

EACL 2014

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



Proceedings of the Conference

April 26-30, 2014
Gothenburg, Sweden

GOLD SPONSORS



SILVER SPONSOR



BRONZE SPONSORS



SUPPORTERS



EXHIBITORS



OTHER SPONSORS



HOSTS



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

ISBN 978-1-937284-78-7

Preface: General Chair

Welcome to EACL 2014, the 14th Conference of the European Chapter of the Association for Computational Linguistics! This is the largest EACL meeting ever: with eighty long papers, almost fifty short ones, thirteen student research papers, twenty-six demos, fourteen workshops and six tutorials, we expect to bring to Gothenburg up to five hundred participants, for a week of excellent science interspersed with entertaining social events.

It is hard to imagine how much work is involved in the preparation of such an event. It takes about three years, from the day the EACL board starts discussing the location and nominating the chairs, until the final details of the budget are resolved. The number of people involved is also huge, and I was fortunate to work with an excellent, dedicated and efficient team, to which I am enormously grateful.

The scientific program was very ably composed by the Program Committee Chairs, Sharon Goldwater and Stefan Riezler, presiding over a team of twenty-four area chairs. Given that this year we had long paper submissions, followed by a rebuttal period, followed by a very stressed short paper reviewing period, this meant *a lot* of work. Overall, Sharon and Stefan handled over five hundred submissions, or over 1,500 reviews! The result of this work is a balanced, high-quality scientific program that I'm sure we will all enjoy. The PC Chairs have also selected the three invited speakers, and we will have the pleasure of attending keynotes delivered by Simon King, Ulrike von Luxburg, and Dan Roth – a great choice of speakers!

The diverse workshop program was put together by the Workshop Chairs, Anja Belz and Reut Tsarfaty, under very strict deadlines due to the fact that as in previous years, workshops were coordinated with other ACL events (this year, ACL and EMNLP). Even in light of the competition, Anja and Reut negotiated a varied and attractive set of fourteen workshops which will keep us busy over the weekend prior to the main conference.

Also on that weekend are the six tutorials, selected from among several submissions by the Tutorial Chairs, Afra Alishahi and Marco Baroni. Again, the tutorials offer a set of diverse and timely topics, covering both core areas of NLP and tangential fields of research.

We included in the program a large number of demonstrations, selected by Marko Tadić and Bogdan Babych, the Demo Chairs. And an integral part of the scientific program is the Student Research Workshop, put together by the SRW Chairs, Desmond Elliott, Konstantina Garoufi, Douwe Kiela, and Ivan Vulić, whose work was supervised by the SRW Faculty Advisor, Sebastian Padó.

The Proceedings that you're reading now were compiled by the Publication Chairs, Gosse Bouma and Yannick Parmentier. Their responsibilities include the preparation of all the proceedings, including the main session, the SRW, the demo session, the workshop proceedings etc. – thousands of pages, all under very strict deadlines.

It has been a very special pleasure for me to work with an excellent local organization team. The Local Organization Chairs, Lars Borin and Aarne Ranta, were assisted by an extremely efficient team, Yvonne Adesam, Martin Kasá and Nina Tahmasebi. Their effort cannot be overestimated: from dealing with the two universities over issues of conference space and funding, through dealing with two professional conference organizers, to corresponding with authors, participants and of course all the other chairs. Add the stress involved in being in charge of a hefty budget that has to be balanced by the end of the conference, and you can only admire the relaxed way in which they took upon themselves this daunting task.

The local team included also Peter Ljunglöf, the Publicity Chair, to whom we should all be grateful for the beautiful web site of the conference and the timely e-mails, tweets and Facebook statuses. The Local Sponsorship Chairs, Sofie Johansson Kokkinakis and Staffan Larsson, worked together with the ACL

Sponsorship Chairs Jochen Leidner and Alessandro Moschitti, to obtain some much needed financial support. Sincere thanks are due to the various sponsors for their generous contribution.

The local team did a wonderful job organizing a social program this year. This includes a reception at the City Hall on Sunday, a catered poster and demo session on Monday, a conference dinner on Tuesday and of course, the famous Cortège at the very end of the conference. A perfect mix of business and pleasure.

