

Coling 2010

**23rd International Conference on
Computational Linguistics**

**Proceedings of the Conference
Volume 1**

Chu-Ren Huang and Dan Jurafsky

23 – 27 August 2010
Beijing International Convention Center
Beijing, China

Published by
Tsinghua University Press
All rights reserved for hard copy production.
Block A, Xue Yan Building, Tsinghua University
Beijing, 100084
China

To order hard copies of this proceedings, contact:

Chinese Information Processing Society of China
No.4, Southern Fourth Street
Haidian District, Beijing, 100190
China
Tel: +86-010-62562916
Fax: +86-010-62562916
cips@iscas.ac.cn

Sponsorship

The COLING 2010 very gratefully acknowledges the following commitments in sponsorship:

Platinum Sponsors

-National Natural Science Foundation of China

-Department of Language Information Administration, Ministry of Education, PRC

Gold Sponsor



BaiDu

Silver Sponsors



Google



富士通 网络世界创意无限

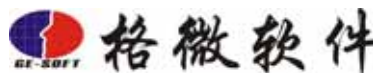
Fujitsu R&D Center CO., LTD.



Microsoft Research



Beijing TRS Information Technology Co., Ltd



Shenyang Globla Envoy software Co.,Ltd.

Supporters



Asian Federation of Natural Language Processing



Institute of Automation
Chinese Academy of Sciences



中国科学院计算技术研究所

Institute of Computing Technology
Chinese Academy of Sciences



中国科学院软件研究所

Institute of Software
Chinese Academy of Sciences



Harbin Institute of Technology



Peking University



Tsinghua University

Preface

You will find in this volume papers from the 23rd International Conference on Computational Linguistics (COLING 2010) held in Beijing, China on August 23-27, 2010 under the auspices of the International Committee on Computational Linguistics (ICCL), and organized by the Chinese Information Processing Society (CIPS) of China. For this prestigious natural language processing conference to be held in China is a significant event for computational linguistics and for colleagues in China, demonstrating both the maturity of our field and the development of academic areas in China.

COLING started as a friendly gathering in New York in 1965, and has grown steadily since. Yet COLING's aspiration to be a different conference remains the same. COLING strives to maintain its key qualities of embracing different theories and encouraging young scholars in spite of its growing size. A new component introduced at COLING 2010 underlines this quality. A RefreshINGenious (RING) session, organized by Aravind Joshi, our General Chair, allows new and un-orthodox ideas to be presented before they are fully developed in order to generate more discussion and stimulate other new ideas. We hope that this can become an important feature of COLING in the future.

The 155 oral papers included in the hardcopy proceedings published by Tsinghua University Press, as well as the 334 papers included in the electronic proceedings (the same 155 oral papers plus 179 poster papers) are selected from among 815 effective submissions among the more than 840 submissions received. The very selective acceptance rate of 19.02% for oral presentations (155/815 submissions) indicates the extremely high quality of the papers. An additional 21.96% (179/815) are selected for poster presentations to bring the overall acceptance rate to 40.98% (334/815).

We would like to thank the program committee area chairs for their dedicated and efficient review work, and our 738 reviewers for giving us very high quality reviews with a very short turnaround time, allowing us to maintain both the review quality and schedule even given the extraordinary number of submissions. Of course we thank the authors of the 840 papers for submitting their labor of love to COLING. Although we were only able to accept a minority of the submitted papers, we do hope that all authors and reviewers benefit from this process of indirect dialogue. We are especially grateful to the incredibly hard-working team of Stanford volunteers Jenny Finkel, Adam Vogel, and Mengqiu Wang, and HIT volunteers Sam Liang and Lemon Liu, who provided timely and efficient support for the two program chairs at every step of the review and publication processes.

Last but not least, we would like to thank the people who made COLING 2010 and this volume possible. We thank local arrangement committee co-chairs Professor Chengqing Zong and Professor Le Sun for their tireless work which will make COLING-2010 a sure success. Our special appreciation goes to the Chinese Information Processing Society (CIPS) and Professor Youqi Cao for their generous support as the COLING 2010 organizer. Lastly, Professor Qin Lu and Professor Tiejun Zhao should be recognized for their meticulous preparation for editing and publication, which brought this volume to reality.

Chu-Ren Huang and Dan Jurafsky,
COLING 2010 Program Committee Co-chairs

July 8, 2010

COLING 2010 is organized by the Chinese Information Processing Society of China (CIPS) and under the auspices of the International Committee on Computational Linguistics (ICCL).

The logo for the International Committee on Computational Linguistics (ICCL) consists of the letters "ICCL" in a bold, blue, sans-serif font.

General Chair:

Aravind K. Joshi (University of Pennsylvania)

Program Chairs:

Chu-Ren Huang (The Hong Kong Polytechnic University)

Dan Jurafsky (Stanford University)

Advisors to Organizing Committee:

Youqi Cao (The Chinese Information Processing Society of China)

Zhendong Dong (The Chinese Information Processing Society of China)

Changning Huang (Microsoft Research Asia)

Sheng Li (Harbin Institute of Technology)

Tianshun Yao (Northeastern University)

Shiwen Yu (Peking University)

Zhiwei Feng (Institute of Applied Linguistics, Ministry of Education)

Kaiying Liu (Shanxi University)

Organization Chairs:

Chengqing Zong (Institute of Automation, Chinese Academy of Sciences)

Le Sun (Institute of Software, Chinese Academy of Sciences)

Publication Chairs:

Qin Lu (The Hong Kong Polytechnic University)

Tiejun Zhao (Harbin Institute of Technology)

Tutorial Chairs:

Dan Gildea (University of Rochester)

Xuanjing Huang (Fudan University)

Workshop Chairs:

Noah Smith (Carnegie Mellon University)

Takenobu Tokunaga (Tokyo Institute of Technology)

Haifeng Wang (Baidu)

Publicity Chairs:

Hal Daumé III (University of Utah)

Bin Wang (Institute of Computing Technology, Chinese Academy of Sciences)

Minghui Dong (Institute for Infocomm Research)

Monica Monachini (Institute for Computational Linguistics)

Aline Villavicencio (Federal University of Rio Grande do Sul)

Sponsorship Chairs:

Tingting He (Huazhong Normal University)

