ACL 2007

Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics

June 23–30, 2007
Prague, Czech Republic

The Association for Computational Linguistics

It is with great sadness that we record the death on 4 April 2007 of Karen Spärck Jones, who among many other honors and awards, was the President of ACL in 1994 and recipient of the ACL Lifetime Achievement Award in 2004.

These proceedings are dedicated to her memory.

Production and Manufacturing by
Omnipress
2600 Anderson Street
Madison, WI 53704
USA

©2007 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

Table of Contents

Preface: General Chair xv

Preface: Program Chairs xvii

Organizers xix

Program Committee xxi

Conference Program xxv

Guiding Statistical Word Alignment Models With Prior Knowledge
Yonggang Deng and Yuqing Gao 1

A Discriminative Syntactic Word Order Model for Machine Translation
Pi-Chuan Chang and Kristina Toutanova 9

Tailoring Word Alignments to Syntactic Machine Translation
John DeNero and Dan Klein 17

Transductive Learning for Statistical Machine Translation
Nicola Ueffing, Gholamreza Haffari and Anoop Sarkar 25

Word Sense Disambiguation Improves Statistical Machine Translation
Yee Seng Chan, Hwee Tou Ng and David Chiang 33

Learning Expressive Models for Word Sense Disambiguation
Lucia Specia, Mark Stevenson and Maria das Graças Volpe Nunes 41

Domain Adaptation with Active Learning for Word Sense Disambiguation
Yee Seng Chan and Hwee Tou Ng 49

Making Lexical Ontologies Functional and Context-Sensitive
Tony Veale and Yanfen Hao 57

A Bayesian Model for Discovering Typological Implications
Hal Daumé III and Lyle Campbell 65

A Discriminative Language Model with Pseudo-negative Samples
Daisuke Okanohara and Jun'ichi Tsujii 73

Detecting Erroneous Sentences using Automatically Mined Sequential Patterns
Guihua Sun, Xiaohua Liu, Gao Cong, Ming Zhou, Zhongyang Xiong, John Lee and Chin-Yew Lin 81

Vocabulary Decomposition for Estonian Open Vocabulary Speech Recognition
Antti Puurula and Mikko Kurimo 89

Phonological Constraints and Morphological Preprocessing for Grapheme-to-Phoneme Conversion
Vera Demberg, Helmut Schmid and Gregor Möhler 96

Redundancy Ratio: An Invariant Property of the Consonant Inventories of the World's Languages
Animesh Mukherjee, Monojit Choudhury, Anupam Basu and Niloy Ganguly 104

Multilingual Transliteration Using Feature based Phonetic Method
Su-Youn Yoon, Kyoung-Young Kim and Richard Sproat 112

Semantic Transliteration of Personal Names
Haizhou Li, Khe Chai Sim, Jin-Shea Kuo and Minghui Dong 120

Generating Complex Morphology for Machine Translation
Einat Minkov, Kristina Toutanova and Hisami Suzuki 128

Assisting Translators in Indirect Lexical Transfer
Bogdan Babych, Anthony Hartley, Serge Sharoff and Olga Mudraya 136

Forest Rescoring: Faster Decoding with Integrated Language Models
Liang Huang and David Chiang 144

Statistical Machine Translation through Global Lexical Selection and Sentence Reconstruction
Srinivas Bangalore, Patrick Haffner and Stephan Kanthak 152

Mildly Context-Sensitive Dependency Languages
Marco Kuhlmann and Mathias Möhl 160

Transforming Projective Bilexical Dependency Grammars into efficiently-parsable CFGs with Unfold-Fold
Mark Johnson 168

Parsing and Generation as Datalog Queries
Makoto Kanazawa 176

Optimizing Grammars for Minimum Dependency Length
Daniel Gildea and David Temperley 184

Generalizing Semantic Role Annotations Across Syntactically Similar Verbs
Andrew Gordon and Reid Swanson 192

A Grammar-driven Convolution Tree Kernel for Semantic Role Classification
Min Zhang, Wanxiang Che, Aiti Aw, Chew Lim Tan, Guodong Zhou, Ting Liu and Sheng Li . . 200

Learning Predictive Structures for Semantic Role Labeling of NomBank
Chang Liu and Hwee Tou Ng 208

A Simple, Similarity-based Model for Selectional Preferences
Katrin Erk 216

SVM Model Tampering and Anchored Learning: A Case Study in Hebrew NP Chunking
Yoav Goldberg and Michael Elhadad 224

Fully Unsupervised Discovery of Concept-Specific Relationships by Web Mining
Dmitry Davidov, Ari Rappoport and Moshe Koppel 232

Adding Noun Phrase Structure to the Penn Treebank
David Vadas and James Curran 240

Formalism-Independent Parser Evaluation with CCG and DepBank
Stephen Clark and James Curran 248

Frustratingly Easy Domain Adaptation
Hal Daumé III 256

Instance Weighting for Domain Adaptation in NLP
Jing Jiang and ChengXiang Zhai 264

The Infinite Tree
Jenny Rose Finkel, Trond Grenager and Christopher D. Manning 272

Guiding Semi-Supervision with Constraint-Driven Learning
Ming-Wei Chang, Lev Ratinov and Dan Roth 280

Supertagged Phrase-Based Statistical Machine Translation
Hany Hassan, Khalil Sima'an and Andy Way 288

Regression for Sentence-Level MT Evaluation with Pseudo References
Joshua Albrecht and Rebecca Hwa 296

Bootstrapping Word Alignment via Word Packing
Yanjun Ma, Nicolas Stroppa and Andy Way 304

Improved Word-Level System Combination for Machine Translation
Antti-Veikko Rosti, Spyros Matsoukas and Richard Schwartz 312

Generating Constituent Order in German Clauses
Katja Filippova and Michael Strube 320

A Symbolic Approach to Near-Deterministic Surface Realisation using Tree Adjoining Grammar
Claire Gardent and Eric Kow 328

Sentence Generation as a Planning Problem
Alexander Koller and Matthew Stone 336

GLEU: Automatic Evaluation of Sentence-Level Fluency
Andrew Mutton, Mark Dras, Stephen Wan and Robert Dale 344

Conditional Modality Fusion for Coreference Resolution
Jacob Eisenstein and Randall Davis 352

The Utility of a Graphical Representation of Discourse Structure in Spoken Dialogue Systems
Mihai Rotaru and Diane Litman 360

Automated Vocabulary Acquisition and Interpretation in Multimodal Conversational Systems
Yi Liu, Joyce Chai and Rong Jin 368

A Multimodal Interface for Access to Content in the Home
Michael Johnston, Luis Fernando D'Haro, Michelle Levine and Bernard Renger 376

Fast Unsupervised Incremental Parsing
Yoav Seginer 384

K-best Spanning Tree Parsing
Keith Hall 392

Is the End of Supervised Parsing in Sight?
Rens Bod 400

An Ensemble Method for Selection of High Quality Parses
Roi Reichart and Ari Rappoport 408