I am grateful to all members of the EACL board for their advice and guidance, and in particular to past Chair Sien Moens, Chair Stephen Clark, Chair-elect Lluís Màrquez and Treasurer Mike Rosner. Many thanks are also due to the ACL Treasurer Graeme Hirst and of course, as always, to the ACL Business Manager Priscilla Rasmussen, who was always there with her vast experience to clear up uncertainties and lend a helping hand.

Finally, let us not forget that this is all about *you*: authors, reviewers, demo presenters, workshop organizers and speakers, tutorial speakers and participants of the conference. Thank you for choosing to be part of EACL-2014, I wish you a very enjoyable conference!

Shuly Wintner, University of Haifa
General Chair
March 2014

Preface: Program Chairs

We are delighted to present you with this volume containing the papers accepted for presentation at the 14th Conference of the European Chapter of the Association for Computational Linguistics, held in Gothenburg, Sweden, from April 26 till April 30 2014.

EACL 2014 introduced a short paper (4 page) format in addition to the usual long paper (8 page) format, which led to the highest total number of submissions of any EACL. We received 317 valid long paper submissions and were able to accept 78 of these papers (an acceptance rate of 24.6%). 49 of the papers (15.4%) were accepted for oral presentation, and 31 (9.8%) for poster presentation. In addition, we received 199 valid short paper submissions and were able to accept 46 of these (an acceptance rate of 23.1%). 33 of the papers (16.6%) were accepted for oral presentation, and 13 (6.5%) for poster presentation. The EACL 2014 schedule also includes oral presentations from two papers published in the Transactions of the Association for Computational Linguistics, a new feature of this year's conference.

The introduction of short papers, handled in a second round of submissions, meant a somewhat higher workload for our program committee, and we are very grateful to our 24 area chairs for recruiting an excellent panel of 434 reviewers from all over the world, and to those reviewers for providing their feedback on the submissions. Each submission was reviewed by at least three reviewers (at least two for short papers), who were then encouraged to discuss any differences of opinion, taking into account the responses of the authors to their initial reviews. Based on the reviews, author response, and reviewer discussion, area chairs provided a ranking for papers in their area. Final selection was made by the program co-chairs after discussion with the area chairs and an independent check of reviews.

Each area chair was also asked to nominate the best long paper and best short paper from his or her area, or to decline to nominate any. Several papers were nominated, and of these the program co-chairs made the final decision on the Best Long Paper and Best Short Paper awards, which will be awarded in a plenary session at the conference.

In addition to the main conference program, EACL 2014 will feature the now traditional Student Research Workshop, 14 other workshops, 6 tutorials and a demo session with 26 presentations. We are also fortunate to have three excellent invited speakers: Dan Roth (University of Illinois at Urbana-Champaign), Ulrike von Luxburg (University of Hamburg), and Simon King (University of Edinburgh).

We would very much like to thank all of the other people who have helped us put together this year's conference. Most importantly, all of the authors who submitted their work to EACL, without whom we would have no conference at all! The number and quality of both long and short paper submissions in many different areas shows that we are maintaining and growing a broad and active community. We are greatly indebted to all the area chairs and reviewers for their hard work, which allowed us to choose from amongst the many high-quality submissions to put together a strong programme and provide useful feedback to authors. The START support team, and especially Rich Gerber, were of great help in swiftly answering all of our technical questions, and occasionally even knowing more about our job than we did! We thank the invited speakers for agreeing to present at EACL, and the publication chairs, Yannick Parmentier and Gosse Bouma, for putting this volume together. The local organizing committee (Lars Borin, Aarne Ranta, Yvonne Adesam, Martin Kasá, and Nina Tahmasebi) have been invaluable in arranging the logistics of the conference and coordinating with us on many organizational issues, and we are grateful to the publicity chair, Peter Ljunglöf, for ensuring up-to-date programme information on the conference web site. We thank also the Student Research Workshop chairs for smoothly coordinating with us on their schedule. Last but not least, we are indebted to the General Chair, Shuly Wintner, for his guidance and support throughout the whole process.

We hope you enjoy the conference!

