

48th Annual Meeting of the Association for Computational Linguistics 2010

(ACL 2010)

**Uppsala, Sweden
11-16 July 2010**

Volume 1 of 2

ISBN: 978-1-61738-808-8

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© (2010) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2010)

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>Efficient Third-Order Dependency Parsers</i>	
Terry Koo and Michael Collins	1
<i>Dependency Parsing and Projection Based on Word-Pair Classification</i>	
Wenbin Jiang and Qun Liu	12
<i>Bitext Dependency Parsing with Bilingual Subtree Constraints</i>	
Wenliang Chen, Jun'ichi Kazama and Kentaro Torisawa	21
<i>Computing Weakest Readings</i>	
Alexander Koller and Stefan Thater	30
<i>Identifying Generic Noun Phrases</i>	
Nils Reiter and Anette Frank	40
<i>Structural Semantic Relatedness: A Knowledge-Based Method to Named Entity Disambiguation</i>	
Xianpei Han and Jun Zhao	50
<i>Correcting Errors in Speech Recognition with Articulatory Dynamics</i>	
Frank Rudzicz	60
<i>Learning to Adapt to Unknown Users: Referring Expression Generation in Spoken Dialogue Systems</i>	
Srinivasan Janarthanam and Oliver Lemon	69
<i>A Risk Minimization Framework for Extractive Speech Summarization</i>	
Shih-Hsiang Lin and Berlin Chen	79
<i>The Human Language Project: Building a Universal Corpus of the World's Languages</i>	
Steven Abney and Steven Bird	88
<i>Bilingual Lexicon Generation Using Non-Aligned Signatures</i>	
Daphna Shezaf and Ari Rappoport	98
<i>Automatic Evaluation Method for Machine Translation Using Noun-Phrase Chunking</i>	
Hiroshi Echizen-ya and Kenji Araki	108
<i>Open Information Extraction Using Wikipedia</i>	
Fei Wu and Daniel S. Weld	118
<i>SystemT: An Algebraic Approach to Declarative Information Extraction</i>	
Laura Chiticariu, Rajasekar Krishnamurthy, Yunyao Li, Sriram Raghavan, Frederick Reiss and Shivakumar Vaithyanathan	128
<i>Extracting Social Networks from Literary Fiction</i>	
David Elson, Nicholas Dames and Kathleen McKeown	138
<i>Pseudo-Word for Phrase-Based Machine Translation</i>	
Xiangyu Duan, Min Zhang and Haizhou Li	148
<i>Hierarchical Search for Word Alignment</i>	
Jason Riesa and Daniel Marcu	157

<i>“Was It Good? It Was Provocative.” Learning the Meaning of Scalar Adjectives</i>	
Marie-Catherine de Marneffe, Christopher D. Manning and Christopher Potts	167
<i>Importance-Driven Turn-Bidding for Spoken Dialogue Systems</i>	
Ethan Selfridge and Peter Heeman	177
<i>Entity-Based Local Coherence Modelling Using Topological Fields</i>	
Jackie Chi Kit Cheung and Gerald Penn	186
<i>Syntactic and Semantic Factors in Processing Difficulty: An Integrated Measure</i>	
Jeff Mitchell, Mirella Lapata, Vera Demberg and Frank Keller	196
<i>Rebanking CCGbank for Improved NP Interpretation</i>	
Matthew Honnibal, James R. Curran and Johan Bos	207
<i>BabelNet: Building a Very Large Multilingual Semantic Network</i>	
Roberto Navigli and Simone Paolo Ponzetto	216
<i>Fully Unsupervised Core-Adjunct Argument Classification</i>	
Omri Abend and Ari Rappoport	226
<i>Towards Open-Domain Semantic Role Labeling</i>	
Danilo Croce, Cristina Giannone, Paolo Annesi and Roberto Basili	237
<i>A Bayesian Method for Robust Estimation of Distributional Similarities</i>	
Jun’ichi Kazama, Stijn De Saeger, Kow Kuroda, Masaki Murata and Kentaro Torisawa	247
<i>Recommendation in Internet Forums and Blogs</i>	
Jia Wang, Qing Li, Yuanzhu Peter Chen and Zhangxi Lin	257
<i>Learning Phrase-Based Spelling Error Models from Clickthrough Data</i>	
Xu Sun, Jianfeng Gao, Daniel Micol and Chris Quirk	266
<i>Inducing Domain-Specific Semantic Class Taggers from (Almost) Nothing</i>	
Ruihong Huang and Ellen Riloff	275
<i>Learning 5000 Relational Extractors</i>	
Raphael Hoffmann, Congle Zhang and Daniel S. Weld	286
<i>Unsupervised Ontology Induction from Text</i>	
Hoifung Poon and Pedro Domingos	296
<i>Exploring Syntactic Structural Features for Sub-Tree Alignment Using Bilingual Tree Kernels</i>	
Jun Sun, Min Zhang and Chew Lim Tan	306
<i>Discriminative Pruning for Discriminative ITG Alignment</i>	
Shujie Liu, Chi-Ho Li and Ming Zhou	316
<i>Fine-Grained Tree-to-String Translation Rule Extraction</i>	
Xianchao Wu, Takuya Matsuzaki and Jun’ichi Tsujii	325
<i>Accurate Context-Free Parsing with Combinatory Categorical Grammar</i>	
Timothy A. D. Fowler and Gerald Penn	335