Opinion Mining using Econometrics: A Case Study on Reputation Systems
Anindya Ghose, Panagiotis Ipeirotis and Arun Sundararajan 416

PageRanking WordNet Synsets: An Application to Opinion Mining
Andrea Esuli and Fabrizio Sebastiani 424

Structured Models for Fine-to-Coarse Sentiment Analysis
Ryan McDonald, Kerry Hannan, Tyler Neylon, Mike Wells and Jeff Reynar 432

Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification
John Blitzer, Mark Dredze and Fernando Pereira 440

Clustering Clauses for High-Level Relation Detection: An Information-theoretic Approach
Samuel Brody 448

Instance-based Evaluation of Entailment Rule Acquisition
Idan Szpektor, Eyal Shnarch and Ido Dagan 456

Statistical Machine Translation for Query Expansion in Answer Retrieval
Stefan Riezler, Alexander Vasserman, Ioannis Tsochantaridis, Vibhu Mittal and Yi Liu 464

A Computational Model of Text Reuse in Ancient Literary Texts
John Lee 472

Finding Document Topics for Improving Topic Segmentation
Olivier Ferret 480

The Utility of Parse-derived Features for Automatic Discourse Segmentation
Seeger Fisher and Brian Roark 488

PERSONAGE: Personality Generation for Dialogue
François Mairesse and Marilyn Walker 496

Making Sense of Sound: Unsupervised Topic Segmentation over Acoustic Input
Igor Malioutov, Alex Park, Regina Barzilay and James Glass 504

Randomised Language Modelling for Statistical Machine Translation
David Talbot and Miles Osborne 512

Bilingual-LSA Based LM Adaptation for Spoken Language Translation
Yik-Cheung Tam, Ian Lane and Tanja Schultz 520

Coreference Resolution Using Semantic Relatedness Information from Automatically Discovered Patterns
Xiaofeng Yang and Jian Su 528

Semantic Class Induction and Coreference Resolution
Vincent Ng 536

Generating a Table-of-Contents
S. R. K. Branavan, Pawan Deshpande and Regina Barzilay 544

Towards an Iterative Reinforcement Approach for Simultaneous Document Summarization and Keyword Extraction
Xiaojun Wan, Jianwu Yang and Jianguo Xiao 552

Fast Semantic Extraction Using a Novel Neural Network Architecture
Ronan Collobert and Jason Weston 560

Improving the Interpretation of Noun Phrases with Cross-linguistic Information
Roxana Girju 568

Learning to Extract Relations from the Web using Minimal Supervision
Razvan Bunescu and Raymond Mooney 576

A Seed-driven Bottom-up Machine Learning Framework for Extracting Relations of Various Complexity
Feiyu Xu, Hans Uszkoreit and Hong Li 584

A Multi-resolution Framework for Information Extraction from Free Text
Mstislav Maslennikov and Tat-Seng Chua 592

Using Corpus Statistics on Entities to Improve Semi-supervised Relation Extraction from the Web
Benjamin Rosenfeld and Ronen Feldman 600

Beyond Projectivity: Multilingual Evaluation of Constraints and Measures on Non-Projective Structures
Jiří Havelka . . . 608

Self-Training for Enhancement and Domain Adaptation of Statistical Parsers Trained on Small Datasets
Roi Reichart and Ari Rappoport . . . 616

HPSG Parsing with Shallow Dependency Constraints
Kenji Sagae, Yusuke Miyao and Jun'ichi Tsujii . . . 624

Constituent Parsing with Incremental Sigmoid Belief Networks
Ivan Titov and James Henderson . . . 632

Corpus Effects on the Evaluation of Automated Transliteration Systems
Sarvnaz Karimi, Andrew Turpin and Falk Scholer . . . 640

Collapsed Consonant and Vowel Models: New Approaches for English-Persian Transliteration and Back-Transliteration
Sarvnaz Karimi, Falk Scholer and Andrew Turpin . . . 648

Alignment-Based Discriminative String Similarity
Shane Bergsma and Grzegorz Kondrak . . . 656

Bilingual Terminology Mining - Using Brain, not Brawn Comparable Corpora
Emmanuel Morin, Béatrice Daille, Koichi Takeuchi and Kyo Kageura . . . 664

Unsupervised Language Model Adaptation Incorporating Named Entity Information
Feifan Liu and Yang Liu . . . 672

Coordinate Noun Phrase Disambiguation in a Generative Parsing Model
Deirdre Hogan . . . 680

A Unified Tagging Approach to Text Normalization
Conghui Zhu, Jie Tang, Hang Li, Hwee Tou Ng and Tiejun Zhao . . . 688

Sparse Information Extraction: Unsupervised Language Models to the Rescue
Doug Downey, Stefan Schoenmackers and Oren Etzioni . . . 696

Forest-to-String Statistical Translation Rules
Yang Liu, Yun Huang, Qun Liu and Shouxun Lin . . . 704

Ordering Phrases with Function Words
Hendra Setiawan, Min-Yen Kan and Haizhou Li . . . 712

A Probabilistic Approach to Syntax-based Reordering for Statistical Machine Translation
Chi-Ho Li, Minghui Li, Dongdong Zhang, Mu Li, Ming Zhou and Yi Guan . . . 720

Machine Translation by Triangulation: Making Effective Use of Multi-Parallel Corpora
Trevor Cohn and Mirella Lapata . . . 728

A Maximum Expected Utility Framework for Binary Sequence Labeling
Martin Jansche 736

A Fully Bayesian Approach to Unsupervised Part-of-Speech Tagging
Sharon Goldwater and Tom Griffiths 744

Computationally Efficient M-Estimation of Log-Linear Structure Models
Noah A. Smith, Douglas L. Vail and John D. Lafferty 752

Guided Learning for Bidirectional Sequence Classification
Libin Shen, Giorgio Satta and Aravind Joshi 760

Different Structures for Evaluating Answers to Complex Questions: Pyramids Won't Topple, and Neither Will Human Assessors
Hoa Trang Dang and Jimmy Lin 768

Exploiting Syntactic and Shallow Semantic Kernels for Question Answer Classification
Alessandro Moschitti, Silvia Quarteroni, Roberto Basili and Suresh Manandhar 776

Language-independent Probabilistic Answer Ranking for Question Answering
Jeongwoo Ko, Teruko Mitamura and Eric Nyberg 784

Learning to Compose Effective Strategies from a Library of Dialogue Components
Martijn Spitters, Marco De Boni, Jakub Zavrel and Remko Bonnema 792

On the Role of Context and Prosody in the Interpretation of 'Okay'
Agustín Gravano, Stefan Benus, Héctor Chávez, Julia Hirschberg and Lauren Wilcox 800

Predicting Success in Dialogue
David Reitter and Johanna D. Moore 808

Resolving It, This, and That in Unrestricted Multi-Party Dialog
Christoph Müller 816

A Comparative Study of Parameter Estimation Methods for Statistical Natural Language Processing
Jianfeng Gao, Galen Andrew, Mark Johnson and Kristina Toutanova 824