Sharon Goldwater and Stefan Riezler
EACL 2014 Programme Chairs

General Chair:

Shuly Wintner, University of Haifa (Israel)

Program Chairs:

Sharon Goldwater, University of Edinburgh (UK)

Stefan Riezler, Heidelberg University (Germany)

Local Organizing Committee:

Lars Borin (chair), University of Gothenburg (Sweden)

Aarne Ranta (chair), University of Gothenburg and Chalmers University of Technology (Sweden)

Yvonne Adesam, University of Gothenburg (Sweden)

Martin Kasá, University of Gothenburg (Sweden)

Nina Tahmasebi, Chalmers University of Technology (Sweden)

Publication Chairs:

Gosse Bouma, University of Groningen (The Netherlands)

Yannick Parmentier, University of Orléans (France)

Workshop Chairs:

Anja Belz, University of Brighton (UK)

Reut Tsarfaty, Uppsala University (Sweden)

Tutorial Chairs:

Afra Alishahi, Tilburg University (The Netherlands)

Marco Baroni, University of Trento (Italy)

Demo Chair:

Marko Tadić, University of Zagreb (Croatia)

Bogdan Babych, University of Leeds (UK)

Student Research Workshop Chairs:

Desmond Elliott, University of Edinburgh (UK)

Konstantina Garoufi, University of Potsdam (Germany)

Douwe Kiela, University of Cambridge (UK)

Ivan Vulić, KU Leuven (Belgium)

Student Research Workshop Faculty advisor:

Sebastian Padó, Heidelberg University (Germany)

Sponsorship Chairs:

Jochen Leidner, Thomson-Reuters/Linguit Ltd. (Switzerland)

Alessandro Moschitti, University of Trento (Italy)
Sofie Johansson Kokkinakis, University of Gothenburg (Sweden)
Staffan Larsson, University of Gothenburg (Sweden)

Publicity Chair:

Peter Ljunglöf, University of Gothenburg and Chalmers University of Technology (Sweden)

Area Chairs:

Enrique Alfonseca, John Blitzer, Aoife Cahill, Vera Demberg, Chris Dyer, Jacob Eisenstein, Micha Elsner, Katrin Erk, Afsaneh Fazly, Katja Filippova, Alexander Fraser, Iryna Gurevych, Chin-Yew Lin, David McClosky, Yusuke Miyao, Hwee Tou Ng, Slav Petrov, Simone Paolo Ponzetto, Sebastian Riedel, Verena Rieser, Helmut Schmid, Izhak Shafran, Hiroya Takamura, Lucy Vanderwende

Reviewers:

Fadi Abu-Sheika, Meni Adler, Nitish Aggarwal, Lars Ahrenberg, Afra Alishahi, Yaser Al-Onaizan, Yasemin Altun, Waleed Ammar, Jacob Andreas, Ion Androutsopoulos, Gabor Angeli, Mihael Arcan, Yoav Artzi, Jordi Atserias Batalla, Michael Auli, Harald Baayen, Timothy Baldwin, David Bamman, Mohit Bansal, Marco Baroni, Loïc Barrault, Núria Bel, Kedar Bellare, Islam Beltagy, Luciana Benotti, Yinon Bentor, Jonathan Berant, Sabine Bergler, Raffaella Bernardi, Clinton Bicknell, Chris Biemann, Arianna Bisazza, Yonatan Bisk, Roi Blanco, Michael Bloodgood, Phil Blunsom, Nathan Bodenstein, Branimir Boguraev, Bernd Bohnet, Gemma Boleda, Danushka Bollegala, Francis Bond, Kalina Bontcheva, Johan Bos, Houda Bouamor, Thorsten Brants, Chloé Braud, Fabienne Braune, Chris Brew, Ted Briscoe, Julian Brooke, Marco Brunello, Paul Buitelaar, Harry Bunt, Aljoscha Burchardt, David Burkett, Stephan Busemann, Bill Byrne, Nicoletta Calzolari, Ivan Cantador, Yunbo Cao, Giuseppe Carenini, Marine Carpuat, Xavier Carreras, John Carroll, Dave Carter, Francisco Casacuberta, Pablo Castells, Nathanael Chambers, Jason Chang, Ming-Wei Chang, David Chen, Hsin-Hsi Chen, Wenliang Chen, Chen Chen, Kehai Chen, Colin Cherry, Jackie Chi Kit Cheung, David Chiang, Christian Chiarcos, Kostadin Cholakov, Christos Christodoulopoulos, Jennifer Chu-Carroll, Cindy Chung, Massimiliano Ciaramita, Philipp Cimiano, Stephen Clark, Shay Cohen, Bonaventura Coppola, Marta R. Costa-jussà, Danilo Croce, Heriberto Cuayahuitl, Walter Daelemans, Cristian Danescu-Niculescu-Mizil, Dipanjan Das, Brian Davis, Munmun De Choudhury, Marie-Catherine de Marneffe, Gerard de Melo, Thierry Declerck, Michael Deisher, Steve DeNeefe, John DeNero, Pascal Denis, Michael Denkowski, Leon Derczynski, Marilena di Bari, Barbara Di Eugenio, Alberto Diaz, Michelangelo Diligenti, Markus Dreyer, Gregory Druck, Jinhua Du, Xiangyu Duan, Kevin Duh, Ewan Dunbar, Nadir Durrani, Marc Dymetman, Judith Eckle-Kohler, Koji Eguchi, Vladimir Eidelman, Andreas Eisele, David Elson, Angela Fahrni, James Fan, Richárd Farkas, Manaal Faruqui, Miriam Fernandez, Raquel Fernandez, Oliver Ferschke, João Filgueiras, Mark Fishel, Jeffrey Flanigan, Radu Florian, Mikel Forcada, Karën Fort, Eric Fosler-Lussier, Victoria Fossom, Jennifer Foster, Gil Francopoulo, Stefan L. Frank, Stella Frank, Francesca Frontini, Alona Fyshe, Michel Galley, Juri Ganitkevitch, Wenxuan Gao, Claire Gardent, Dan Garrette, Guillermo Garrido, Albert Gatt, Georgi Georgiev, Andrea Gesmundo, Arnab Ghoshal, George Giannakopoulos, Daniel Gildea, Kevin Gimpel, Jonathan Ginzburg, Yoav Goldberg, Julio Gonzalo, Spence Green, Edward Grefenstette, Camille Guinaudeau, Sonal Gupta, Francisco Guzman, Nizar Habash, Barry Haddow, John Hale, David Hall, Keith Hall, Greg Hanneman, Sanda Harabagiu, Christian Hardmeier, Matthias Hartung, mohammed hasanuzzaman, Katsuhiko Hayashi, Zhongjun He, Michael Heilman, James Henderson, John Henderson, Aurélie Herbelot, Ulf Hermjakob, Raquel Hervas, Graeme Hirst, Hieu Hoang, Johannes Hoffart, Mark Hopkins, Veronique Hoste, Fei Huang, Xiaojiang Huang, Xuanjing Huang, Rebecca Hwa, Nancy Ide, Gonzalo Iglesias, Diana Inkpen, Ann Irvine, Jagadeesh Jagarlamudi, Srinivasan