<i>Faster Parsing by Supertagger Adaptation</i>	
Jonathan K. Kummerfeld, Jessika Roesner, Tim Dawborn, James Haggerty, James R. Curran and Stephen Clark	345
<i>Using Smaller Constituents Rather Than Sentences in Active Learning for Japanese Dependency Parsing</i>	
Manabu Sassano and Sadao Kurohashi	356
<i>Conditional Random Fields for Word Hyphenation</i>	
Nikolaos Trogkanis and Charles Elkan	366
<i>Enhanced Word Decomposition by Calibrating the Decision Threshold of Probabilistic Models and Using a Model Ensemble</i>	
Sebastian Spiegler and Peter A. Flach	375
<i>Word Representations: A Simple and General Method for Semi-Supervised Learning</i>	
Joseph Turian, Lev-Arie Ratinov and Yoshua Bengio	384
<i>Identifying Text Polarity Using Random Walks</i>	
Ahmed Hassan and Dragomir Radev	395
<i>Sentiment Learning on Product Reviews via Sentiment Ontology Tree</i>	
Wei Wei and Jon Atle Gulla	404
<i>Employing Personal/Impersonal Views in Supervised and Semi-Supervised Sentiment Classification</i>	
Shoushan Li, Chu-Ren Huang, Guodong Zhou and Sophia Yat Mei Lee	414
<i>A Latent Dirichlet Allocation Method for Selectional Preferences</i>	
Alan Ritter, Mausam and Oren Etzioni	424
<i>Latent Variable Models of Selectional Preference</i>	
Diarmuid Ó Séaghdha	435
<i>Improving the Use of Pseudo-Words for Evaluating Selectional Preferences</i>	
Nathanael Chambers and Dan Jurafsky	445
<i>Syntax-to-Morphology Mapping in Factored Phrase-Based Statistical Machine Translation from English to Turkish</i>	
Reyyan Yeniterzi and Kemal Oflazer	454
<i>Hindi-to-Urdu Machine Translation through Transliteration</i>	
Nadir Durrani, Hassan Sajjad, Alexander Fraser and Helmut Schmid	465
<i>Training Phrase Translation Models with Leaving-One-Out</i>	
Joern Wuebker, Arne Mauser and Hermann Ney	475
<i>Efficient Staggered Decoding for Sequence Labeling</i>	
Nobuhiro Kaji, Yasuhiro Fujiwara, Naoki Yoshinaga and Masaru Kitsuregawa	485
<i>Minimized Models and Grammar-Informed Initialization for Supertagging with Highly Ambiguous Lexicons</i>	
Sujith Ravi, Jason Baldridge and Kevin Knight	495
<i>Practical Very Large Scale CRFs</i>	
Thomas Lavergne, Olivier Cappé and François Yvon	504