Hinrich Schütze (University of Stuttgart)

Key-Sun Choi (Korea Advanced Institute of Science and Technology)

Rion Snow (Stanford University)

Takehito Utsuro (University of Tsukuba)

Renata Vieira (Pontifical Catholic University of Rio Grande do Sul)

Hao Yu (Fujitsu(China) R&D Center)

Demo Chairs:

Ting Liu (Harbin Institute of Technology)

Yang Liu (The University of Texas at Dallas)

Program Committee Members/AreaChairs:

Nianwen Xue (Brandeis University)

Rajeev Sangal (India Institute of Information Technology)

Roger Levy (University of California, San Diego)

Justine Cassell (Northwestern University)

Caroline Sporleder (Saarland University)

Gary Geunbae Lee (Pohang University of Science and Technology)

Rosie Jones (Yahoo! Research)

Nicoletta Calzolari (Istituto di Linguistica Computazionale)

Chris Callison-Burch (Johns Hopkins University)

Qun Liu (Institute of Computing Technology)

Pierre Isabelle (National Research Council of Canada)

Sadao Kurohashi (Kyoto University)

Sebastian Padó (Universität Stuttgart)

Takenobu Tokunaga (Tokyo Institute of Technology)

Bo Pang (Yahoo! Research)

Haizhou Li (Institute for Infocomm Research)

David A Smith (University of Massachusetts, Amherst)

Donia Scott (University of Sussex)

Emily Bender (University of Washington)

Ming Zhou (Microsoft Research Asia)

Organization Committee Members:

Juanzi Li (Tsinghua University)

KWONG Olivia (Hong Kong City University)

Bin Sun (Peking University)

Houfeng Wang (Peking University)

Xiaojie Wang (Beijing University of Posts and Telecommunications)

Endong Xun (Beijing Language and Culture University)

Erhong Yang (Beijing Language and Culture University)

Jun Zhao (Institute of Automation, Chinese Academy of Sciences)

Jingbo Zhu (Northeastern University)

Program Committee Members/Reviewers:

Rob Abbott	Jordan Boyd-Graber	Yejin Choi
Takeshi Abekawa	S. R. K. Branavan	Kenneth Church
Omri Abend	Antonio Branco	Massimiliano Ciaramita
Steven Abney	Eric Breck	Philipp Cimiano
Eytan Adar	Chris Brew	Shay Cohen
Guadalupe Aguado	Sabine Buchholz	Kevyn Collins-Thompson
Khalid Al-Kofahi	Stefan Buettcher	John Conroy
Cecilia Ovesdotter Alm	Hung Bui	Bonaventura Coppola
Omar Alonso	Razvan Bunescu	Ed Cormany
Bharat Ram Ambati	Aljoscha Burchardt	Dan Cristea
Massih-Reza Amini	Stephan Busemann	Bruce Croft
Xiangdong An	Miriam Butt	Dick Crouch
Sophia Ananiadou	William Byrne	Montse Cuadros
Jaime Arguello	Lynne Cahill	Silviu-Petru Cucerzan
Masayuki Asahara	Jamie Callan	Oliver Čulo
Nicholas Asher	Bin Cao	Aron Culotta
Tania Avgustinova	Yunbo Cao	Beatrice Daille
Necip Fazil Ayan	Guiseppe Carenini	Hercules Dalianis
Lakshmi Bai	Jean Carletta	Sandipan Dandapat
Jason Baldridge	Marine Carpuat	Hoa Dang
Timothy Baldwin	Xavier Carreras	Dana Dannells
Carmen Banea	John Carroll	Dipanjan Das
Srinivas Bangalore	Ben Carterette	Sajib Dasgupta
Colin Bannard	Francisco Casacuberta	Hal Daume
Ken Barker	Steve Cassidy	Dmitry Davidov
Marco Baroni	Eric Castelli	Adria de Gispert
Regina Barzilay	Alexandru Ceausu	Eric De La Clergerie
John Bateman	Nick Cercone	Valeria de Paiva
Beata Beigman Klebanov	Jeong-Won Cha	Maarten de Rijke
Daisuke Bekki	Vineet Chaitanya	Thierry Declerck
Nria Bel	Yllias Chali	Vera Demberg
Yinon Bentor	Nate Chambers	Steve DeNeeffe
Sabine Bergler	Baobao Chang	John DeNero
Aditya Bhargava	Pi-Chuan Chang	Pascal Denis
Pushpak Bhattacharyya	Wanxiang Che	Mona Diab
Timothy Bickmore	Ciprian Chelba	Georgiana Dinu
Klinton Bicknell	Aitao Chen	Pinar Donmez
Alexandra Birch	Bin Chen	Iustin Dornescu
Philippe Blache	Boxing Chen	Ascander Dost
Alan Black	Chien Chin Chen	Antoine Doucet
John Blitzer	Harr Chen	Doug Downey
Michael Bloodgood	Hsin-Hsi Chen	Markus Dreyer
Phil Blunsom	Jinying Chen	Gregory Druck
Ondrej Bojar	Keh-Jiann Chen	Xiangyu Duan
Gemma Boleda	Wenliang Chen	Amit Dubey
Francis Bond	Xiao Chen	Kevin Duh
Johan Bos	Charibeth K. Cheng	Michael Dukes
Marisa Boston	Colin Cherry	Chris Dyer
Pierrette Bouillon	Tee Kiah Chia	Marc Dymetman
Julien Bourdaillet	David Chiang	Markus Egg