Janarthanam, Lifeng Jia, Richard Johansson, Doug Jones, Laura Kallmeyer, Jaap Kamps, Evangelos Kanoulas, Damianos Karakos, Graham Katz, Simon Keizer, Frank Keller, Shahram Khadivi, Adam Kilgarriff, Jin-Dong Kim, Seungyeon Kim, Katrin Kirchhoff, Philipp Koehn, Alexander Koller, Terry Koo, Anna Korhonen, Zornitsa Kozareva, Emiel Krahmer, Marco Kuhlmann, Roland Kuhn, Shankar Kumar, Jonathan Kummerfeld, Patrik Lambert, Phillippe Langlais, Guy Lapalme, Egoitz Laparra, Mirella Lapata, Staffan Larsson, Thomas Lavergne, Alon Lavie, Florian Laws, Lillian Lee, Junhui Li, lishuang li, Zhenghua Li, Maria Liakata, Chu-Cheng Lin, Krister Linden, Xiao Ling, Bing Liu, Jing Liu, Qun Liu, Yang Liu, Karen Livescu, Peter Ljunglöf, Elena Lloret, Adam Lopez, Annie Louis, Wei Lu, Yanjun Ma, Ji Ma, Klaus Macherey, Wolfgang Macherey, Bernardo Magnini, Inderjeet Mani, Chris Manning, Daniel Marcu, José B. Mariño, André F. T. Martins, Yuval Marton, Rebecca Mason, Yuji Matsumoto, Takuya Matsuzaki, Cettolo Mauro, Arne Mauser, Chandler May, Diana McCarthy, Ryan McDonald, Bob McMurray, Yashar Mehdad, Edgar Meij, Arul Menezes, Florian Metzger, Christian M. Meyer, Jeffrey Micher, Bonan Min, Margaret Mitchell, Behrang Mohit, Karo Moilanen, Monica Monachini, Christof Monz, Raymond Mooney, Andrea Moro, Alessandro Moschitti, Thomas Mueller, Smaranda Muresan, Brian Murphy, Seung-Hoon Na, Tetsuji Nakagawa, Toshiaki Nakazawa, Preslav Nakov, Ramesh Nallapati, Vivi Nastase, Tetsuya Nasukawa, Roberto Navigli, Mark-Jan Nederhof, Sapna Negi, Matteo Negri, Ani Nenkova, Graham Neubig, Vincent Ng, Jian-Yun Nie, Jan Niehues, Joakim Nivre, Brendan O'Connor, Stephan Oepen, Kemal Oflazer, Naoaki Okazaki, Gozde Ozbal, Sebastian Padó, Martha Palmer, Patrick Pantel, Cecile Paris, Christopher Parisien, Rebecca J. Passonneau, Alexandre Passos, Siddharth Patwardhan, Michael Paul, Michael J. Paul, Adam Pauls, Sasa Petrovic, Daniele Pighin, Andrei Popescu-Belis, Maja Popović, Fred Popowich, Matt Post, Sameer Pradhan, John Prager, Stephen Pulman, Matthew Purver, Sampo Pyysalo, Behrang Qasemzadeh, Ariadna Quattoni, Chris Quirk, Altaf Rahman, Owen Rambow, Ari Rappoport, Sujith Ravi, Alexis Raykhel, Michaela Regneri, Roi Reichart, Ehud Reiter, Jason Riesa, German Rigau, Alan Ritter, Stephen Roller, Laurent Romary, Carolyn Rose, Michael Roth, Dana Rubinstein, Rachel Rudinger, Vasile Rus, Alexander M. Rush, Graham Russell, Delia Rusu, Kenji Sagae, Horacio Saggion, Kazi Saidul Hasan, Hassan Sajjad, Mark Sammons, Baskaran Sankaran, Felix Sasaki, Giorgio Satta, Hassan Sawaf, David Schlangen, Nathan Schneider, Björn Schuller, Sabine Schulte im Walde, Yohei Seki, Hendra Setiawan, Aliaksei Severyn, Serge Sharoff, Libin Shen, Shuming Shi, Hiroyuki Shindo, Ekaterina Shutova, Advait Siddharthan, Carina Silberer, Mario J. Silva, Khalil Sima'an, Michel Simard, Kiril Simov, Serra Sinem Tekiroglu, Sameer Singh, Olivier Siohan, Gabriel Skantze, Nathaniel Smith, Stephen Soderland, Anders Søgaard, Tamar Solorio, Hagen Soltau, Swapna Somasundaran, Lucia Specia, Valentin Spitzkovsky, Caroline Sporleder, Edward Stabler, Mark Steedman, Josef Steinberger, Georg Stemmer, Amanda Stent, Mark Stevenson, Matthew Stone, Veselin Stoyanov, Carlo Strapparava, Michael Strube, Sara Stymne, Keh-Yih Su, Katsuhito Sudoh, Weiwei Sun, Mihai Surdeanu, Jun Suzuki, Mary Swift, Stan Szpakowicz, Whitney Tabor, Partha Pratim Talukdar, Joel Tetreault, Simone Teufel, Stefan Thater, Mariët Theune, Blaise Thomson, Jörg Tiedemann, Christoph Tillmann, Kristina Toutanova, David Traum, Ming-Feng Tsai, Richard Tzong-Han Tsai, Ioannis Tsochantaridis, Yulia Tsvetkov, Dan Tufiş, Masao Utiyama, Tim Van de Cruys, Antal van den Bosch, Benjamin Van Durme, Josef van Genabith, Paola Velardi, David Vilar, Andreas Vlachos, Stephan Vogel, Clare Voss, Stephen Wan, Haifeng Wang, Kai Wang, ling wang, Wen Wang, Pidong Wang, Yu-Chun Wang, Taro Watanabe, Bonnie Webber, Jason Williams, Philip Williams, Colin Wilson, Travis Wolfe, Dekai Wu, Sander Wubben, Fei Xia, Deyi Xiong, Deyi Xiong, Peng Xu, Bishan Yang, Hui Yang, Muyun Yang, tae yano, Limin Yao, Dani Yogatama, François Yvon, Beñat Zepirain, Richard Zens, Torsten Zesch, Luke Zettlemoyer, Feifei Zhai, Hui Zhang, Joy Ying Zhang, Lei Zhang, Min Zhang, Yi Zhang, Yue Zhang, Meng Zhang, Liu Zhanyi, Shiqi Zhao, Tiejun Zhao, Xin Zhao, Xin Zhao, Muhua Zhu, Chengqing Zong