Grammar Approximation by Representative Sublanguage: A New Model for Language Learning
Smaranda Muresan and Owen Rambow 832

Chinese Segmentation with a Word-Based Perceptron Algorithm
Yue Zhang and Stephen Clark 840

Unsupervised Coreference Resolution in a Nonparametric Bayesian Model
Aria Haghighi and Dan Klein 848

Pivot Language Approach for Phrase-Based Statistical Machine Translation
Hua Wu and Haifeng Wang 856

Bootstrapping a Stochastic Transducer for Arabic-English Transliteration Extraction
Tarek Sherif and Grzegorz Kondrak 864

Benefits of the Passively Parallel Rosetta Stone? Cross-Language Information Retrieval with over 30 Languages
Peter Chew and Ahmed Abdelali 872

A Re-examination of Machine Learning Approaches for Sentence-Level MT Evaluation
Joshua Albrecht and Rebecca Hwa 880

Automatic Acquisition of Ranked Qualia Structures from the Web
Philipp Cimiano and Johanna Wenderoth 888

A Sequencing Model for Situation Entity Classification
Alexis Palmer, Elias Ponvert, Jason Baldridge and Carlota Smith 896

Words and Echoes: Assessing and Mitigating the Non-Randomness Problem in Word Frequency Distribution Modeling
Baroni Marco and Evert Stefan 904

A System for Large-Scale Acquisition of Verbal, Nominal and Adjectival Subcategorization Frames from Corpora
Judita Preiss, Ted Briscoe and Anna Korhonen 912

A Language-Independent Unsupervised Model for Morphological Segmentation
Vera Demberg 920

Using Mazurkiewicz Trace Languages for Partition-Based Morphology
François Barthélemy 928

Much ado about nothing: A Social Network Model of Russian Paradigmatic Gaps
Robert Daland, Andrea D. Sims and Janet Pierrehumbert 936

Substring-Based Transliteration
Tarek Sherif and Grzegorz Kondrak 944

Pipeline Iteration
Kristy Hollingshead and Brian Roark 952

Learning Synchronous Grammars for Semantic Parsing with Lambda Calculus
Yuk Wah Wong and Raymond Mooney 960

Generalizing Tree Transformations for Inductive Dependency Parsing
Jens Nilsson, Joakim Nivre and Johan Hall 968

Learning Multilingual Subjective Language via Cross-Lingual Projections
Rada Mihalcea, Carmen Banea and Janyce Wiebe 976

Sentiment Polarity Identification in Financial News: A Cohesion-based Approach
Ann Devitt and Khurshid Ahmad 984

Weakly Supervised Learning for Hedge Classification in Scientific Literature
Ben Medlock and Ted Briscoe 992

Text Analysis for Automatic Image Annotation
Koen Deschacht and Marie-Francine Moens 1000

User Requirements Analysis for Meeting Information Retrieval Based on Query Elicitation
Vincenzo Pallotta, Violeta Seretan and Marita Ailomaa 1008

Combining Multiple Knowledge Sources for Dialogue Segmentation in Multimedia Archives
Pei-Yun Hsueh and Johanna D. Moore 1016

Topic Analysis for Psychiatric Document Retrieval
Liang-Chih Yu, Chung-Hsien Wu, Chin-Yew Lin, Eduard Hovy and Chia-Ling Lin 1024

What to be? - Electronic Career Guidance Based on Semantic Relatedness
Iryna Gurevych, Christof Müler and Torsten Zesch 1032

Extracting Social Networks and Biographical Facts From Conversational Speech Transcripts
Hongyan Jing, Nanda Kambhatla and Salim Roukos 1040

Author Index 1049

Preface: General Chair

On behalf of the organizing committee I am delighted to welcome you to the 45th Annual Meeting of the Association for Computational Linguistics, in Prague.

Setting up and running the ACL conference involves a lot of work by many people. Some of them are officially identified as being responsible for various aspects of the conference, while the contributions of others are less visible. I would like to say a warm thank you to the people named below, with apologies to anyone I have overlooked.

The Program Chairs, **Antal van den Bosch** and **Annie Zaenen** have done a great job in managing the almost 600 submissions for the main conference and putting together a high quality program. Through this process Antal has become a 'grandmaster' of the START paper management system and has given a lot of help to other chairs who have had to deal with their own sets of submissions. Many thanks also to their Area Chairs and the program committee of reviewers, and to **Florence Reeder** for coordinating the pre-submission mentoring service. (Antal and Annie reflect on their PC experience overleaf).

Sophia Ananiadou is Chair of the Demo/Poster part of the conference, and has overseen a separate review process to select a high quality set of presentations.

The Student Research Workshop Chairs, **Chris Biemann**, **Violeta Seretan** and **Ellen Riloff** have assembled an excellent program of papers and posters. I encourage everyone to attend the student workshop to hear about the exciting work being carried out by researchers just starting out on their careers in computational linguistics.

Workshops Chair **Simone Teufel** is overseeing 15 workshops – the most ever at an ACL conference – chosen (with the help of **Beth Ann Hockey**, **Katja Markert** and **Dekai Wu**) from a total of 27 proposals. The scale of the workshop program can be gauged by the fact that they received an aggregate total of 470 submissions (without even counting IWPT and EMNLP-CoNLL). **Joakim Nivre**, as Tutorials Chair, has assembled a program of 5 attractive and complementary tutorials, selected from 20 proposals with advice from **Walter Daelemans**, **Robert Dale**, **Nancy Ide**, **Diane Litman** and **Chris Manning**.

One of the most demanding yet least noticed organizational roles is that of Publications Chair. **Su Jian** has done a fantastic job in producing the hardcopy and electronic record of the conference, supported by his team—with notable contributions from **Upali Kohomban** who has cheerfully helped at all stages and whenever needed, day and night, weekdays and weekends.

Sponsorship is another success story, thanks to the Sponsorship Chairs **Martha Palmer**, **Gabor Proszeky**, **Jan Hajic** and **Jun'ichi Tsujii**, who have recruited 12 corporate sponsors. We are very grateful for their financial support.

Eva Hajicova, the Local Arrangements Chair, assisted by **Jan Hajic** and **Anna Kotesovcova**, have put in an enormous amount of detailed work to make this conference a success, ably supported by the local team of **Milan Fucik**, **Jaroslava Hlavacova**, **Marketa Lopatkova**, **Jiri Mirovsky**, **Pavel Pecina**,

Pavel Schlesinger, **Juraj Simlovic**, **Miroslav Spousta**, **Pavel Stranak**, **Zlatka Subrova**, **Jan Votrubec**, **Zdenek Zabokrtsky** and **Daniel Zeman**. From the ACL itself, **Kathy McCoy**, **Dragomir Radev**, **Priscilla Rasmussen**, **Mark Steedman** and **Jun'ichi Tsujii** have played strong roles in making decisions and giving advice, and kept everything on track while I was out of action due to ill health last year.