<i>On the Computational Complexity of Dominance Links in Grammatical Formalisms</i> Sylvain Schmitz	514
<i>Optimal Rank Reduction for Linear Context-Free Rewriting Systems with Fan-Out Two</i> Benoît Sagot and Giorgio Satta	525
<i>The Importance of Rule Restrictions in CCG</i> Marco Kuhlmann, Alexander Koller and Giorgio Satta	534
<i>Automatic Evaluation of Linguistic Quality in Multi-Document Summarization</i> Emily Pitler, Annie Louis and Ani Nenkova	544
<i>Identifying Non-Explicit Citing Sentences for Citation-Based Summarization.</i> Vahed Qazvinian and Dragomir R. Radev	555
<i>Automatic Generation of Story Highlights</i> Kristian Woodsend and Mirella Lapata	565
<i>Sentence and Expression Level Annotation of Opinions in User-Generated Discourse</i> Cigdem Toprak, Niklas Jakob and Iryna Gurevych	575
<i>Generating Focused Topic-Specific Sentiment Lexicons</i> Valentin Jijkoun, Maarten de Rijke and Wouter Weerkamp	585
<i>Evaluating Multilanguage-Comparability of Subjectivity Analysis Systems</i> Jungi Kim, Jin-Ji Li and Jong-Hyeok Lee	595
<i>Error Detection for Statistical Machine Translation Using Linguistic Features</i> Deyi Xiong, Min Zhang and Haizhou Li	604
<i>TrustRank: Inducing Trust in Automatic Translations via Ranking</i> Radu Soricut and Abdessamad Echihabi	612
<i>Bridging SMT and TM with Translation Recommendation</i> Yifan He, Yanjun Ma, Josef van Genabith and Andy Way	622
<i>On Jointly Recognizing and Aligning Bilingual Named Entities</i> Yufeng Chen, Chengqing Zong and Keh-Yih Su	631
<i>Generating Templates of Entity Summaries with an Entity-Aspect Model and Pattern Mining</i> Peng Li, Jing Jiang and Yinglin Wang	640
<i>Comparable Entity Mining from Comparative Questions</i> Shasha Li, Chin-Yew Lin, Young-In Song and Zhoujun Li	650
<i>Towards Robust Multi-Tool Tagging. An OWL/DL-Based Approach</i> Christian Chiarcos	659
<i>Temporal Information Processing of a New Language: Fast Porting with Minimal Resources</i> Francisco Costa and António Branco	671
<i>A Taxonomy, Dataset, and Classifier for Automatic Noun Compound Interpretation</i> Stephen Tratz and Eduard Hovy	678
<i>Models of Metaphor in NLP</i> Ekaterina Shutova	688

