# Table of Contents

Preface: Conference Chair ..... iii

Preface: Program Committee Chair ..... v

Organizers ..... vii

## Invited and Plenary Talks

*A Step toward Compositional Semantics: E-HowNet a Lexical Semantic Representation System*
Keh-Jiann Chen and Shu-Ling Huang ..... 1

*Gesture Theory is Linguistics: On Modelling Multimodality as Prosody*
Dafydd Gibbon ..... 9

*Automatic Lexical Classification – Balancing between Machine Learning and Linguistics*
Anna Korhonen ..... 19

*Resultatives as Causal Relations between Events*
Ik-Hwan Lee ..... 29

*Developing Speech Recognition and Synthesis Technologies to Support Computer-Aided Pronunciation Training for Chinese Learners of English*
Helen Meng ..... 40

*Generating Chinese Couplets and Quatrain Using a Statistical Approach*
Ming Zhou, Long Jiang and Jing He ..... 43

## Regular Papers

*Capturing Lexical Variation in MT Evaluation Using Automatically Built Sense-Cluster Inventories*
Marianna Apidianaki, Yifan He and Andy Way ..... 53

*Dependency Grammar Based English Subject-Verb Agreement Evaluation*
Dongfeng Cai, Yonghua Hu, Xuelei Miao and Yan Song ..... 63

*WikiSense: Supersense Tagging of Wikipedia Named Entities Based WordNet*
Joseph Chang, Richard Tzong-Han Tsai and Jason S. Chang ..... 72

*An Integrated Approach to Heterogeneous Data for Information Extraction*
Ying Chen, Sophia Y. M. Lee and Chu-Ren Huang ..... 82

*Are Emotions Enumerable or Decomposable? And its Implications for Emotion Processing*
Ying Chen, Sophia Y. M. Lee and Chu-Ren Huang ..... 92

*Scope and Anaphoric Links in Dynamic Discourse Representation Theory*
So-Woo Chung and Jungmin Lee ..... 101

*Coupling an Annotated Corpus and a Morphosyntactic Lexicon for State-of-the-Art POS Tagging with Less Human Effort*
Pascal Denis and Benoît Sagot ..... 110

*Voted Approach for Part of Speech Tagging in Bengali*
Asif Ekbal, Md. Hasanuzzaman and Sivaji Bandyopadhyay . . . . . 120

*Adjective Density as a Text Formality Characteristic for Automatic Text Classification: A Study Based on the British National Corpus*
Alex Chengyu Fang and Jing Cao . . . . . 130

*Correcting Errors Using the Framework of Argumentation: Towards Generating Argumentative Correction Propositions from Error Annotation Schemas*
Marie Garnier, Arnaud Rykner and Patrick Saint-Dizier . . . . . 140

*Shallow Semantic Parsing of Persian Sentences*
Azadeh Kamel Ghalibaf, Saeed Rahati and Azam Estaji . . . . . 150

*LogisticLDA: Regularizing Latent Dirichlet Allocation by Logistic Regression*
Jia-Cheng Guo, Bao-Liang Lu, Zhiwei Li and Lei Zhang . . . . . 160

*Dependency Relations as Source Context in Phrase-Based SMT*
Rejwanul Haque, Sudip Kumar Naskar, Antal van den Bosch and Andy Way . . . . . 170

*Multi-Task Learning in Conditional Random Fields for Chunking in Shallow Semantic Parsing*
Saike He, Xiaojie Wang, Yuan Dong, Taozheng Zhang and Xue Bai . . . . . 180

*Query-Focused Multi-Document Summarization Using Co-Training Based Semi-Supervised Learning*
Po Hu, Donghong Ji, Hai Wang and Chong Teng . . . . . 190

*Review Classification Using Semantic Features and Run-Time Weighting*
Chung-chi Huang, Meng-chiech Lee, Zhe-nan Lin and Jason S. Chang . . . . . 200

*A Key Word Analysis of English Intensifying Adverbs in Male and Female Speech in ICE-GB*
Jieun Jeon and Jae-Woong Choe . . . . . 210

*Gender and Animacy Knowledge Discovery from Web-Scale N-Grams for Unsupervised Person Mention Detection*
Heng Ji and Dekang Lin . . . . . 220