And finally, many thanks to the authors and presenters in the main conference, workshops and co-located events... and to all the participants. I hope you enjoy the formally organized aspects of the conference, take advantage of the opportunity to network with colleagues old and new, and that you also have a chance to appreciate the history and sights of Prague while you are here.

John Carroll
ACL 2007 General Chair

Preface: Program Chairs

The number of submissions for this ACL broke a new record: the program committee's selection of 131 papers was based on 588 submissions (after withdrawals). An updated program design with four parallel sessions and 25-minute papers allowed an acceptance rate of 22.3%, and an acceptance of all submissions that were recommended with priority by the area chairs.

First and foremost, we thank all the authors for submitting papers describing their recent work; the sheer amount of submissions reflects how active our field is. We thank **Florence Reeder** for provided mentoring to 17 author teams who felt they needed some writing support. For the selection, we are indebted to the 332 program committee members, who produced one to eleven reviews per reviewer, for a total of close to 1,800 reviews, and to the ten area chairs on whose shoulders rested most of the work of organizing the review process. We decided to work without an area chairs meeting: the two program co-chairs met for two days at Tilburg University, and interacted during that time vigorously with the area chairs by email and sometimes by phone.

As usual the main program will run for three days: there will be four parallel sessions of main session presentations, a demo/poster session organized by **Sophia Ananiadou**, nameMiroslav Spousta and **Zdenek Zabokrtsky**, and a Student Research Workshop – thanks to **Ellen Riloff**, **Violeta Seretan** and **Chris Biemann** for organizing it. Also as usual the conference is flanked by tutorial sessions and workshops; our thanks go to **Joakim Nivre** and **Simone Teufel** for organizing and compiling an excellent package.

The announcements of the ACL Lifetime Award and of the Best Paper Award will provide the customary suspense. They will take place in plenary sessions. Other plenary sessions will be devoted to the business meeting and the two invited talks, which this year will be delivered by **Tom Mitchell** and **Barney Pell**. We are grateful for their kind acceptation of our invitation.

We thank **John Carroll**, General Conference Chair, the Local Arrangements Committee headed by **Eva Hajicova**, and the ACL executive, especially **Dragomir Radev**, for their help and advice, and last year's co-chairs, **Claire Cardie** and **Pierre Isabelle**, for sharing their experience. Our sincere thanks go to **Su Jian** for putting together the proceedings.

Enjoy the conference,

Antal van den Bosch and Annie Zaenen
ACL-2007 Program Chairs

Organizers

General Chair:

John Carroll, University of Sussex, UK

Local Arrangements Chair:

Eva Hajicova, Charles University, Czech Republic

Program Chairs:

Antal van den Bosch, Tilburg University, The Netherlands
Annie Zaenen, Palo Alto Research Center (PARC), USA

Student Research Workshop:

Violeta Seretan, University of Geneva, Switzerland
Chris Biemann, University of Leipzig, Germany
Ellen Riloff (Faculty Advisor), University of Utah, USA

Workshop Chair:

Simone Teufel, University of Cambridge, UK

Tutorial Chair:

Joakim Nivre, Växjö University, Sweden

Demo/Poster Chair:

Sophia Ananiadou, The University of Manchester, UK

Exhibits Chairs:

Jaroslava Hlavacova, Charles University, Czech Republic
Pavel Pecina, Charles University, Czech Republic

Sponsorship Chairs:

Martha Palmer, University of Colorado, USA
Gabor Proszeky, Morphologic, Hungary
Jan Hajic, Charles University, Czech Republic
Jun'ichi Tsujii, University of Tokyo, Japan

Publicity Chairs:

Pavel Stranak, Charles University, Czech Republic
Jiri Mirovsky, Charles University, Czech Republic

Publication Chair:

Su Jian, Institute for Infocomm Research (I2R), Singapore

Mentoring Service:

Florence Reeder, The MITRE Corporation, USA

Student Volunteers:

Marketa Lopatkova, Charles University, Czech Republic

Webmasters:

Zlatka Subrova, Czech Republic
Juraj Simlovic, Czech Republic

Secretariat:

Anna Kotesovcova, Charles University, Czech Republic

Registration:

Priscilla Rasmussen, Association for Computational Linguistics (ACL)

ACL Executive Committee:

Mark Steedman, University of Edinburgh, UK
Bonnie Dorr, University of Maryland, USA
Steven Bird, University of Melbourne, Australia
Jun'ichi Tsujii, The University of Tokyo, Japan
Kathleen F. McCoy, University of Delaware, USA
Dragomir R. Radev, University of Michigan, USA
Owen Rambow, Columbia University, USA
Alex Lascarides, University of Edinburgh, UK
Keh-Yih Su, Behavior Design Corporation, Taiwan
Claire Cardie, Cornell University, USA
Nicoletta Calzolari, Istituto di Linguistica Computazionale del CNR, Italy

Program Committee

Program Chairs:

Antal van den Bosch, Tilburg University, The Netherlands
Annie Zaenen, Palo Alto Research Center, USA

Area Chairs:

Tim Baldwin, University of Melbourne, Australia
Kees van Deemter, University of Aberdeen, UK
Barbara Di Eugenio, University of Illinois at Chicago, USA
Josef van Genabith, Dublin City University, Ireland
Claire Grover, University of Edinburgh, UK
Diana McCarthy, University of Sussex, UK
Dan Roth, University of Illinois at Urbana-Champaign, USA
Richard Sproat, University of Illinois at Urbana-Champaign, USA
Marc Swerts, Tilburg University, The Netherlands
Andy Way, Dublin City University, Ireland

Program Committee Members:

Takeshi Abekawa, Eneko Agirre, Gregory Aist, Cyril Allauzen, Elisabeth André, Shlomo Argamon, Ron Artstein,

Collin Baker, Srinivas Bangalore, Colin Bannard, Regina Barzilay, Roberto Basili, John Bateman, Kenneth Beesley, Anja Belz, Slaven Bilac, Steven Bird, Guido Boella, Francis Bond, Kalina Bontcheva, Johan Bos, Gosse Bouma, Susan Brennan, Chris Brew, Ted Briscoe, Ralf Brown, Paul Buitelaar, Razvan Bunescu, Harry Bunt,

Aoife Cahill, Janet Cahn, Mary Elaine Califf, Charles Callaway, Chris Callison-Burch, Nicoletta Calzolari, Giuseppe Carenini, Michael Carl, Jean Carletta, Xavier Carreras, Justine Cassell, Joyce Chai, Jing-Shin Chang, Ming-Wei Chang, Jinying Chen, Keh-Jiann Chen, Colin Cherry, David Chiang, Grzegorz Chrupala, Jennifer Chu-Carroll, Ilyas Cicekli, Philipp Cimiano, Stephen Clark, Michael Collins, Michael Connor, Mark Core, Mathias Creutz, James Curran,

Walter Daelemans, Ido Dagan, Hercules Dalianis, Hal Daume, Eric de la Clergerie, Maarten de Rijke, Barbara Di Eugenio, Mona Diab, Gael Diaz, Bonnie Dorr, Laila Dybkjær,