<i>A Game-Theoretic Model of Metaphorical Bargaining</i>	
Beata Beigman Klebanov and Eyal Beigman	698
<i>Kernel Based Discourse Relation Recognition with Temporal Ordering Information</i>	
WenTing Wang, Jian Su and Chew Lim Tan	710
<i>Hierarchical Joint Learning: Improving Joint Parsing and Named Entity Recognition with Non-Jointly Labeled Data</i>	
Jenny Rose Finkel and Christopher D. Manning	720
<i>Detecting Errors in Automatically-Parsed Dependency Relations</i>	
Markus Dickinson	729
<i>Boosting-Based System Combination for Machine Translation</i>	
Tong Xiao, Jingbo Zhu, Muhua Zhu and Huizhen Wang	739
<i>Fine-Grained Genre Classification Using Structural Learning Algorithms</i>	
Zhili Wu, Katja Markert and Serge Sharoff	749
<i>Metadata-Aware Measures for Answer Summarization in Community Question Answering</i>	
Mattia Tomasoni and Minlie Huang	760
<i>A Hybrid Rule/Model-Based Finite-State Framework for Normalizing SMS Messages</i>	
Richard Beaufort, Sophie Roekhaut, Louise-Amélie Cougnon and Cédric Fairon	770
<i>Letter-Phoneme Alignment: An Exploration</i>	
Sittichai Jiampojarn and Grzegorz Kondrak	780
<i>Using Document Level Cross-Event Inference to Improve Event Extraction</i>	
Shasha Liao and Ralph Grishman	789
<i>Now, Where Was I? Resumption Strategies for an In-Vehicle Dialogue System</i>	
Jessica Villing	798
<i>Learning to Follow Navigational Directions</i>	
Adam Vogel and Dan Jurafsky	806
<i>A Hybrid Hierarchical Model for Multi-Document Summarization</i>	
Asli Celikyilmaz and Dilek Hakkani-Tur	815
<i>Improving Statistical Machine Translation with Monolingual Collocation</i>	
Zhanyi Liu, Haifeng Wang, Hua Wu and Sheng Li	825
<i>Bilingual Sense Similarity for Statistical Machine Translation</i>	
Boxing Chen, George Foster and Roland Kuhn	834
<i>Untangling the Cross-Lingual Link Structure of Wikipedia</i>	
Gerard de Melo and Gerhard Weikum	844
<i>Bucking the Trend: Large-Scale Cost-Focused Active Learning for Statistical Machine Translation</i>	
Michael Bloodgood and Chris Callison-Burch	854
<i>Creating Robust Supervised Classifiers via Web-Scale N-Gram Data</i>	
Shane Bergsma, Emily Pitler and Dekang Lin	865

<i>Convolution Kernel over Packed Parse Forest</i>	
Min Zhang, Hui Zhang and Haizhou Li	875
<i>Estimating Strictly Piecewise Distributions</i>	
Jeffrey Heinz and James Rogers	886
<i>String Extension Learning</i>	
Jeffrey Heinz	897
<i>Compositional Matrix-Space Models of Language</i>	
Sebastian Rudolph and Eugenie Giesbrecht	907
<i>Cross-Language Document Summarization Based on Machine Translation Quality Prediction</i>	
Xiaojun Wan, Huiying Li and Jianguo Xiao	917
<i>A New Approach to Improving Multilingual Summarization Using a Genetic Algorithm</i>	
Marina Litvak, Mark Last and Menahem Friedman	927
<i>Bayesian Synchronous Tree-Substitution Grammar Induction and Its Application to Sentence Compression</i>	
Elif Yamangil and Stuart M. Shieber	937
<i>Contextualizing Semantic Representations Using Syntactically Enriched Vector Models</i>	
Stefan Thater, Hagen Fürstenau and Manfred Pinkal	948
<i>Bootstrapping Semantic Analyzers from Non-Contradictory Texts</i>	
Ivan Titov and Mikhail Kozhevnikov	958
<i>Open-Domain Semantic Role Labeling by Modeling Word Spans</i>	
Fei Huang and Alexander Yates	968
<i>Learning Script Knowledge with Web Experiments</i>	
Michaela Regneri, Alexander Koller and Manfred Pinkal	979
<i>Starting from Scratch in Semantic Role Labeling</i>	
Michael Connor, Yael Gertner, Cynthia Fisher and Dan Roth	989
<i>Modeling Norms of Turn-Taking in Multi-Party Conversation</i>	
Kornel Laskowski	999
<i>Optimising Information Presentation for Spoken Dialogue Systems</i>	
Verena Rieser, Oliver Lemon and Xingkun Liu	1009
<i>Combining Data and Mathematical Models of Language Change</i>	
Morgan Sonderegger and Partha Niyogi	1019
<i>Finding Cognate Groups Using Phylogenies</i>	
David Hall and Dan Klein	1030
<i>An Exact A* Method for Deciphering Letter-Substitution Ciphers</i>	
Eric Corlett and Gerald Penn	1040
<i>A Statistical Model for Lost Language Decipherment</i>	
Benjamin Snyder, Regina Barzilay and Kevin Knight	1048