*Layer-Based Dependency Parsing*
Ping Jian and Chengqing Zong . . . . . 230

*Mediatory Summary Generation: Summary-Passage Extraction for Information Credibility on the Web*
Koichi Kaneko, Hideyuki Shibuki, Masahiro Nakano, Rintaro Miyazaki, Madoka Ishioroshi and Tatsunori Mori . . . . . 240

*An Experimental Syntactic Study of Binding: A Case Study of Korean Long-Distance Anaphor caki*
Ji-Hye Kim and James H. Yoon . . . . . 250

*Method of Extracting Is-A and Part-Of Relations Using Pattern Pairs in Mass Corpus*
Se-Jong Kim, Yong-Hun Lee and Jong-Hyeok Lee . . . . . 260

*On the Scope Interaction of Japanese Indefinites: An Epsilon Calculus Approach*
Masahiro Kobayashi and Hiroaki Nakamura . . . . . 269

*Pattern Lattice as a Model for Linguistic Knowledge and Performance*
Kow Kuroda . . . . . 278

*Chinese WordNet Domains: Bootstrapping Chinese WordNet with Semantic Domain Labels*
Lung-Hao Lee, Yu-Ting Yu and Chu-Ren Huang . . . . . 288

*Cause Event Representations for Happiness and Surprise*
Sophia Yat Mei Lee, Ying Chen and Chu-Ren Huang . . . . . 297

*Sentiment Classification Considering Negation and Contrast Transition*
Shoushan Li and Chu-Ren Huang . . . . . 307

*Modal Verbs for the Advice Move in Advice Columns*
Ying-shu Liao and Ting-gen Liao . . . . . 317

*Approach to Selecting Best Development Set for Phrase-Based Statistical Machine Translation*
Peng Liu, Yu Zhou and Chengqing Zong . . . . . 325

*Using Extra-Linguistic Material for Mandarin-French Verbal Constructions Comparison*
Pierre Magistry, Laurent Prévot, Hintat Cheung, Chien-yun Shiao, Yann Desalle and Bruno Gaume . . . . . 335

*Improving Unsegmented Dialogue Turns Annotation with N-gram Transducers*
Carlos-D. Martínez-Hinarejos, Vicent Tamarit and José-Miguel Benedí . . . . . 345

*Using Tree Kernels for Classifying Temporal Relations between Events*
Seyed Abolghasem Mirroshandel, Gholamreza Ghassem-Sani and Mahdy Khayyamian . . . . . 355

*How Language Processing Constrains (Computational) Natural Language Processing: A Cognitive Perspective*
Prakash Mondal . . . . . 365

*Extending Bilingual WordNet via Hierarchical Word Translation Classification*
Tzu-yi Nien, Tsun Ku, Chung-chi Huang, Mei-hua Chen and Jason S. Chang . . . . . 375

*Extended GL and Japanese Postposition No*
Sumiyo Nishiguchi . . . . . 385

*Note on Japanese Epistemic Verb Constructions: A Surface-Compositional Analysis*
Akira Ohtani and Mark Steedman . . . . . 395

*On the So-Called Thematic Use of Wa: Reconsideration and Reconciliation*
David Y. Oshima . . . . . 405

*Chinese Semantic Class Learning from Web Based on Concept-Level Characteristics*
Wenbo Pang, Xiaozhong Fan, Jiangde Yu and Yuxiang Jia . . . . . 415

*On Pseudogapping in HPSG*
Dong-woo Park . . . . . 425

*Resolving Multiple Wh-Fragments: A Syntax-Pragmatics Approach*
Sang-Hee Park . . . . . 435

*A Lexicalized Tree Adjoining Grammar for Thai*
Siripong Potisuk . . . . . 444

*Incorporate Web Search Technology to Solve Out-of-Vocabulary Words in Chinese Word Segmentation*
Wei Qiao and Maosong Sun . . . . . 454

*A Hybrid Model for Sense Guessing of Chinese Unknown Words*
Likun Qiu, Kai Zhao and Changjian Hu . . . . . 464

*Collaborative Summarization: When Collaborative Filtering Meets Document Summarization*
Yang Qu and Qunxiu Chen . . . . . 474

