiting Domain Dictionary and Dynamically Estimated Domains of Unknown Words</i>	
Chikara Hashimoto and Sadao Kurohashi	69
<i>Mixture Model POMDPs for Efficient Handling of Uncertainty in Dialogue Management</i>	
James Henderson and Oliver Lemon	73
<i>Recent Improvements in the CMU Large Scale Chinese-English SMT System</i>	
Almut Silja Hildebrand, Kay Rottmann, Mohamed Noamany, Quin Gao, Sanjika Hewavitharana, Nguyen Bach and Stephan Vogel	77
<i>Machine Translation System Combination using ITG-based Alignments</i>	
Damianos Karakos, Jason Eisner, Sanjeev Khudanpur and Markus Dreyer	81
<i>Dictionary Definitions based Homograph Identification using a Generative Hierarchical Model</i>	
Anagha Kulkarni and Jamie Callan	85
<i>A Novel Feature-based Approach to Chinese Entity Relation Extraction</i>	
Wenjie Li, Peng Zhang, Furu Wei, Yuexian Hou and Qin Lu	89
<i>Using Structural Information for Identifying Similar Chinese Characters</i>	
Chao-Lin Liu and Jen-Hsiang Lin	93
<i>You've Got Answers: Towards Personalized Models for Predicting Success in Community Question Answering</i>	
Yandong Liu and Eugene Agichtein	97
<i>Self-Training for Biomedical Parsing</i>	
David McClosky and Eugene Charniak	101
<i>A Unified Syntactic Model for Parsing Fluent and Disfluent Speech</i>	
Tim Miller and William Schuler	105
<i>The Good, the Bad, and the Unknown: Morphosyllabic Sentiment Tagging of Unseen Words</i>	
Karo Moilanen and Stephen Pulman	109
<i>Kernels on Linguistic Structures for Answer Extraction</i>	
Alessandro Moschitti and Silvia Quarteroni	113

<i>Arabic Morphological Tagging, Diacritization, and Lemmatization Using Lexeme Models and Feature Ranking</i>	
Ryan Roth, Owen Rambow, Nizar Habash, Mona Diab and Cynthia Rudin	117
<i>Using Automatically Transcribed Dialogs to Learn User Models in a Spoken Dialog System</i>	
Umar Syed and Jason Williams	121
<i>Robust Extraction of Named Entity Including Unfamiliar Word</i>	
Masatoshi Tsuchiya, Shinya Hida and Seiichi Nakagawa	125
<i>In-Browser Summarisation: Generating Elaborative Summaries Biased Towards the Reading Context</i>	
Stephen Wan and Cécile Paris	129
<i>Lyric-based Song Sentiment Classification with Sentiment Vector Space Model</i>	
Yunqing Xia, Linlin Wang, Kam-Fai Wong and Mingxing Xu	133
<i>Mining Wikipedia Revision Histories for Improving Sentence Compression</i>	
Elif Yamangil and Rani Nelken	137
<i>Smoothing a Tera-word Language Model</i>	
Deniz Yuret	141
<i>Event Matching Using the Transitive Closure of Dependency Relations</i>	
Daniel M. Bikel and Vittorio Castelli	145
<i>A Linguistically Annotated Reordering Model for BTG-based Statistical Machine Translation</i>	
Deyi Xiong, Min Zhang, Aiti Aw and Haizhou Li	149
<i>Segmentation for English-to-Arabic Statistical Machine Translation</i>	
Ibrahim Badr, Rabih Zbib and James Glass	153
<i>Exploiting N-best Hypotheses for SMT Self-Enhancement</i>	
Boxing Chen, Min Zhang, Aiti Aw and Haizhou Li	157
<i>Partial Matching Strategy for Phrase-based Statistical Machine Translation</i>	
Zhongjun He, Qun Liu and Shouxun Lin	161
<i>Unsupervised Learning of Acoustic Sub-word Units</i>	
Balakrishnan Varadarajan, Sanjeev Khudanpur and Emmanuel Dupoux	165
<i>High Frequency Word Entrainment in Spoken Dialogue</i>	
Ani Nenkova, Agustín Gravano and Julia Hirschberg	169
<i>Distributed Listening: A Parallel Processing Approach to Automatic Speech Recognition</i>	
Yolanda McMillian and Juan Gilbert	173
<i>Learning Semantic Links from a Corpus of Parallel Temporal and Causal Relations</i>	
Steven Bethard and James H. Martin	177