<i>Efficient Inference through Cascades of Weighted Tree Transducers</i>	
Jonathan May, Kevin Knight and Heiko Vogler	1058
<i>A Tree Transducer Model for Synchronous Tree-Adjoining Grammars</i>	
Andreas Maletti	1067
<i>Dynamic Programming for Linear-Time Incremental Parsing</i>	
Liang Huang and Kenji Sagae	1077
<i>Hard Constraints for Grammatical Function Labelling</i>	
Wolfgang Seeker, Ines Rehbein, Jonas Kuhn and Josef Van Genabith	1087
<i>Simple, Accurate Parsing with an All-Fragments Grammar</i>	
Mohit Bansal and Dan Klein	1098
<i>Joint Syntactic and Semantic Parsing of Chinese</i>	
Junhui Li, Guodong Zhou and Hwee Tou Ng	1108
<i>Cross-Language Text Classification Using Structural Correspondence Learning</i>	
Peter Prettenhofer and Benno Stein	1118
<i>Cross-Lingual Latent Topic Extraction</i>	
Duo Zhang, Qiaozhu Mei and ChengXiang Zhai	1128
<i>Topic Models for Word Sense Disambiguation and Token-Based Idiom Detection</i>	
Linlin Li, Benjamin Roth and Caroline Sporleder	1138
<i>PCFGs, Topic Models, Adaptor Grammars and Learning Topical Collocations and the Structure of Proper Names</i>	
Mark Johnson	1148
<i>A Cognitive Cost Model of Annotations Based on Eye-Tracking Data</i>	
Katrin Tomanek, Udo Hahn, Steffen Lohmann and Jürgen Ziegler	1158
<i>A Rational Model of Eye Movement Control in Reading</i>	
Klinton Bicknell and Roger Levy	1168
<i>The Influence of Discourse on Syntax: A Psycholinguistic Model of Sentence Processing</i>	
Amit Dubey	1179
<i>Complexity Metrics in an Incremental Right-Corner Parser</i>	
Stephen Wu, Asaf Bachrach, Carlos Cardenas and William Schuler	1189
<i>“Ask Not What Textual Entailment Can Do for You...”</i>	
Mark Sammons, V.G. Vinod Vydiswaran and Dan Roth	1199
<i>Assessing the Role of Discourse References in Entailment Inference</i>	
Shachar Mirkin, Ido Dagan and Sebastian Pado	1209
<i>Global Learning of Focused Entailment Graphs</i>	
Jonathan Berant, Ido Dagan and Jacob Goldberger	1220
<i>Modeling Semantic Relevance for Question-Answer Pairs in Web Social Communities</i>	
Baoxun Wang, Xiaolong Wang, Chengjie Sun, Bingquan Liu and Lin Sun	1230