Koji Eguchi	Tianxia Gong	Juan Huete
David Eichmann	Julio Gonzalo	Sarmad Hussain
Andreas Eisele	Cyril Goutte	Rebecca Hwa
Jacob Eisenstein	Brigitte Grau	Gonzalo Iglesias
Jason Eisner	Stephan Greene	Steven Ikier
Noemie Elhadad	Mark Greenwood	Diana Inkpen
Charles Elkan	Gregory Grefenstette	Kentaro Inui
Micha Elsner	Ralph Grishman	Elena Irimia
Alexandre entiev	Cecile Grivaz	Amy Isard
Katrin Erk	Jiafeng Guo	Abe Ittycheriah
Andrea Esuli	Ben Hachey	Tatsuya Izuha
Richard Evans	Aria Haghighi	Adam Jatowt
Roger Evans	Udo Hahn	Jiwoon Jeon
Patrick Fan	Jan Hajic	Heng Ji
Alex Fang	Eva Hajicova	Sittichai Jiampojarn
Benoit Favre	Dilek Hakkani-tur	Daxin Jiang
Marcello Federico	John Hale	Hongfei Jiang
Christiane Fellbaum	David Hall	Jiepu Jiang
Katja Filippova	Greg Hanneman	Jing Jiang
Radu Florian	Max Harper	Long Jiang
Sandiway Fong	Chikara Hashimoto	Wenbin Jiang
George Foster	Ahmed Hassan	Valentin Jijkoun
Gil Francopoulo	Claudia Hauff	Richard Johansson
Stefan Frank	Yoshihiko Hayashi	Howard Johnson
Alex Fraser	Ben He	Rie Johnson
Marjorie Freedman	daan he	Kristiina Jokinen
Maria Fuentes	Daqing He	Doug Jones
Hagen Fuerstenau	Xiaodong He	Rosie Jones
Atsushi Fujii	Yulan He	Hanmin Jung
Sumio Fujita	John Henderson	Vijay K. Shanker
Pascale Fung	Iris Hendrickx	Min-Yen Kan
Ryan Gabbard	Amac Herdagdelen	Hiroshi Kanayama
Evgeniy Gabrilovich	Ulf Hermjakob	Damianos Karakos
Karthik Gali	Raquel Hervas	Nikiforos Karamanis
Michel Galley	Dirk Heylen	Rohit Kate
Michael Gamon	Andrew Hickl	Tsuneaki Kato
Kuzman Ganchev	Graeme Hirst	Daisuke Kawahara
Jianfeng Gao	Jerry Hobbs	Tatsuya Kawahara
Albert Gatt	Katja Hofmann	Junichi Kazama
Eric Gaussier	Steven Chu Hong Hoi	Shahram Khadivi
Ruifang Ge	Kristy Hollingshead	Chloe Kiddon
Byron Georgantopoulos	Mark Hopkins	Jin-Dong Kim
Pablo Gervas	Eric Horvitz	Min Kim
Sanjukta Ghosh	Veronique Hoste	Chunyu Kit
Daniel Gildea	Yunhua Hu	Kevin Knight
Dan Gillick	Jimmy Huang	Youngjoong Ko
Jesus Gimenez	XuanJing Huang	Philipp Koehn
Kevin Gimpel	Yifen Huang	Rob Koeling
Roxana Girju	Yun Huang	Dimitrios Kokkinakis
Sharon Goldwater	Matt Huenerfauth	Greg Kondrak

Moshe Koppel	Wen Jie Li	David McClosky
Valia Kordoni	Zhifei Li	Ryan McDonald
Lili Kotlerman	Dekang Lin	Tara McIntosh
Eric Kow	Lucian Vlad Lita	Kathy McKeown
Zornitsa Kozareva	Diane Litman	Susan McRoy
Emiel Kramer	Bing Liu	Yashar Mehdad
Steven Krauwer	Shui Liu	Qiaozhu Mei
Gerhard Kremer	Ting Liu	Chris Mellish
Canasai Kruengkrai	Xiaohua Liu	Amlia Mendes
Lun-Wei Ku	Yan Liu	Helen Meng
Taku Kudo	Yang Liu	Donald Metzler
Jonas Kuhn	Yang Liu	Haitao Mi
Roland Kuhn	Yupeng Liu	Jun Miao
Peter Khnlein	Zhanyi Liu	Lukas Michelbacher
Amba Kulkarni	Zhaopengx Liu	Jeffrey Micher
Ravi Kumar	Anna Lobanova	Eleni Miltsakaki
Shankar Kumar	Qiu Long	David Mimno
A Kumaran	Adam Lopez	Zhaoyan Ming
Oren Kurland	Qin Lu	Shachar Mirkin
K. L. Kwok	Yue Lu	Jeff Mitchell
Olivia Kwong	Harald Lngen	Vibhu Mittal
Wai Lam	Xiaoqiang Luo	Yusuke Miyao
Man Lan	Tomoya Lwakura	Marie-Francine Moens
Philippe Langlais	Bin Ma	Dan Moldovan
Francois Lareau	Qing Ma	Diego Moll Aliod
Martha Larson	Yanjun Ma	Christof Monz
Alex Lascarides	Wolfgang Macherey	Raymond J. Mooney
Alberto Lavelli	Nitin Madnani	Bob Moore
Alon Lavie	Hala Maghout	Roser Morante
Florian Laws	B. Mallikarjun	Louis-Philippe Morency
Guy Lebanon	Gideon Mann	Alessandro Moschitti
Changki Lee	Daniel Marcu	Isabelle Moulinier
John Lee	Mitch Marcus	Animesh Mukherjee
Joo Young Lee	Joseph Mariani	Stefan Miller
Lillian Lee	Katja Markert	Art Munson
Seungwoo Lee	Erwin Marsi	Dragos Munteanu
Tan Lee	M. Antonia Marti Antonin	Masaki Murata
Jochen Leidner	Jean-Claude MARTIN	Vanessa Murdock
Yves Lepage	Andre Martins	Smaranda Muresan
Kevin Lerman	Yuval Marton	Gabriel Murray
James Lester	Yuji Masumoto	Pradeep Muthukrishnan
Gregor Leusch	Shigeki Matsubara	Sobha Nair
Gina-Anne Levow	Tomoko Matsui	Testuji Nakagawa
Hang Li	Yuji Matsumoto	Vivi Nastase
Jiye Li	Takuya Matsuzaki	Martina Naughton
Lei Li	Evgeny Matusov	Roberto Navigli
Mu Li	Mausam	Mark-Jan Nederhof
Shiqi Li	Arne Mauser	Vasek Nemcik
Shoushan Li	Marshall Mayberry	Ani Nenkova
Wei Li	Diana McCarthy	Hwee Tou Ng