Table of Contents

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

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

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

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

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

Conference Program

Monday, April 28

(8:30-9:00) Opening Session

(9:00-10:15) Invited Talk: Dan Roth

Session L1A: (10:45-12:25) Machine Translation

Improving Word Alignment Using Linguistic Code Switching Data

Fei Huang and Alexander Yates

Undirected Machine Translation with Discriminative Reinforcement Learning

Andrea Gesmundo and James Henderson

Minimum Translation Modeling with Recurrent Neural Networks

Yuening Hu, Michael Auli, Qin Gao and Jianfeng Gao

Maximizing Component Quality in Bilingual Word-Aligned Segmentations

Spyros Martzoukos, Christof Monz and Christophe Costa Florencio

Session L1B: (10:45-12:25) Semantics and Discourse

A Joint Model for Quotation Attribution and Coreference Resolution

Mariana S. C. Almeida, Miguel B. Almeida and André F. T. Martins

A Hierarchical Bayesian Model for Unsupervised Induction of Script Knowledge

Lea Frermann, Ivan Titov and Manfred Pinkal

Inducing Example-based Semantic Frames from a Massive Amount of Verb Uses

Daisuke Kawahara, Daniel Peterson, Octavian Popescu and Martha Palmer

Automated Verb Sense Labelling Based on Linked Lexical Resources

Kostadin Cholakov, Judith Eckle-Kohler and Iryna Gurevych

Monday, April 28 (continued)

Session L1C: (10:45-12:25) Social Media and Sentiment

Multi-Granular Aspect Aggregation in Aspect-Based Sentiment Analysis

John Pavlopoulos and Ion Androutsopoulos

Simple, Robust and (almost) Unsupervised Generation of Polarity Lexicons for Multiple Languages

Iñaki San Vicente, Rodrigo Agerri and German Rigau

Mapping Dialectal Variation by Querying Social Media

Gabriel Doyle

Modeling the Use of Graffiti Style Features to Signal Social Relations within a Multi-Domain Learning Paradigm

Mario Piergallini, Seza Dogruoz, Phani Gadde, David Adamson and Carolyn Rose

Session L1D: (10:45-12:25) Syntax and Parsing

Modelling the Lexicon in Unsupervised Part of Speech Induction

Gregory Dubbin and Phil Blunsom

Generalizing a Strongly Lexicalized Parser using Unlabeled Data

Tejaswini Deoskar, Christos Christodoulopoulos, Alexandra Birch and Mark Steedman

Special Techniques for Constituent Parsing of Morphologically Rich Languages

Zsolt Szántó and Richárd Farkas

Leveraging Verb-Argument Structures to Infer Semantic Relations

Eduardo Blanco and Dan Moldovan

Monday, April 28 (continued)

Session S1A: (14:00-15:00) see SP schedule

Session S1B: (14:00-15:00) see SP schedule

Session S1C: (14:00-15:00) see SP schedule

Session S1D: (14:00-15:00) see SP schedule

(15:15-18:00) Posters

LP Posters

Structured and Unstructured Cache Models for SMT Domain Adaptation

Annie Louis and Bonnie Webber

Regularized Structured Perceptron: A Case Study on Chinese Word Segmentation, POS Tagging and Parsing

Kaixu Zhang, Jinsong Su and Changle Zhou

About Inferences in a Crowdsourced Lexical-Semantic Network

Mathieu Lafourcade, Manel Zarrouk and Alain Joubert

Incremental Query Generation

Laura Perez-Beltrachini, Claire Gardent and Enrico Franconi

PARADIGM: Paraphrase Diagnostics through Grammar Matching

Jonathan Weese, Juri Ganitkevitch and Chris Callison-Burch

Translation memory retrieval methods

Michael Bloodgood and Benjamin Strauss

Frame Semantic Tree Kernels for Social Network Extraction from Text

Apoorv Agarwal, Sriramkumar Balasubramanian, Anup Kotalwar, Jiehan Zheng and Owen Rambow

Statistical Script Learning with Multi-Argument Events

Karl Pichotta and Raymond Mooney

Monday, April 28 (continued)

Improving Distributional Semantic Vectors through Context Selection and Normalisation