<i>How Many Words Is a Picture Worth? Automatic Caption Generation for News Images</i> Yansong Feng and Mirella Lapata	1239
<i>Generating Image Descriptions Using Dependency Relational Patterns</i> Ahmet Aker and Robert Gaizauskas	1250
<i>Incorporating Extra-Linguistic Information into Reference Resolution in Collaborative Task Dialogue</i> Ryu Iida, Syumpei Kobayashi and Takenobu Tokunaga	1259
<i>Reading between the Lines: Learning to Map High-Level Instructions to Commands</i> S.R.K. Branavan, Luke Zettlemoyer and Regina Barzilay	1268
<i>Profiting from Mark-Up: Hyper-Text Annotations for Guided Parsing</i> Valentin I. Spitzkovsky, Daniel Jurafsky and Hiyan Alshawhi	1278
<i>Phylogenetic Grammar Induction</i> Taylor Berg-Kirkpatrick and Dan Klein	1288
<i>Improved Unsupervised POS Induction through Prototype Discovery</i> Omri Abend, Roi Reichart and Ari Rappoport	1298
<i>Extraction and Approximation of Numerical Attributes from the Web</i> Dmitry Davidov and Ari Rappoport	1308
<i>Learning Word-Class Lattices for Definition and Hypernym Extraction</i> Roberto Navigli and Paola Velardi	1318
<i>On Learning Subtypes of the Part-Whole Relation: Do Not Mix Your Seeds</i> Ashwin Ittoo and Gosse Bouma	1328
<i>Understanding the Semantic Structure of Noun Phrase Queries</i> Xiao Li	1337
<i>Multilingual Pseudo-Relevance Feedback: Performance Study of Assisting Languages</i> Manoj Kumar Chinnakotla, Karthik Raman and Pushpak Bhattacharyya	1346
<i>Wikipedia as Sense Inventory to Improve Diversity in Web Search Results</i> Celina Santamaría, Julio Gonzalo and Javier Artiles	1357
<i>A Unified Graph Model for Sentence-Based Opinion Retrieval</i> Binyang Li, Lanjun Zhou, Shi Feng and Kam-Fai Wong	1367
<i>Generating Fine-Grained Reviews of Songs from Album Reviews</i> Swati Tata and Barbara Di Eugenio	1376
<i>A Study of Information Retrieval Weighting Schemes for Sentiment Analysis</i> Georgios Paltoglou and Mike Thelwall	1386
<i>Supervised Noun Phrase Coreference Research: The First Fifteen Years</i> Vincent Ng	1396
<i>Unsupervised Event Coreference Resolution with Rich Linguistic Features</i> Cosmin Bejan and Sanda Harabagiu	1412
<i>Coreference Resolution across Corpora: Languages, Coding Schemes, and Preprocessing Information</i> Marta Recasens and Eduard Hovy	1423

<i>Constituency to Dependency Translation with Forests</i>	
Haitao Mi and Qun Liu	1433
<i>Learning to Translate with Source and Target Syntax</i>	
David Chiang	1443
<i>Discriminative Modeling of Extraction Sets for Machine Translation</i>	
John DeNero and Dan Klein	1453
<i>Detecting Experiences from Weblogs</i>	
Keun Chan Park, Yoonjae Jeong and Sung Hyon Myaeng	1464
<i>Experiments in Graph-Based Semi-Supervised Learning Methods for Class-Instance Acquisition</i>	
Partha Pratim Talukdar and Fernando Pereira	1473
<i>Learning Arguments and Supertypes of Semantic Relations Using Recursive Patterns</i>	
Zornitsa Kozareva and Eduard Hovy	1482
<i>A Transition-Based Parser for 2-Planar Dependency Structures</i>	
Carlos Gómez-Rodríguez and Joakim Nivre	1492
<i>Viterbi Training for PCFGs: Hardness Results and Competitiveness of Uniform Initialization</i>	
Shay Cohen and Noah A Smith	1502
<i>A Generalized-Zero-Preserving Method for Compact Encoding of Concept Lattices</i>	
Matthew Skala, Victoria Krakovna, János Kramár and Gerald Penn	1512
<i>Knowledge-Rich Word Sense Disambiguation Rivaling Supervised Systems</i>	
Simone Paolo Ponzetto and Roberto Navigli	1522
<i>All Words Domain Adapted WSD: Finding a Middle Ground between Supervision and Unsupervision</i>	
Mitesh Khapra, Anup Kulkarni, Saurabh Sohoney and Pushpak Bhattacharyya	1532
<i>Combining Orthogonal Monolingual and Multilingual Sources of Evidence for All Words WSD</i>	
Weiwei Guo and Mona Diab	1542
<i>Phrase-Based Statistical Language Generation Using Graphical Models and Active Learning</i>	
Francois Mairesse, Milica Gasic, Filip Jurcicek, Simon Keizer, Blaise Thomson, Kai Yu and Steve Young	1552
<i>Plot Induction and Evolutionary Search for Story Generation</i>	
Neil McIntyre and Mirella Lapata	1562
<i>Automated Planning for Situated Natural Language Generation</i>	
Konstantina Garoufi and Alexander Koller	1573
<i>Beyond NomBank: A Study of Implicit Arguments for Nominal Predicates</i>	
Matthew Gerber and Joyce Chai	1583