Vincent Ng	Laurent Prevot	Stephanie Seneff
Huyen Nguyen Thi Minh	Matthew Purver	Violeta Seretan
Patrick Nguyen	Haoliang Qi	Hendra Setiawan
Alex Niculescu-Mizil	Toh Zhi Qiang	Chirag Shah
Jian-Yun Nie	Tao Qin	Dipti Sharma
Rodney Nielsen	Yang Qu	Dou Shen
Takashi Ninomiya	Chris Quirk	Libin Shen
Toyoaki Nishida	Bhiksha Raj	Wade Shen
Cheng Niu	S. Rajendran	Kiyooki Shirai
Zheng-Yu Niu	Dan Ramage	Eyal Shnarch
Diarmuid Saghda	S.V. Ramanan	Advait Siddharthan
Franz Och	P. V. S. Rambabu	Candy Sidner
Michael O'Donnell	Owen Rambow	Khalil Simaan
Stephan Oepen	Delip Rao	Michel Simard
Naoki Okazaki	Ari Rappoport	Kiril Simov
Lilja Ovreliid	Paul Rayson	Anil Kumar Singh
Derya Ozkan	Michaela Regneri	Smriti Singh
Ulrike Pado	Nils Reiter	Samar Sinha
Alexis Palmer	Norbert Reithinger	Sharon Small
Sinno Jialin Pan	Giuseppe Riccardi	Jason Smith
Patrick Pantel	Sebastian Riedel	Nathaniel Smith
Simone Paolo Ponzetto	German Rigau	Matthew Snover
Jong Park	Hae-Chang Rim	Benjamin Snyder
Kristen Parton	Lucia Rino	Stephen Soderland
Marius Pasca	Hammam Riza	Swapna Somasundaran
Siddharth Patwardhan	Horacio Rodríguez	Yan Song
Michael Paul	Gerhard Rolletschek	Young-In Song
Soma Paul	Lorenza Romano	Virach Sornlertlamvanich
Matthias Paulik	Mathias Rossignol	Lucia Specia
Yves Peirsman	Antti-Veikko Rosti	Jennifer Spenader
Gerald Penn	Michael Roth	Valentin Spitzkovsky
Marco Pennacchiotti	Alex Rudnick	Richard Sproat
Bhaskararao Peri	Marta Ruiz	Ed Stabler
Aina Peris	Kenji Sagae	Manfred Stede
Wim Peters	Horacio Saggion	Mark Steedman
Kay Peterson	Harabagiu Sanda	Amanda Stent
Slav Petrov	Baskaran Sankaran	Mark Stevenson
Daniele Pighin	Ratna Sanyal	Matthew Stone
Prasad Pingali	Murat Saraclar	Svetlana Stoyanchev
Stelios Piperidis	Sudeshna Sarkar	Veselin Stoyanov
Guillaume Pitel	Manabu Sassano	Kristina Striegnitz
Paul Piwek	Sekine Satoshi	Michael Strube
Luiz Pizzato	Roser Sauri	Tomek Strzalkowski
Massimo Poesio	Helmut Schmid	Jian Su
Hoifung Poon	William Schuler	Jinsong Su
Ana-Maria Popescu	Sabine Schulte im Walde	Keh-Yih Su
Maja Popovic	Ineke Schuurman	Nam Kim Su
Christopher Potts	Lane Schwartz	Rajen Subba
Richard Power	Wolfgang Seeker	K. V. Subbarao
Rashmi Prasad	sekine Sekine	L. V. Subramaniam

Zhifang Sui	David Vilar	Bei Yu
Eiichiro Sumita	Aline Villavicencio	Jianxing Yu
Congkai Sun	Stephan Vogel	Kun Yu
Jun Sun	Clare Voss	Mo Yu
Maosong Sun	Piek Vossen	Yisong Yue
Hisami Suzuki	Marilyn Walker	Zdenek Zabokrtsky
Gyorgy Szarvas	Hanna Wallach	Taras Zagibalov
Hiroya Takamura	Xiaojun Wan	Omar Zaidan
Koichi Takeuchi	Wenting Wang	Roberto Zanolli
David Talbot	Haifeng Wang	Fabio Zanzotto
Chew Lim Tan	Hsin-Min Wang	Alessandra Zarcone
Jie Tang	Kai Wang	Richard Zens
Chua Tat-Seng	Rui Wang	Torsten Zesch
Paul Tepper	Wei Wang	Luke Zettlemyer
Simone Teufel	Xuanhui Wang	Hao Zhang
Stefan Thater	Leo Wanner	Hui Zhang
Mariet Theune	Nigel Ward	Jiajun Zhang
Paul Thomas	Taro Watanabe	Min Zhang
Vu Thuy	Yotaro Watanabe	Ruiqiang Zhang
Joerg Tiedemann	Nick Webb	Yujie Zhang
Christoph Tillmann	Bonnie Webber	Ziqi Zhang
Ivan Titov	Eric Wehrli	Hai Zhao
Katrin Tomanek	Kilian Weinberger	Jun Zhao
Yoichi Tomiura	David Weir	Bowen Zhou
Sara Tonelli	Michael White	Guodong Zhou
Kentaro Torisawa	Michael Wick	Yiping Zhou
Kristina Toutanova	Jan Wiebe	Conghui Zhu
Roy Tromble	Yorick Wilks	Jerry Zhu
Huihsin Tseng	Theresa Wilson	Tao Zhuang
Shu-Chuan Tseng	Shuly Wintner	Heike Zinsmeister
Benjamin Tsou	Andreas Witt	Imed Zitouni
Yuta Tsuboi	Peter Wittek	Andreas Zollmann
Jun-ichi Tsujii	Magda Wolska	Chengqing Zong
Koji Tsukamoto	Chung-Hsien Wu	Ingrid Zukerman
Yoshimasa Tsuruoka	Dan Wu	
Dan Tufis	Dekai Wu	
Gokhan Tur	Hua Wu	
Kiyotaka Uchimoto	Xiaoyun Wu	
Yuya Unno	Yunfang Wu	
Lonneke Van der Plas	Fei Xia	
ielka van der sluis	Rui Xia	
Josef van Genabith	Xinyan Xiao	
Gertjan van Noord	Deyi Xiong	
Lucy Vanderwende	Peng Xu	
Sebastian Varges	Ruifeng Xu	
Sriram Venkatapathy	Bert Xue	
Marc Verhagen	Yongxin Yan	
Yannick Versley	Muyun Yang	
Cristina Vertan	Qiang Yang	
Marta Vila	Ainur Yessenalina	