Jason Eisner, Noemie Elhadad, Katrin Erk, Dominique Estival, Stefan Evert,

David Farwell, Afsaneh Fazly, Marcello Federico, Katja Filippova, Kate Forbes Riley, Mikel Forcada, George Foster, Mary Ellen Foster, Anette Frank, Reva Freedman, Guohong Fu, Pascale Fung,

Rob Gaizauskas, Jianfeng Gao, Kallirroi Georgila, Daniel Gildea, Roxana Girju, Alfio Gliozzo, Sharon Goldwater, Fernando Gomez, Nancy Green, Mark Greenwood, Ralph Grishman, Declan Groves,

Nizar Habash, Keith Hall, Susan Haller, Sanda Harabagiu, Henk Harkema, Tony Hartley, Sven Hartrumpf, Mary Hearne, Marti Hearst, Peter Heeman, James Henderson, Mark Hepple, Ryuichiro Higashinaka, Julia Hirschberg, Graeme Hirst, Barbora Hladka, Julia Hockenmaier, Deirdre Hogan, Baden Hughes,

Diana Inkpen, Kentaro Inui,

Martin Jansche, Valentin Jijkoun, Mark Johnson, Kristiina Jokinen, Pamela Jordan,

Laura Kallmeyer, Stephan Kanthak, Vangelis Karkaletsis, Daisuke Kawahara, John Kelleher, Frank Keller, Andre Kempe, Rodger Kibble, Adam Kilgarriff, Tracy King, Ewan Klein, Alexandre Klementiev, Kevin Knight, Alistair Knott, Philipp Koehn, Rob Koeling, Alexander Koller, Grzegorz Kondrak, Moshe Koppel, Valia Kordoni, Anna Korhonen, Emiel Krahmer, Sandra Kuebler, Jonas Kuhn, Sadao Kurohashi,

Philippe Langlais, Guy Lapalme, Mirella Lapata, Staffan Larsson, Alberto Lavelli, Alon Lavie, Lillian Lee, Jochen Leidner, Oliver Lemon, Yves Lepage, Leonardo Lesmo, Gina Levow, Ian Lewin, William Lewis, Mu Li, Dekang Lin, Jimmy Lin, Diane Litman, Ting Liu, Saturnino Luz,

Bernardo Magnini, Gideon Mann, Chris Manning, Daniel Marcu, Katja Markert, Lluis Marquez, Erwin Marsi, David Martinez, Carlos Martin-Vide, Evgeny Matusov, John Maxwell, Mike Maxwell, Diana Maynard, Andrew McCallum, Ryan McDonald, Kathy McKeown, Susan McRoy, Mike McTear, Dan Melamed, Chris Mellish, Arul Menezes, Helen Meng, Paola Merlo, Detmar Meurers, Rada Mihalcea, Maria Milosavljevic, Eleni Miltsakaki, David Milward, Yusuke Miyao, Dan Moldovan, Diego Molla, Raymond Mooney, Roger Moore, Alessandro Moschitti, Karin Mueller, Reinhard Muskens,

Hiroshi Nakagawa, Hiromi Nakaiwa, Roberto Navigli, Mark-Jan Nederhof, Goran Nenadic, John Nerbonne, Hwee Tou Ng, Vincent Ng, Jian-Yun Nie, Malvina Nissim, Joakim Nivre, David Novick,

Jon Oberlander, Franz Och, Stephan Oepen, Kemal Oflazer, Ian O'Neill, Miles Osborne, Mari Ostendorf,

Tim Paek, Martha Palmer, Patrick Pantel, Ivandre Paraboni, Becky Passonneau, Michael Paul, Fernando Pereira, Scott Piao, Manfred Pinkal, Paul Piwek, Massimo Poesio, Joe Polifroni, Richard Power, John Prager, Stephen Pulman, Vasin Punyakanok,

Chris Quirk,

Reinhard Rapp, Lev Ratinov, Ehud Reiter, Martin Reynaert, Stefan Riezler, German Rigau, Ellen Riloff, Brian Roark, Laurent Romary,

Bogdan Sacaleanu, Horacio Saggion, Anna Sagvall Hein, Mark Sanderson, Anoop Sarkar, Avik Sarkar, Helmut Schmid, Gerold Schneider, Patrick Schone, Hinrich Schuetze, Sabine Schulte im Walde, Satoshi Sekine, Stephanie Seneff, Izhak Shafran, Advaith Siddharthan, Khalil Simaan, Kevin Small, Noah A. Smith, Harold Somers, Virach Sornlertlamvanich, Manfred Stede, Mark Steedman, Amanda Stent, Mark Stevenson, Carlo Strapparava, Oliver Streiter, Nicolas Stroppa, Michael Strube, Jian Su, Rajen Subba, Stan Szpakowicz,

Maite Taboada, Hiroya Takamura, Kumiko Tanaka-Ishii, Louis ten Bosch, Joel Tetreault, Simone Teufel, Mariet Theune, Joerg Tiedemann, Takenobu Tokunaga, Kristina Toutanova, Isabel Trancoso, Richard Tsai, Gokhan Tur,

Nicola Ueffing, Nathan Vaillette, Gertjan van Noord, Menno van Zaanen, Sebastian Varges, Paola Velardi, Renata Vieira, Begona Villada Moiron, Carl Vogel, Stephan Vogel, Piek Vossen,

Marilyn Walker, Haifeng Wang, Xinglong Wang, Bonnie Webber, David Weir, Ben Wellner, Richard Wicentowski, Janyce Wiebe, Ross Wilkinson, Yorick Wilks, Theresa Wilson, Shuly Wintner, Dekai Wu,

Fei Xia, Nianwen Xue, Endong Xun,

Muyun Yang, Scott Wen-tau Yih, Anssi Yli-Jyrä,

Dmitry Zelenko, Richard Zens, ChengXiang Zhai, Min Zhang, Tong Zhang, Guodong Zhou

Conference Program

Monday, June 25, 2007

8:45–9:00 Opening Session

9:00–10:00 Invited Talk by Tom Mitchell

10:00–10:30 Break

Session 1A: Machine Translation 1

10:30–10:55 *Guiding Statistical Word Alignment Models With Prior Knowledge*
Yonggang Deng and Yuqing Gao

10:55–11:20 *A Discriminative Syntactic Word Order Model for Machine Translation*
Pi-Chuan Chang and Kristina Toutanova

11:20–11:45 *Tailoring Word Alignments to Syntactic Machine Translation*
John DeNero and Dan Klein

11:45–12:10 *Transductive Learning for Statistical Machine Translation*
Nicola Ueffing, Gholamreza Haffari and Anoop Sarkar

Session 1B: Word Sense Disambiguation

10:30–10:55 *Word Sense Disambiguation Improves Statistical Machine Translation*
Yee Seng Chan, Hwee Tou Ng and David Chiang