Tamara Polajnar and Stephen Clark

Source-side Preordering for Translation using Logistic Regression and Depth-first Branch-and-Bound Search

Laura Jehl, Adrià de Gispert, Mark Hopkins and Bill Byrne

Incremental Bayesian Learning of Semantic Categories

Lea Frermann and Mirella Lapata

Word Ordering with Phrase-Based Grammars

Adrià de Gispert, Marcus Tomalin and Bill Byrne

Iterative Constrained Clustering for Subjectivity Word Sense Disambiguation

Cem Akkaya, Janyce Wiebe and Rada Mihalcea

Identifying fake Amazon reviews as learning from crowds

Tommaso Fornaciari and Massimo Poesio

Assessing the relative reading level of sentence pairs for text simplification

Sowmya Vajjala and Detmar Meurers

Subcategorisation Acquisition from Raw Text for a Free Word-Order Language

Will Roberts, Markus Egg and Valia Kordoni

Classifying Temporal Relations with Simple Features

Paramita Mirza and Sara Tonelli

Using idiolects and sociolects to improve word prediction

Wessel Stoop and Antal van den Bosch

Dynamic Topic Adaptation for Phrase-based MT

Eva Hasler, Phil Blunsom, Philipp Koehn and Barry Haddow

Deterministic Parsing using PCFGs

Mark-Jan Nederhof and Martin McCaffery

Monday, April 28 (continued)

Empirically-motivated Generalizations of CCG Semantic Parsing Learning Algorithms

Jesse Glass and Alexander Yates

Correcting Grammatical Verb Errors

Alla Rozovskaya, Dan Roth and Vivek Srikumar

Fast Statistical Parsing with Parallel Multiple Context-Free Grammars

Krasimir Angelov and Peter Ljunglöf

Sentiment Propagation via Implicature Constraints

Lingjia Deng and Janyce Wiebe

Acquisition of Noncontiguous Class Attributes from Web Search Queries

Marius Pasca

Learning from Post-Editing: Online Model Adaptation for Statistical Machine Translation

Michael Denkowski, Chris Dyer and Alon Lavie

Predicting and Characterising User Impact on Twitter

Vasileios Lampos, Nikolaos Aletras, Daniel Preoŧiuc-Pietro and Trevor Cohn

A Knowledge-based Representation for Cross-Language Document Retrieval and Categorization

Marc Franco-Salvador, Paolo Rosso and Roberto Navigli

Dependency Tree Abstraction for Long-Distance Reordering in Statistical Machine Translation

Chenchen Ding and Yuki Arase

Improving the Lexical Function Composition Model with Pathwise Optimized Elastic-Net Regression

Jiming Li, Marco Baroni and Georgiana Dinu

Is Machine Translation Getting Better over Time?

Yvette Graham, Timothy Baldwin, Alistair Moffat and Justin Zobel

Monday, April 28 (continued)

SP Posters (see SP schedule)

Tuesday, April 29

(9:00-10:15) Invited Talk: Simon King

Session L2A: (10:45-12:25) Lexicon and Lexical Representation

Learning Dictionaries for Named Entity Recognition using Minimal Supervision
Arvind Neelakantan and Michael Collins

Improving Vector Space Word Representations Using Multilingual Correlation
Manaal Faruqui and Chris Dyer

Using Distributional Similarity of Multi-way Translations to Predict Multiword Expression Compositionality
Bahar Salehi, Paul Cook and Timothy Baldwin

Word Embeddings through Hellinger PCA
Rémi Lebret and Ronan Collobert

Session L2B: (10:45-12:25) Semantics and Discourse

A Latent Variable Model for Discourse-aware Concept and Entity Disambiguation
Angela Fahrni and Michael Strube

Topical PageRank: A Model of Scientific Expertise for Bibliographic Search
James Jardine and Simone Teufel

Distributional Lexical Entailment by Topic Coherence
Laura Rimell

Information Structure Prediction for Visual-world Referring Expressions
Micha Elsner, Hannah Rohde and Alasdair Clarke

Tuesday, April 29 (continued)

Session L2C: (10:45-12:25) Language Resources

Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality

Jey Han Lau, David Newman and Timothy Baldwin

What Substitutes Tell Us - Analysis of an "All-Words" Lexical Substitution Corpus