Table of Contents

<i>Testing SDRT's Right Frontier</i> Stergos Afantenos and Nicholas Asher	1
<i>Identifying Multi-word Expressions by Leveraging Morphological and Syntactic Idiosyncrasy</i> Hassan Al-Haj and Shuly Wintner	10
<i>Robust Measurement and Comparison of Context Similarity for Finding Translation Pairs</i> Daniel Andrade, Tetsuya Nasukawa and Junichi Tsujii	19
<i>Multilingual Subjectivity: Are More Languages Better?</i> Carmen Banea, Rada Mihalcea and Janyce Wiebe	28
<i>Plagiarism Detection across Distant Language Pairs</i> Alberto Barrón-Cedeño, Paolo Rosso, Eneko Agirre and Gorka Labaka	37
<i>Automatic Detection of Non-deverbal Event Nouns for Quick Lexicon Production</i> Nuria Bel, Maria Coll and Gabriela Resnik	46
<i>Fast and Accurate Arc Filtering for Dependency Parsing</i> Shane Bergsma and Colin Cherry	53
<i>A Hierarchical Classifier Applied to Multi-way Sentiment Detection</i> Adrian Bickerstaffe and Ingrid Zukerman	62
<i>Fluency Constraints for Minimum Bayes-Risk Decoding of Statistical Machine Translation Lattices</i> Graeme Blackwood, Adrià de Gispert and William Byrne	71
<i>Self-Annotation for fine-grained geospatial relation extraction</i> Andre Blessing and Hinrich Schütze	80
<i>Top Accuracy and Fast Dependency Parsing is not a Contradiction</i> Bernd Bohnet	89
<i>Broad Coverage Multilingual Deep Sentence Generation with a Stochastic Multi-Level Realizer</i> Bernd Bohnet, Leo Wanner, Simon Mill and Alicia Burga	98
<i>Towards an optimal weighting of context words based on distance</i> Bernard Brosseau-Villeneuve, Jian-Yun Nie and Noriko Kando	107
<i>Measuring the Non-compositionality of Multiword Expressions</i> Fan Bu, Xiaoyan Zhu and Ming Li	116
<i>A Utility-Driven Approach to Question Ranking in Social QA</i> Razvan Bunescu and Yunfeng Huang	125
<i>Simultaneous Ranking and Clustering of Sentences: A Reinforcement Approach to Multi-Document Summarization</i> Xiaoyan Cai, Wenjie Li, You Ouyang and Hong Yan	134

<i>End-to-End Coreference Resolution via Hypergraph Partitioning</i>	
Jie Cai and Michael Strube	143
<i>Exploiting Background Knowledge for Relation Extraction</i>	
Yee Seng Chan and Dan Roth	152
<i>Jointly Modeling WSD and SRL with Markov Logic</i>	
Wanxiang Che and Ting Liu	161
<i>Bipolar Person Name Identification of Topic Documents Using Principal Component Analysis</i>	
Chien Chin Chen and Chen-Yuan Wu	170
<i>Emotion Cause Detection with Linguistic Constructions</i>	
Ying Chen, Sophia Yat Mei Lee, Shoushan Li and Chu-Ren Huang	179
<i>A Twin-Candidate Based Approach for Event Pronoun Resolution using Composite Kernel</i>	
Bin Chen, Jian Su and Chew Lim Tan	188
<i>Unsupervised Synthesis of Multilingual Wikipedia Articles</i>	
Yuncong Chen and Pascale Fung	197
<i>Simplicity is Better: Revisiting Single Kernel PPI Extraction</i>	
Sung-Pil Choi and Sung-Hyon Myaeng	206
<i>An ontology-driven system for detecting global health events</i>	
Nigel Collier, Reiko Matsuda Goodwin, John McCrae, Son Doan, Ai Kawazoe, Mike Conway, Asanee Kawtrakul, Koichi Takeuchi and Dinh Dien	215
<i>Constraining robust constructions for broad-coverage parsing with precision grammars</i>	
Bart Cramer and Yi Zhang	223
<i>Local lexical adaptation in Machine Translation through triangulation: SMT helping SMT</i>	
Josep Maria Crego, Aurélien Max and François Yvon	232
<i>Automated Translation of Semantic Relationships</i>	
Dmitry Davidov and Ari Rappoport	241
<i>Comparison of different algebras for inducing the temporal structure of texts</i>	
Pascal Denis and Philippe Muller	250
<i>Generating Learner-Like Morphological Errors in Russian</i>	
Markus Dickinson	259
<i>Resolving Object and Attribute Coreference in Opinion Mining</i>	
Xiaowen Ding and Bing Liu	268
<i>Entity Disambiguation for Knowledge Base Population</i>	
Mark Dredze, Paul McNamee, Delip Rao, Adam Gerber and Tim Finin	277
<i>A Discriminative Latent Variable-Based "DE" Classifier for Chinese-English SMT</i>	
Jinhua Du and Andy Way	286