10:55–11:20 *Learning Expressive Models for Word Sense Disambiguation*
Lucia Specia, Mark Stevenson and Maria das Graças Volpe Nunes

11:20–11:45 *Domain Adaptation with Active Learning for Word Sense Disambiguation*
Yee Seng Chan and Hwee Tou Ng

11:45–12:10 *Making Lexical Ontologies Functional and Context-Sensitive*
Tony Veale and Yanfen Hao

Monday, June 25, 2007 (continued)

Session 1C: Language Modeling 1

10:30–10:55 *A Bayesian Model for Discovering Typological Implications*
Hal Daumé III and Lyle Campbell

10:55–11:20 *A Discriminative Language Model with Pseudo-negative Samples*
Daisuke Okanohara and Jun'ichi Tsujii

11:20–11:45 *Detecting Erroneous Sentences using Automatically Mined Sequential Patterns*
Guihua Sun, Xiaohua Liu, Gao Cong, Ming Zhou, Zhongyang Xiong, John Lee and Chin-Yew Lin

11:45–12:10 *Vocabulary Decomposition for Estonian Open Vocabulary Speech Recognition*
Antti Puurula and Mikko Kurimo

Session 1D: Phonology and Morphology 1

10:30–10:55 *Phonological Constraints and Morphological Preprocessing for Grapheme-to-Phoneme Conversion*
Vera Demberg, Helmut Schmid and Gregor Möhler

10:55–11:20 *Redundancy Ratio: An Invariant Property of the Consonant Inventories of the World's Languages*
Animesh Mukherjee, Monojit Choudhury, Anupam Basu and Niloy Ganguly

11:20–11:45 *Multilingual Transliteration Using Feature based Phonetic Method*
Su-Youn Yoon, Kyoung-Young Kim and Richard Sproat

11:45–12:10 *Semantic Transliteration of Personal Names*
Haizhou Li, Khe Chai Sim, Jin-Shea Kuo and Minghui Dong

12:10–13:30 Lunch

Session 2A: Machine Translation 2

13:30–13:55 *Generating Complex Morphology for Machine Translation*
Einat Minkov, Kristina Toutanova and Hisami Suzuki

13:55–14:20 *Assisting Translators in Indirect Lexical Transfer*
Bogdan Babych, Anthony Hartley, Serge Sharoff and Olga Mudraya

Monday, June 25, 2007 (continued)

14:20–14:45 *Forest Rescoring: Faster Decoding with Integrated Language Models*
Liang Huang and David Chiang

14:45–15:10 *Statistical Machine Translation through Global Lexical Selection and Sentence Reconstruction*
Srinivas Bangalore, Patrick Haffner and Stephan Kanthak

Session 2B: Grammars

13:30–13:55 *Mildly Context-Sensitive Dependency Languages*
Marco Kuhlmann and Mathias Möhl

13:55–14:20 *Transforming Projective Bilexical Dependency Grammars into Efficiently-parsable CFGs with Unfold-Fold*
Mark Johnson

14:20–14:45 *Parsing and Generation as Datalog Queries*
Makoto Kanazawa

14:45–15:10 *Optimizing Grammars for Minimum Dependency Length*
Daniel Gildea and David Temperley

Session 2C: Semantic Role Labeling

13:30–13:55 *Generalizing Semantic Role Annotations Across Syntactically Similar Verbs*
Andrew Gordon and Reid Swanson

13:55–14:20 *A Grammar-driven Convolution Tree Kernel for Semantic Role Classification*
Min Zhang, Wanxiang Che, Aiti Aw, Chew Lim Tan, Guodong Zhou, Ting Liu and Sheng Li

14:20–14:45 *Learning Predictive Structures for Semantic Role Labeling of NomBank*
Chang Liu and Hwee Tou Ng

14:45–15:10 *A Simple, Similarity-based Model for Selectional Preferences*
Katrin Erk

Monday, June 25, 2007 (continued)

Session 2D: Language Resources

13:30–13:55 *SVM Model Tampering and Anchored Learning: A Case Study in Hebrew NP Chunking*
Yoav Goldberg and Michael Elhadad

13:55–14:20 *Fully Unsupervised Discovery of Concept-Specific Relationships by Web Mining*
Dmitry Davidov, Ari Rappoport and Moshe Koppel

14:20–14:45 *Adding Noun Phrase Structure to the Penn Treebank*
David Vadas and James Curran

14:45–15:10 *Formalism-Independent Parser Evaluation with CCG and DepBank*
Stephen Clark and James Curran

15:10–15:45 Break

Session 3A, Machine Learning Methods 1

15:45–16:10 *Frustratingly Easy Domain Adaptation*
Hal Daumé III

16:10–16:35 *Instance Weighting for Domain Adaptation in NLP*
Jing Jiang and ChengXiang Zhai

16:35–17:00 *The Infinite Tree*
Jenny Rose Finkel, Trond Grenager and Christopher D. Manning

17:00–17:25 *Guiding Semi-Supervision with Constraint-Driven Learning*
Ming-Wei Chang, Lev Ratinov and Dan Roth

Session 3B: Machine Translation 3

15:45–16:10 *Supertagged Phrase-Based Statistical Machine Translation*
Hany Hassan, Khalil Sima'an and Andy Way

16:10–16:35 *Regression for Sentence-Level MT Evaluation with Pseudo References*
Joshua S. Albrecht and Rebecca Hwa

16:35–17:00 *Bootstrapping Word Alignment via Word Packing*
Yanjun Ma, Nicolas Stroppa and Andy Way

17:00–17:25 *Improved Word-Level System Combination for Machine Translation*
Antti-Veikko Rosti, Spyros Matsoukas and Richard Schwartz

Monday, June 25, 2007 (continued)

Session 3C: Generation

15:45–16:10 *Generating Constituent Order in German Clauses*
Katja Filippova and Michael Strube

16:10–16:35 *A Symbolic Approach to Near-Deterministic Surface Realisation using Tree Adjoining Grammar*
Claire Gardent and Eric Kow

16:35–17:00 *Sentence Generation as a Planning Problem*
Alexander Koller and Matthew Stone

17:00–17:25 *GLEU: Automatic Evaluation of Sentence-Level Fluency*
Andrew Mutton, Mark Dras, Stephen Wan and Robert Dale

Session 3D: Multimodality 1

15:45–16:10 *Conditional Modality Fusion for Coreference Resolution*
Jacob Eisenstein and Randall Davis

16:10–16:35 *The Utility of a Graphical Representation of Discourse Structure in Spoken Dialogue Systems*
Mihai Rotaru and Diane Litman

16:35–17:00 *Automated Vocabulary Acquisition and Interpretation in Multimodal Conversational Systems*
Yi Liu, Joyce Chai and Rong Jin

17:00–17:25 *A Multimodal Interface for Access to Content in the Home*
Michael Johnston, Luis Fernando D'Haro, Michelle Levine and Bernard Renger