*Finding Answers to Definition Questions Using Web Knowledge Bases*
Han Ren, Donghong Ji, Jing Wan and Chong Teng . . . . . 484

*Incorporating Statistical Information of Lexical Dependency into a Rule-Based Parser*
Yoon-Hyung Roh, Ki-Young Lee and Young-Gil Kim . . . . . 493

*An Automated Thematic Role Labeler and Generalizer for Filipino Verb Arguments*
Briane Paul Samson, Bianca Pamela Alcera, Ed Oswald Go, Czarina Meg Gonzales and Nathalie Rose Lim . . . . . 501

*Hybrid N-gram Probability Estimation in Morphologically Rich Languages*
Hyopil Shin and Hyunjo You . . . . . 511

*Summarization Approaches Based on Document Probability Distributions*
Sandeep Sripada and Jagadeesh Jagarlamudi . . . . . 521

*Chinese Function Tag Labeling*
Weiwei Sun and Zhifang Sui . . . . . 530

*What L2 Processing Strategy Reveals about the Prototypical Relationship: A Case of Japanese Modal Markers of Possibility*
Mizuho Tamaji . . . . . 540

*HMM Expanded to Multiple Interleaved Chains as a Model for Word Sense Disambiguation*
Denis Turdakov and Dmitry Lizorkin . . . . . 549

*Passage Retrieval Using Answer Type Profiles in Question Answering*
Surya Ganesh Veeravalli and Vasudeva Varma . . . . . 559

*A Continuum-Based Approach for Tightness Analysis of Chinese Semantic Units*
Ying Xu, Christoph Ringlstetter and Randy Goebel . . . . . 569

*A Framework for Effectively Integrating Hard and Soft Syntactic Rules into Phrase Based Translation*
Jiajun Zhang and Chengqing Zong . . . . . 579

*A Bootstrapping Method for Finer-Grained Opinion Mining Using Graph Model*
Shu Zhang, Yingju Xia, Yao Meng and Hao Yu . . . . . 589

*SESS: A Self-Supervised and Syntax-Based Method for Sentiment Classification*
Weishi Zhang, Kai Zhao, Likun Qiu and Changjian Hu . . . . . 596

## Poster Papers

*Summarizing Opinions in Blog Threads*
Alexandra Balahur, Mijail Kabadjov, Josef Steinberger, Ralf Steinberger and Andrés Montoyo . . . . . 606

*Constraint Based Hybrid Approach to Parsing Indian Languages*
Akshar Bharati, Samar Husain, Meher Vijay, Kalyan Deepak, Dipti Misra Sharma and Rajeev Sangal . . . . . 614

*Interpolated PLSI for Learning Plausible Verb Arguments*
Hiram Calvo, Kentaro Inui and Yuji Matsumoto . . . . . 622

*An EM Algorithm for Context-Based Searching and Disambiguation with Application to Synonym Term Alignment*
Jing-Shin Chang and Shih-Jay Chiou . . . . . 630

*VocabAnalyzer: A Referred Word List Analyzing Tool with Keyword, Concordancing and N-gram Functions*
Siaw-Fong Chung, F.Y. August Chao and Yi-Chen Hsieh . . . . . 638

*Building Online Corpora of Philippine Languages*
Shirley N. Dita, Rachel Edita O. Roxas and Paul Inventado . . . . . 646

*Machine Translation Using Automatically Inferred Construction-Based Correspondence and Language Models*
Shimon Edelman and Zach Solan . . . . . 654

*Latin Etymologies as Features on BNC Text Categorization*
Alex Chengyu Fang, Wanyin Li and Nancy Ide . . . . . 662

*Experiments on Domain Adaptation for English–Hindi SMT*
Rejwanul Haque, Sudip Kumar Naskar, Josef van Genabith and Andy Way . . . . . 670

*Extraction of English Ditransitive Constructions Using Formal Concept Analysis*
Yoichiro Hasebe and Kow Kuroda . . . . . 678

*Bridging the Gap between Graph Modeling and Developmental Psycholinguistics: An Experiment on Measuring Lexical Proximity in Chinese Semantic Space*
Shu-Kai Hsieh, Chun-Han Chang, Ivy Kuo, Hintat Cheung, Chu-Ren Huang and Bruno Gaume . . . . . 686