<i>An Empirical Study on Learning to Rank of Tweets</i> Yajuan Duan, Long Jiang, Tao Qin, Ming Zhou and Heung-Yeung Shum	295
<i>Translation Model Generalization using Probability Averaging for Machine Translation</i> Nan Duan, Hong Sun and Ming Zhou	304
<i>Mixture Model-based Minimum Bayes Risk Decoding using Multiple Machine Translation Systems</i> Nan Duan, Mu Li, Dongdong Zhang and Ming Zhou	313
<i>Multi-Sentence Compression: Finding Shortest Paths in Word Graphs</i> Katja Filippova	322
<i>Enriching Dictionaries with Images from the Internet - Targeting Wikipedia and a Japanese Semantic Lexicon: Lexeed -</i> Sanae Fujita and Masaaki Nagata	331
<i>Opinosis: A Graph Based Approach to Abstractive Summarization of Highly Redundant Opinions</i> Kavita Ganesan, ChengXiang Zhai and Jiawei Han	340
<i>EMDC: A Semi-supervised Approach for Word Alignment</i> Qin Gao, Francisco Guzman and Stephan Vogel	349
<i>A Large Scale Ranker-Based System for Search Query Spelling Correction</i> Jianfeng Gao, Xiaolong Li, Daniel Micol, Chris Quirk and Xu Sun	358
<i>RTG based surface realisation for TAG</i> Claire Gardent and Laura Perez-Beltrachini	367
<i>Automatically Learning Source-side Reordering Rules for Large Scale Machine Translation</i> Dmitriy Genzel	376
<i>Comparing Language Similarity across Genetic and Typologically-Based Groupings</i> Ryan Georgi, Fei Xia and William Lewis	385
<i>Better Arabic Parsing: Baselines, Evaluations, and Analysis</i> Spence Green and Christopher D. Manning	394
<i>Paraphrase Alignment for Synonym Evidence Discovery</i> Gintarè Grigonytė, João Paulo Cordeiro, Gaël Dias, Rumen Moraliyski and Pavel Brazdil ...	403
<i>Finding the Storyteller: Automatic Spoiler Tagging using Linguistic Cues</i> Sheng Guo and Naren Ramakrishnan	412
<i>Detection of Simple Plagiarism in Computer Science Papers</i> Yaakov HaCohen-Kerner, Aharon Tayeb and Natan Ben-Dror	421
<i>A Structured Vector Space Model for Hidden Attribute Meaning in Adjective-Noun Phrases</i> Matthias Hartung and Anette Frank	430
<i>Hierarchical Phrase-based Machine Translation with Word-based Reordering Model</i> Katsuhiko Hayashi, Hajime Tsukada, Katsuhito Sudoh, Kevin Duh and Seiichi Yamamoto ...	439

<i>A Novel Reordering Model Based on Multi-layer Phrase for Statistical Machine Translation</i>	
Yanqing He, Yu Zhou, Chengqing Zong and Huilin Wang	447
<i>Standardizing Wordnets in the ISO Standard LMF: Wordnet-LMF for GermaNet</i>	
Verena Henrich and Erhard Hinrichs	456
<i>Normal-form parsing for Combinatory Categorical Grammars with generalized composition and type-raising</i>	
Julia Hockenmaier and Yonatan Bisk	465
<i>An Empirical Study on Web Mining of Parallel Data</i>	
Gumwon Hong, Chi-Ho Li, Ming Zhou and Hae-Chang Rim	474
<i>Enhancing Cross Document Coreference of Web Documents with Context Similarity and Very Large Scale Text Categorization</i>	
Jian Huang, Pucktada Treeratpituk, Sarah Taylor and C. Lee Giles	483
<i>Feature-Rich Discriminative Phrase Rescoring for SMT</i>	
Fei Huang and Bing Xiang	492
<i>FactRank: Random Walks on a Web of Facts</i>	
Alpa Jain and Patrick Pantel	501
<i>Open Entity Extraction from Web Search Query Logs</i>	
Alpa Jain and Marco Pennacchiotti	510
<i>Reranking Models in Fine-grained Opinion Analysis</i>	
Richard Johansson and Alessandro Moschitti	519
<i>Unsupervised phonemic Chinese word segmentation using Adaptor Grammars</i>	
Mark Johnson and Katherine Demuth	528
<i>Data-Driven Parsing with Probabilistic Linear Context-Free Rewriting Systems</i>	
Laura Kallmeyer and Wolfgang Maier	537
<i>Learning to Predict Readability using Diverse Linguistic Features</i>	
Rohit Kate, Xiaoqiang Luo, Siddharth Patwardhan, Martin Franz, Radu Florian, Raymond Mooney, Salim Roukos and Chris Welty	546
<i>Value for Money: Balancing Annotation Effort, Lexicon Building and Accuracy for Multilingual WSD</i>	
Mitesh Khapra, Saurabh Sohoney, Anup Kulkarni and Pushpak Bhattacharyya	555
<i>A Cross-lingual Annotation Projection Approach for Relation Detection</i>	
Seokhwan Kim, Minwoo Jeong, Jonghoon Lee and Gary Geunbae Lee	564
<i>Evaluating N-gram based Evaluation Metrics for Automatic Keyphrase Extraction</i>	
Su Nam Kim, Timothy Baldwin and Min-Yen Kan	572
<i>Improving the Quality of Text Understanding by Delaying Ambiguity Resolution</i>	
Doo Soon Kim, Ken Barker and Bruce Porter	581

<i>Automatic generation of inter-passage links based on semantic similarity</i>	
Petr Knoth, Jakub Novotny and Zdenek Zdrahal	590
<i>Dependency-driven Anaphoricity Determination for Coreference Resolution</i>	
Fang Kong, Guodong Zhou, Longhua Qian and Qiaoming Zhu	599
<i>Phrase Clustering for Smoothing TM Probabilities - or, How to Extract Paraphrases from Phrase Tables</i>	
Roland Kuhn, Boxing Chen, George Foster and Evan Stratford	608
<i>Revisiting Context-based Projection Methods for Term-Translation Spotting in Comparable Corpora</i>	
Audrey Laroche and Philippe Langlais	617
<i>Constituent Reordering and Syntax Models for English-to-Japanese Statistical Machine Translation</i>	
Young-Suk Lee, Bing Zhao and Xiaoqian Luo	626
<i>Sentiment Classification and Polarity Shifting</i>	
Shoushan Li, Sophia Y. M. Lee, Ying Chen, Chu-Ren Huang and Guodong Zhou	635
<i>Improving Corpus Comparability for Bilingual Lexicon Extraction from Comparable Corpora</i>	
Bo Li and Eric Gaussier	644
<i>Structure-Aware Review Mining and Summarization</i>	
Fangtao Li, Chao Han, Minlie Huang, Xiaoyan Zhu, Ying-Ju Xia, Shu Zhang and Hao Yu ...	653
<i>Adaptive Development Data Selection for Log-linear Model in Statistical Machine Translation</i>	
Mu Li, Yinggong Zhao, Dongdong Zhang and Ming Zhou	662
<i>Learning the Scope of Negation via Shallow Semantic Parsing</i>	
Junhui Li, Guodong Zhou, Hongling Wang and Qiaoming Zhu	671
<i>Filtered Ranking for Bootstrapping in Event Extraction</i>	
Shasha Liao and Ralph Grishman	680
<i>Exploring variation across biomedical subdomains</i>	
Tom Lippincott, Diarmuid Ó Séaghdha, Lin Sun and Anna Korhonen	689
<i>Semantic Role Labeling for News Tweets</i>	
Xiaohua Liu, Kuan Li, Bo Han, Ming Zhou, Long Jiang, Zhongyang Xiong and Changning Huang	698
<i>Joint Parsing and Translation</i>	
Yang Liu and Qun Liu	707
<i>Semantic Role Features for Machine Translation</i>	
Ding Liu and Daniel Gildea	716
<i>TimeML Events Recognition and Classification: Learning CRF Models with Semantic Roles</i>	
Hector Llorens, Estela Saquete and Borja Navarro-Colorado	725
<i>Exploiting Structured Ontology to Organize Scattered Online Opinions</i>	
Yue Lu, Huizhong Duan, Hongning Wang and ChengXiang Zhai	734