Tuesday, June 26, 2007

Session 4A, Parsing 1

09:00–09:25 *Fast Unsupervised Incremental Parsing*
Yoav Seginer

09:25–09:50 *K-best Spanning Tree Parsing*
Keith Hall

09:50–10:15 *Is the End of Supervised Parsing in Sight?*
Rens Bod

10:15–10:40 *An Ensemble Method for Selection of High Quality Parses*
Roi Reichart and Ari Rappoport

Session 4B: Sentiment 1

09:00–09:25 *Opinion Mining using Econometrics: A Case Study on Reputation Systems*
Anindya Ghose, Panagiotis Ipeirotis and Arun Sundararajan

09:25–09:50 *PageRanking WordNet Synsets: An Application to Opinion Mining*
Andrea Esuli and Fabrizio Sebastiani

09:50–10:15 *Structured Models for Fine-to-Coarse Sentiment Analysis*
Ryan McDonald, Kerry Hannan, Tyler Neylon, Mike Wells and Jeff Reynar

10:15–10:40 *Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification*
John Blitzer, Mark Dredze and Fernando Pereira

Session 4C: Paraphrasing, Textual Entailment

09:00–09:25 *Clustering Clauses for High-Level Relation Detection: An Information-theoretic Approach*
Samuel Brody

09:25–09:50 *Instance-based Evaluation of Entailment Rule Acquisition*
Idan Szpektor, Eyal Shnarch and Ido Dagan

09:50–10:15 *Statistical Machine Translation for Query Expansion in Answer Retrieval*
Stefan Riezler, Alexander Vasserman, Ioannis Tsochantaridis, Vibhu Mittal and Yi Liu

10:15–10:40 *A Computational Model of Text Reuse in Ancient Literary Texts*
John Lee

Tuesday, June 26, 2007 (continued)

Session 4D: Discourse and Dialog 1

09:00–09:25 *Finding Document Topics for Improving Topic Segmentation*
Olivier Ferret

09:25–09:50 *The Utility of Parse-derived Features for Automatic Discourse Segmentation*
Seeger Fisher and Brian Roark

09:50–10:15 *PERSONAGE: Personality Generation for Dialogue*
François Mairesse and Marilyn Walker

10:15–10:40 *Making Sense of Sound: Unsupervised Topic Segmentation over Acoustic Input*
Igor Malioutov, Alex Park, Regina Barzilay and James Glass

10:40–11:10 Break

11:10-12:10 LifeTime Achievement Award

12:10–13:30 Lunch

13:00-14:30 ACL Business Meeting

Session 5A: Language Modeling 2

14:30–14:55 *Randomised Language Modelling for Statistical Machine Translation*
David Talbot and Miles Osborne

14:55–15:20 *Bilingual-LSA Based LM Adaptation for Spoken Language Translation*
Yik-Cheung Tam, Ian Lane and Tanja Schultz

Session 5B: Coreference

14:30–14:55 *Coreference Resolution Using Semantic Relatedness Information from Automatically Discovered Patterns*
Xiaofeng Yang and Jian Su

14:55–15:20 *Semantic Class Induction and Coreference Resolution*
Vincent Ng

Tuesday, June 26, 2007 (continued)

Session 5C: Summarization

14:30–14:55 *Generating a Table-of-Contents*
S. R. K. Branavan, Pawan Deshpande and Regina Barzilay

14:55–15:20 *Towards an Iterative Reinforcement Approach for Simultaneous Document Summarization and Keyword Extraction*
Xiaojun Wan, Jianwu Yang and Jianguo Xiao

Session 5D: Semantic Relations

14:30–14:55 *Fast Semantic Extraction Using a Novel Neural Network Architecture*
Ronan Collobert and Jason Weston

14:55–15:20 *Improving the Interpretation of Noun Phrases with Cross-linguistic Information*
Roxana Girju

15:20–15:45 Break

Session 6A: Information Extraction

15:45–16:10 *Learning to Extract Relations from the Web using Minimal Supervision*
Razvan Bunescu and Raymond Mooney

16:10–16:35 *A Seed-driven Bottom-up Machine Learning Framework for Extracting Relations of Various Complexity*
Feiyu Xu, Hans Uszkoreit and Hong Li

16:35–17:00 *A Multi-resolution Framework for Information Extraction from Free Text*
Mstislav Maslennikov and Tat-Seng Chua

17:00–17:25 *Using Corpus Statistics on Entities to Improve Semi-supervised Relation Extraction from the Web*
Benjamin Rosenfeld and Ronen Feldman

Tuesday, June 26, 2007 (continued)

Session 6B: Parsing 2

15:45–16:10 *Beyond Projectivity: Multilingual Evaluation of Constraints and Measures on Non-Projective Structures*
Jiří Havelka

16:10–16:35 *Self-Training for Enhancement and Domain Adaptation of Statistical Parsers Trained on Small Datasets*
Roi Reichart and Ari Rappoport

16:35–17:00 *HPSG Parsing with Shallow Dependency Constraints*
Kenji Sagae, Yusuke Miyao and Jun'ichi Tsujii

17:00–17:25 *Constituent Parsing with Incremental Sigmoid Belief Networks*
Ivan Titov and James Henderson

Session 6C: Multilinguality 1

15:45–16:10 *Corpus Effects on the Evaluation of Automated Transliteration Systems*
Sarvnaz Karimi, Andrew Turpin and Falk Scholer

16:10–16:35 *Collapsed Consonant and Vowel Models: New Approaches for English-Persian Transliteration and Back-Transliteration*
Sarvnaz Karimi, Falk Scholer and Andrew Turpin

16:35–17:00 *Alignment-Based Discriminative String Similarity*
Shane Bergsma and Grzegorz Kondrak

17:00–17:25 *Bilingual Terminology Mining - Using Brain, not brawn comparable corpora*
Emmanuel Morin, Béatrice Daille, Koichi Takeuchi and Kyo Kageura

Session 6D: Language Modeling 3

15:45–16:10 *Unsupervised Language Model Adaptation Incorporating Named Entity Information*
Feifan Liu and Yang Liu

16:10–16:35 *Coordinate Noun Phrase Disambiguation in a Generative Parsing Model*
Deirdre Hogan

16:35–17:00 *A Unified Tagging Approach to Text Normalization*
Conghui Zhu, Jie Tang, Hang Li, Hwee Tou Ng and Tiejun Zhao

17:00–17:25 *Sparse Information Extraction: Unsupervised Language Models to the Rescue*
Doug Downey, Stefan Schoenmackers and Oren Etzioni

Wednesday, June 27, 2007

Session 7A: Machine Translation 4

09:00–09:25 *Forest-to-String Statistical Translation Rules*
Yang Liu, Yun Huang, Qun Liu and Shouxun Lin

09:25–09:50 *Ordering Phrases with Function Words*
Hendra Setiawan, Min-Yen Kan and Haizhou Li

09:50–10:15 *A Probabilistic Approach to Syntax-based Reordering for Statistical Machine Translation*
Chi-Ho Li, Minghui Li, Dongdong Zhang, Mu Li, Ming Zhou and Yi Guan