Gerhard Kremer, Katrin Erk, Sebastian Padó and Stefan Thater

Weighted Krippendorff's alpha is a more reliable metrics for multi-coders ordinal annotations: experimental studies on emotion, opinion and coreference annotation

Jean-Yves Antoine, Jeanne Villaneau and Anaïs Lefeuivre

Discriminating Rhetorical Analogies in Social Media

Christoph Lofi, Christian Nieke and Nigel Collier

Session L2D: (10:45-12:25) Morphology

Semi-supervised learning of morphological paradigms and lexicons

Mans Hulden, Markus Forsberg and Malin Ahlberg

How to Produce Unseen Teddy Bears: Improved Morphological Processing of Compounds in SMT

Fabienne Cap, Alexander Fraser, Marion Weller and Aoife Cahill

Type-Supervised Domain Adaptation for Joint Segmentation and POS-Tagging

Meishan Zhang, Yue Zhang, Wanxiang Che and Ting Liu

Session L2D: (10:45-12:25) TACL: Joint Morphological and Syntactic Analysis for Richly Inflected Languages. Bernd Bohnet, Joakim Nivre, Igor Boguslavsky, Richárd Farkas, Filip Ginter, Jan Hajič

Tuesday, April 29 (continued)

(14:00-15:00) Business Meeting

Session S2A: (15:00-16:00) see SP schedule

Session S2B: (15:00-16:00) see SP schedule

Session S2C: (15:00-16:00) see SP schedule

Session S2D: (15:00-16:00) see SP schedule

(16:30-18:00) Student Research Workshop

Wednesday, April 30

(9:00-10:15) Invited Talk: Ulrike v. Luxburg

Session L3A: (10:45-12:25) Machine Translation

Applying the semantics of negation to SMT through n-best list re-ranking
Federico Fancellu and Bonnie Webber

Bilingual Sentiment Consistency for Statistical Machine Translation
Boxing Chen and Xiaodan Zhu

Augmenting Translation Models with Simulated Acoustic Confusions for Improved Spoken Language Translation
Yulia Tsvetkov, Florian Metze and Chris Dyer

Wednesday, April 30 (continued)

Session L3A: (10:45-12:25) TACL: Automatic Detection and Language Identification of Multilingual Documents. Marco Lui, Jey Han Lau, and Timothy Baldwin

Session L3B: (10:45-12:25) Pragmatics/Discourse

A Generative Model for User Simulation in a Spatial Navigation Domain
Aciel Eshky, Ben Allison, Subramanian Ramamoorthy and Mark Steedman

Verbose, Laconic or Just Right: A Simple Computational Model of Content Appropriateness under Length Constraints
Annie Louis and Ani Nenkova

Discovering Implicit Discourse Relations Through Brown Cluster Pair Representation and Coreference Patterns
Attapol Rutherford and Nianwen Xue

“I Object!” Modeling Latent Pragmatic Effects in Courtroom Dialogues
Dan Goldwasser and Hal Daumé III

Session L3C: (10:45-12:25) Information Extraction and Applications

Encoding Semantic Resources in Syntactic Structures for Passage Reranking
Kateryna Tymoshenko, Alessandro Moschitti and Aliaksei Severyn

Automatic Food Categorization from Large Unlabeled Corpora and Its Impact on Relation Extraction
Michael Wiegand, Benjamin Roth and Dietrich Klakow

Redundancy Detection in ESL Writings
Huichao Xue and Rebecca Hwa

Fast Recursive Multi-class Classification of Pairs of Text Entities for Biomedical Event Extraction
Xiao Liu, Antoine Bordes and Yves Grandvalet

Wednesday, April 30 (continued)

Session L3D: (10:45-12:25) Generation and Summarization

Cluster-based Prediction of User Ratings for Stylistic Surface Realisation

Nina Dethlefs, Heriberto Cuayáhuil, Helen Hastie, Verena Rieser and Oliver Lemon

Improving the Estimation of Word Importance for News Multi-Document Summarization

Kai Hong and Ani Nenkova

Hybrid text simplification using synchronous dependency grammars with hand-written and automatically harvested rules

Advaith Siddharthan and Angrosh Mandya

A Summariser based on Human Memory Limitations and Lexical Competition

Yimai Fang and Simone Teufel

(14:00-15:00) Best paper awards

Best long paper

Learning part-of-speech taggers with inter-annotator agreement loss

Barbara Plank, Dirk Hovy and Anders Søgaard

Best short paper (see SP schedule)