<i>Enhancing Morphological Alignment for Translating Highly Inflected Languages</i>	
Minh-Thang Luong and Min-Yen Kan	743
<i>Automatic analysis of semantic similarity in comparable text through syntactic tree matching</i>	
Erwin Marsi and Emiel Kraemer	752
<i>Controlling Listening-oriented Dialogue using Partially Observable Markov Decision Processes</i>	
Toyomi Meguro, Ryuichiro Higashinaka, Yasuhiro Minami and Kohji Dohsaka	761
<i>Recognising Entailment within Discourse</i>	
Shachar Mirkin, Jonathan Berant, Ido Dagan and Eyal Shnarch	770
<i>Evaluating Dependency Representations for Event Extraction</i>	
Makoto Miwa, Sampo Pyysalo, Tadayoshi Hara and Jun'ichi Tsujii	779
<i>Entity-Focused Sentence Simplification for Relation Extraction</i>	
Makoto Miwa, Rune Sætre, Yusuke Miyao and Jun'ichi Tsujii	788
<i>Using Cross-Lingual Projections to Generate Semantic Role Labeled Annotated Corpus for Urdu - A Resource Poor Language</i>	
Smruthi Mukund, Debanjan Ghosh and Rohini Srihari	797
<i>Recognition of Affect, Judgment, and Appreciation in Text</i>	
Alena Neviarouskaya, Helmut Prendinger and Mitsuru Ishizuka	806
<i>Nonparametric Word Segmentation for Machine Translation</i>	
ThuyLinh Nguyen, Stephan Vogel and Noah A. Smith	815
<i>Automatic Discovery of Feature Sets for Dependency Parsing</i>	
Peter Nilsson and Pierre Nugues	824
<i>Evaluation of Dependency Parsers on Unbounded Dependencies</i>	
Joakim Nivre, Laura Rimell, Ryan McDonald and Carlos Gómez Rodríguez	833
<i>Co-STAR: A Co-training Style Algorithm for Hyponymy Relation Acquisition from Structured and Unstructured Text</i>	
Jong-Hoon Oh, Ichiro Yamada, Kentaro Torisawa and Stijn De Saeger	842
<i>Simple and Efficient Algorithm for Approximate Dictionary Matching</i>	
Naoaki Okazaki and Jun'ichi Tsujii	851
<i>Latent Mixture of Discriminative Experts for Multimodal Prediction Modeling</i>	
Derya Ozkan, Kenji Sagae and Louis-Philippe Morency	860
<i>Text Summarization of Turkish Texts using Latent Semantic Analysis</i>	
Makbule Ozsoy, Ilyas Cicekli and Ferda Alpaslan	869
<i>Knowing What to Believe (when you already know something)</i>	
Jeff Pasternack and Dan Roth	877
<i>Using Web-scale N-grams to Improve Base NP Parsing Performance</i>	
Emily Pitler, Shane Bergsma, Dekang Lin and Kenneth Church	886

<i>Citation Summarization Through Keyphrase Extraction</i>	
Vahed Qazvinian, Dragomir R. Radev and Arzucan Ozgur	895
<i>2D Trie for Fast Parsing</i>	
Xian Qian, Qi Zhang, Xuangjing Huang and Lide Wu	904
<i>The Bag-of-Opinions Method for Review Rating Prediction from Sparse Text Patterns</i>	
Lizhen Qu, Georgiana Ifrim and Gerhard Weikum	913
<i>An Exploration of Features for Recognizing Word Emotion</i>	
Changqin Quan and Fuji Ren	922
<i>Inducing Fine-Grained Semantic Classes via Hierarchical and Collective Classification</i>	
Altat Rahman and Vincent Ng	931
<i>Fast, Greedy Model Minimization for Unsupervised Tagging</i>	
Sujith Ravi, Ashish Vaswani, Kevin Knight and David Chiang	940
<i>Bringing Active Learning to Life</i>	
Ines Rehbein, Josef Ruppenhofer and Alexis Palmer	949
<i>Computing EM-based Alignments of Routes and Route Directions as a Basis for Natural Language Generation</i>	
Michael Roth and Anette Frank	958
<i>A Multiple-Domain Ontology Builder</i>	
Sara Salem and Samir AbdelRahman	967
<i>Weakly Supervised Morphology Learning for Agglutinating Languages Using Small Training Sets.</i>	
Ksenia Shalnova and Bruno Golenia	976
<i>Multi-Document Summarization via the Minimum Dominating Set</i>	
Chao Shen and Tao Li	984
<i>Corpus-based Semantic Class Mining: Distributional vs. Pattern-Based Approaches</i>	
Shuming Shi, Huibin Zhang, Xiaojie Yuan and Ji-Rong Wen	993
<i>Metaphor Identification Using Verb and Noun Clustering</i>	
Ekaterina Shutova, Lin Sun and Anna Korhonen	1002
<i>Explore the Structure of Social Tags by Subsumption Relations</i>	
Xiance Si, Zhiyuan Liu and Maosong Sun	1011
<i>Ukwabelana - An open-source morphological Zulu corpus</i>	
Sebastian Spiegler, Andrew van der Spuy and Peter A. Flach	1020
<i>EMMA: A novel Evaluation Metric for Morphological Analysis</i>	
Sebastian Spiegler and Christian Monson	1029
<i>Modeling Socio-Cultural Phenomena in Discourse</i>	
Tomek Strzalkowski, George Aaron Broadwell, Jennifer Stromer-Galley, Samira Shaikh, Sarah Taylor and Nick Webb	1038