10:15–10:40 *Machine Translation by Triangulation: Making Effective Use of Multi-Parallel Corpora*
Trevor Cohn and Mirella Lapata

Session 7B: Sequence Processing

09:00–09:25 *A Maximum Expected Utility Framework for Binary Sequence Labeling*
Martin Jansche

09:25–09:50 *A Fully Bayesian Approach to Unsupervised Part-of-speech Tagging*
Sharon Goldwater and Tom Griffiths

09:50–10:15 *Computationally Efficient M-Estimation of Log-Linear Structure Models*
Noah A. Smith, Douglas L. Vail and John D. Lafferty

10:15–10:40 *Guided Learning for Bidirectional Sequence Classification*
Libin Shen, Giorgio Satta and Aravind Joshi

Session 7C: Question Answering

09:00–09:25 *Different Structures for Evaluating Answers to Complex Questions: Pyramids Won't Topple, and Neither Will Human Assessors*
Hoa Trang Dang and Jimmy Lin

09:25–09:50 *Exploiting Syntactic and Shallow Semantic Kernels for Question Answer Classification*
Alessandro Moschitti, Silvia Quarteroni, Roberto Basili and Suresh Manandhar

09:50–10:15 *Language-independent Probabilistic Answer Ranking for Question Answering*
Jeongwoo Ko, Teruko Mitamura and Eric Nyberg

Wednesday, June 27, 2007 (continued)

Session 7D: Discourse and Dialog 2

09:00–09:25 *Learning to Compose Effective Strategies from a Library of Dialogue Components*
Martijn Spitters, Marco De Boni, Jakub Zavrel and Remko Bonnema

09:25–09:50 *On the Role of Context and Prosody in the Interpretation of 'okay'*
Agustín Gravano, Stefan Benus, Héctor Chávez, Julia Hirschberg and Lauren Wilcox

09:50–10:15 *Predicting Success in Dialogue*
David Reitter and Johanna D. Moore

10:15–10:40 *Resolving It, This, and That in Unrestricted Multi-Party Dialog*
Christoph Müller

10:40–11:10 Break

11:10–12:10 Invited Talk by Barney Pell

12:10–13:30 Lunch

Session 8A: Machine Learning Methods 2

13:30–13:55 *A Comparative Study of Parameter Estimation Methods for Statistical Natural Language Processing*
Jianfeng Gao, Galen Andrew, Mark Johnson and Kristina Toutanova

13:55–14:20 *Grammar Approximation by Representative Sublanguage: A New Model for Language Learning*
Smaranda Muresan and Owen Rambow

14:20–14:45 *Chinese Segmentation with a Word-Based Perceptron Algorithm*
Yue Zhang and Stephen Clark

14:45–15:10 *Unsupervised Coreference Resolution in a Nonparametric Bayesian Model*
Aria Haghighi and Dan Klein

Wednesday, June 27, 2007 (continued)

Session 8B: Machine Translation and Multilinguality

13:30–13:55 *Pivot Language Approach for Phrase-Based Statistical Machine Translation*
Hua Wu and Haifeng Wang

13:55–14:20 *Bootstrapping a Stochastic Transducer for Arabic-English Transliteration Extraction*
Tarek Sherif and Grzegorz Kondrak

14:20–14:45 *Benefits of the Passively Parallel Rosetta Stone? Cross-Language Information Retrieval with over 30 Languages*
Peter A. Chew and Ahmed Abdelali

14:45–15:10 *A Re-examination of Machine Learning Approaches for Sentence-Level MT Evaluation*
Joshua S. Albrecht and Rebecca Hwa

Session 8C: Lexicon and Lexical Semantics

13:30–13:55 *Automatic Acquisition of Ranked Qualia Structures from the Web*
Philipp Cimiano and Johanna Wenderoth

13:55–14:20 *A Sequencing Model for Situation Entity Classification*
Alexis Palmer, Elias Ponvert, Jason Baldridge and Carlota Smith

14:20–14:45 *Words and Echoes: Assessing and Mitigating the Non-Randomness Problem in Word Frequency Distribution Modeling*
Marco Baroni and Stefan Evert

14:45–15:10 *A System for Large-Scale Acquisition of Verbal, Nominal and Adjectival Subcategorization Frames from Corpora*
Judita Preiss, Ted Briscoe and Anna Korhonen

Wednesday, June 27, 2007 (continued)

Session 8D: Phonology and Morphology 2

13:30–13:55 *A Language-Independent Unsupervised Model for Morphological Segmentation*
Vera Demberg

13:55–14:20 *Using Mazurkiewicz Trace Languages for Partition-Based Morphology*
François Barthélemy

14:20–14:45 *Much ado about nothing: A Social Network Model of Russian Paradigmatic Gaps*
Robert Daland, Andrea D. Sims and Janet Pierrehumbert

14:45–15:10 *Substring-Based Transliteration*
Tarek Sherif and Grzegorz Kondrak

15:10–15:45 Break

Session 9A: Parsing 3

15:45–16:10 *Pipeline Iteration*
Kristy Hollingshead and Brian Roark

16:10–16:35 *Learning Synchronous Grammars for Semantic Parsing with Lambda Calculus*
Yuk Wah Wong and Raymond Mooney

16:35–17:00 *Generalizing Tree Transformations for Inductive Dependency Parsing*
Jens Nilsson, Joakim Nivre and Johan Hall

Session 9B: Sentiment 2

15:45–16:10 *Learning Multilingual Subjective Language via Cross-Lingual Projections*
Rada Mihalcea, Carmen Banea and Janyce Wiebe

16:10–16:35 *Sentiment Polarity Identification in Financial News: A Cohesion-based Approach*
Ann Devitt and Khurshid Ahmad

16:35–17:00 *Weakly Supervised Learning for Hedge Classification in Scientific Literature*
Ben Medlock and Ted Briscoe

Wednesday, June 27, 2007 (continued)

Session 9C: Multimodality 2

15:45–16:10 *Text Analysis for Automatic Image Annotation*
Koen Deschacht and Marie-Francine Moens

16:10–16:35 *User Requirements Analysis for Meeting Information Retrieval Based on Query Elicitation*
Vincenzo Pallotta, Violeta Seretan and Marita Ailomaa

16:35–17:00 *Combining Multiple Knowledge Sources for Dialogue Segmentation in Multimedia Archives*
Pei-Yun Hsueh and Johanna D. Moore

Session 9D: Text mining and Retrieval

15:45–16:10 *Topic Analysis for Psychiatric Document Retrieval*
Liang-Chih Yu, Chung-Hsien Wu, Chin-Yew Lin, Eduard Hovy and Chia-Ling Lin

16:10–16:35 *What to be? - Electronic Career Guidance Based on Semantic Relatedness*
Iryna Gurevych, Christof Müller and Torsten Zesch

16:35–17:00 *Extracting Social Networks and Biographical Facts From Conversational Speech Transcripts*
Hongyan Jing, Nanda Kambhatla and Salim Roukos

17:10 Best Paper Award (Plenary Session)