<i>Evolving New Lexical Association Measures Using Genetic Programming</i> Jan Šnajder, Bojana Dalbelo Bašić, Saša Petrović and Ivan Sikirić	181
<i>Semantic Types of Some Generic Relation Arguments: Detection and Evaluation</i> Sophia Katrenko and Pieter Adriaans	185
<i>Mapping between Compositional Semantic Representations and Lexical Semantic Resources: Towards Accurate Deep Semantic Parsing</i> Sergio Roa, Valia Kordoni and Yi Zhang	189
<i>Query-based Sentence Fusion is Better Defined and Leads to More Preferred Results than Generic Sentence Fusion</i> Emiel Krahmer, Erwin Marsi and Paul van Pelt	193
<i>Intrinsic vs. Extrinsic Evaluation Measures for Referring Expression Generation</i> Anja Belz and Albert Gatt	197
<i>Correlation between ROUGE and Human Evaluation of Extractive Meeting Summaries</i> Feifan Liu and Yang Liu	201
<i>FastSum: Fast and Accurate Query-based Multi-document Summarization</i> Frank Schilder and Ravikumar Kondadadi	205
<i>Construct State Modification in the Arabic Treebank</i> Ryan Gabbard and Seth Kulick	209
<i>Unlexicalised Hidden Variable Models of Split Dependency Grammars</i> Gabriele Antonio Musillo and Paola Merlo	213
<i>Computing Confidence Scores for All Sub Parse Trees</i> Feng Lin and Fuliang Weng	217
<i>Adapting a WSJ-Trained Parser to Grammatically Noisy Text</i> Jennifer Foster, Joachim Wagner and Josef van Genabith	221
<i>Enriching Spoken Language Translation with Dialog Acts</i> Vivek Kumar Rangarajan Sridhar, Srinivas Bangalore and Shrikanth Narayanan	225
<i>Speakers' Intention Prediction Using Statistics of Multi-level Features in a Schedule Management Domain</i> Donghyun Kim, Hyunjung Lee, Choong-Nyoung Seon, Harksoo Kim and Jungyun Seo	229
<i>Active Learning with Confidence</i> Mark Dredze and Koby Crammer	233
<i>splitSVM: Fast, Space-Efficient, non-Heuristic, Polynomial Kernel Computation for NLP Applications</i> Yoav Goldberg and Michael Elhadad	237
<i>Extracting a Representation from Text for Semantic Analysis</i> Rodney D. Nielsen, Wayne Ward, James H. Martin and Martha Palmer	241

<i>Efficient Processing of Underspecified Discourse Representations</i>	
Michaela Regneri, Markus Egg and Alexander Koller	245
<i>Choosing Sense Distinctions for WSD: Psycholinguistic Evidence</i>	
Susan Windisch Brown	249
<i>Decompounding query keywords from compounding languages</i>	
Enrique Alfonseca, Slaven Bilac and Stefan Pharies	253
<i>Multi-domain Sentiment Classification</i>	
Shoushan Li and Chengqing Zong	257
<i>Evaluating Word Prediction: Framing Keystroke Savings</i>	
Keith Trnka and Kathleen McCoy	261
<i>Pairwise Document Similarity in Large Collections with MapReduce</i>	
Tamer Elsayed, Jimmy Lin and Douglas Oard	265
<i>Text Segmentation with LDA-Based Fisher Kernel</i>	
Qi Sun, Runxin Li, Dingsheng Luo and Xihong Wu	269