*Effective Use of Chinese Structural Auxiliaries for Chinese Parsing*
Yun Jin, Qing Li, Yingshun Wu and Young-Gil Kim . . . . . 694

*Towards Conceptual Indexing of the Blogosphere through Wikipedia Topic Hierarchy*
Mariko Kawaba, Daisuke Yokomoto, Hiroyuki Nakasaki, Takehito Utsuro and Tomohiro Fukuhara . . . . . 702

*Syntactic Category Prediction for Improving Translation Quality in English-Korean Machine Translation*
Sung-Dong Kim . . . . . 710

*Customizing an English-Korean Machine Translation System for Patent/Technical Documents Translation*
Oh-Woog Kwon, Sung-Kwon Choi, Ki-Young Lee, Yoon-Hyung Roh and Young-Gil Kim .... 718

*Word Boundary Decision with CRF for Chinese Word Segmentation*
Shoushan Li and Chu-Ren Huang ..... 726

*Extracting Keyphrases from Chinese News Articles Using TextRank and Query Log Knowledge*
Weiming Liang, Chang-Ning Huang, Mu Li and Bao-Liang Lu ..... 733

*Modeling the Relationship among Linguistic Typological Features with Hierarchical Dirichlet Process*
Chu-Cheng Lin, Yu-Chun Wang and Richard Tzong-Han Tsai ..... 741

*Document Re-ranking via Wikipedia Articles for Definition/Biography Type Questions*
Maofu Liu, Fang Fang and Donghong Ji ..... 748

*Towards Bilingual Term Extraction in Comparable Patents*
Bin Lu and Benjamin K. Tsou ..... 755

*Factors Affecting Part-of-Speech Tagging for Tagalog*
Erlyn Manguilimotan and Yuji Matsumoto ..... 763

*GuideLink: A Corpus Annotation System that Integrates the Management of Annotation Guidelines*
Kenta Oouchida, Jin-Dong Kim, Toshihisa Takagi and Jun'ichi Tsujii ..... 771

*Using Wikipedia for Hierarchical Finer Categorization of Named Entities*
Aasish Pappu ..... 779

*A Novel Method of Sentence Ordering Based on Support Vector Machine*
Gongfu Peng, Yanxiang He, Ye Tian, Yingsheng Tian and Weidong Wen ..... 787

*Supertagging with Factorial Hidden Markov Models*
Srivatsan Ramanujam and Jason Baldridge ..... 795

*Identifying and Utilizing the Class of Monosemous Japanese Functional Expressions in Machine Translation*
Akiko Sakamoto, Taiji Nagasaka, Takehito Utsuro and Suguru Matsuyoshi ..... 803

*Named Entity Recognition for Manipuri Using Support Vector Machine*
Thoudam Doren Singh, Kishorjit Nongmeikapam, Asif Ekbal and Sivaji Bandyopadhyay ..... 811

*BaseNP Supersense Tagging for Japanese Texts*
Hirotoshi Taira, Sen Yoshida and Masaaki Nagata ..... 819

*Which is More Suitable for Chinese Word Segmentation, the Generative Model or the Discriminative One?*
Kun Wang, Chengqing Zong and Keh-Yih Su ..... 827

*Design of Chinese HPSG Framework for Data-Driven Parsing*
Xiangli Wang, Shunya Iwasawa, Yusuke Miyao, Takuya Matsuzaki, Kun Yu and Jun'ichi Tsujii ..... 835

*Rule-based Korean Grapheme to Phoneme Conversion Using Sound Patterns*
Yu-Chun Wang and Richard Tzong-Han Tsai ..... 843

*Discovery of Dependency Tree Patterns for Relation Extraction*
Hongzhi Xu, Changjian Hu and Guoyang Shen . . . . . 851

*Generalizable Features Help Semantic Role Labeling*
Li Yang . . . . . 859

*Utilizing Features of Verbs in Statistical Zero Pronoun Resolution for Japanese Speech*
Sen Yoshida and Masaaki Nagata . . . . . 867

*Towards Establishing a Hierarchy in the Japanese Sentence Structure*
Kei Yoshimoto, Chidori Nakamura and Alastair Butler . . . . . 875

Author Index . . . . . 883