<i>Discriminative Induction of Sub-Tree Alignment using Limited Labeled Data</i>	
Jun Sun, Min Zhang and Chew Lim Tan	1047
<i>Investigating the cross-linguistic potential of VerbNet-style classification</i>	
Lin Sun, Thierry Poibeau, Anna Korhonen and Cedric Messiant	1056
<i>Semi-supervised dependency parsing using generalized tri-training</i>	
Anders Søgaard and Christian Rishøj	1065
<i>SemanticRank: Ranking Keywords and Sentences Using Semantic Graphs</i>	
George Tsatsaronis, Iraklis Varlamis and Kjetil Nørvåg	1074
<i>Chinese CCGbank: extracting CCG derivations from the Penn Chinese Treebank</i>	
Daniel Tse and James R. Curran	1083
<i>Dependency Forest for Statistical Machine Translation</i>	
Zhaopeng Tu, Yang Liu, Young-Sook Hwang, Qun Liu and Shouxun Lin	1092
<i>Large Scale Parallel Document Mining for Machine Translation</i>	
Jakob Uszkoreit, Jay Ponte, Ashok Popat and Moshe Dubiner	1101
<i>Hungarian Corpus of Light Verb Constructions</i>	
Veronika Vincze and János Csirik	1110
<i>Syntax Based Reordering with Automatically Derived Rules for Improved Statistical Machine Translation</i>	
Karthik Visweswariah, Jiri Navratil, Jeffrey Sorensen, Vijil Chenthamarakshan and Nandakishore Kambhatla	1119
<i>Efficient Statement Identification for Automatic Market Forecasting</i>	
Henning Wachsmuth, Peter Prettenhofer and Benno Stein	1128
<i>Towards a Unified Approach to Simultaneous Single-Document and Multi-Document Summarizations</i>	
Xiaojun Wan	1137
<i>"Got You!": Automatic Vandalism Detection in Wikipedia with Web-based Shallow Syntactic-Semantic Modeling</i>	
William Yang Wang and Kathleen McKeown	1146
<i>Exploiting Salient Patterns for Question Detection and Question Retrieval in Community-based Question Answering</i>	
Kai Wang and Tat-Seng Chua	1155
<i>Probabilistic Tree-Edit Models with Structured Latent Variables for Textual Entailment and Question Answering</i>	
Mengqiu Wang and Christopher Manning	1164
<i>A Character-Based Joint Model for Chinese Word Segmentation</i>	
Kun Wang, Chengqing Zong and Keh-Yih Su	1173
<i>Near-synonym Lexical Choice in Latent Semantic Space</i>	
Tong Wang and Graeme Hirst	1182

<i>Disambiguating Dynamic Sentiment Ambiguous Adjectives</i> Yunfang Wu and Miaomiao Wen	1191
<i>Joint Tokenization and Translation</i> Xinyan Xiao, Yang Liu, YoungSook Hwang, Qun Liu and Shouxun Lin	1200
<i>Build Chinese Emotion Lexicons Using A Graph-based Algorithm and Multiple Resources</i> Ge Xu, Xinfan Meng and Houfeng Wang	1209
<i>A Methodology for Automatic Identification of Nocuous Ambiguity</i> Hui Yang, Anne De Roeck, Alistair Willis and Bashar Nuseibeh	1218
<i>Contextual Modeling for Meeting Translation Using Unsupervised Word Sense Disambiguation</i> Mei Yang and Katrin Kirchhoff	1227
<i>A Working Report on Statistically Modeling Dative Variation in Mandarin Chinese</i> Yao Yao and Feng-hsi Liu	1236
<i>Kernel Slicing: Scalable Online Training with Conjunctive Features</i> Naoki Yoshinaga and Masaru Kitsuregawa	1245
<i>Discriminative Training for Near-Synonym Substitution</i> Liang-Chih Yu, Hsiu-Min Shih, Yu-Ling Lai, Jui-Feng Yeh and Chung-Hsien Wu	1254
<i>Estimating Linear Models for Compositional Distributional Semantics</i> Fabio Massimo Zanzotto, Ioannis Korkontzelos, Francesca Fallucchi and Suresh Manandhar	1263
<i>Grouping Product Features Using Semi-Supervised Learning with Soft-Constraints</i> Zhongwu Zhai, Bing Liu, Hua Xu and Peifa Jia	1272
<i>Forest-guided Supertagger Training</i> Yao-zhong Zhang, Takuya Matsuzaki and Jun'ichi Tsujii	1281
<i>Entity Linking Leveraging Automatically Generated Annotation</i> Wei Zhang, Jian Su, Chew Lim Tan and Wen Ting Wang	1290
<i>A Rhetorical Syntax-Driven Model for Speech Summarization</i> Jian Zhang and Pascale Fung	1299
<i>Maximum Metric Score Training for Coreference Resolution</i> Shanheng Zhao and Hwee Tou Ng	1308
<i>Paraphrasing with Search Engine Query Logs</i> Shiqi Zhao, Haifeng Wang and Ting Liu	1317
<i>Leveraging Multiple MT Engines for Paraphrase Generation</i> Shiqi Zhao, Haifeng Wang, Xiang Lan and Ting Liu	1326
<i>Resolving Surface Forms to Wikipedia Topics</i> Yiping Zhou, Lan Nie, Omid Rouhani-Kalleh, Flavian Vasile and Scott Gaffney	1335

<i>Heterogeneous Parsing via Collaborative Decoding</i>	
Muhua Zhu, Jingbo Zhu and Tong Xiao	1344
<i>A Monolingual Tree-based Translation Model for Sentence Simplification</i>	
Zhemín Zhu, Delphine Bernhard and Iryna Gurevych	1353
<i>A Minimum Error Weighting Combination Strategy for Chinese Semantic Role Labeling</i>	
Tao Zhuang and Chengqing Zong	1362
<i>Detecting Speech Repairs Incrementally Using a Noisy Channel Approach</i>	
Simon Zwarts, Mark Johnson and Robert Dale	1371
<i>Syntactic Scope Resolution in Uncertainty Analysis</i>	
Lilja Øvrelid, Erik Velldal and Stephan Oepen	1379