Table of Contents

Foreword ................................................................. iii
Organizers ................................................................. iv

Invited and Plenary Talks

Symposium on Language Resources

PodCastle: A Spoken Document Retrieval Service Improved by Anonymous User Contributions
Masataka Goto and Jun Ogata ........................................... 3

The Future Role of Language Resources for Natural Language Parsing
(We Won’t Be Able to Rely on Pierre Vinken Rorever… or Will We Have to?)
Julia Hockenmaier .......................................................... 13

The Acquisition of Word Order in a Topic-prominent Language:
Corpus Findings and Experimental Investigation
Thomas Hun-Tak Lee ........................................................ 15

Symposium: Toward a Linguistic Theory of Ba: Semantics and Pragmatics of Ba

How and Why Two Strangers Can Co-create a Story:
An Application of the ‘Ba’-Theory Based Approach to the Discourse
Sachiko Ide ................................................................. 19

Situated Language: Case of Japanese
Hideyuki Nakashima ...................................................... 25

Listening In
Stanley Peters ............................................................. 31

Plenary Talks

Basic Units of Lexicons and Ontologies: Words, Senses and Concepts
Hee-Rahk Chae ........................................................... 35

A Morphosyntactic Analysis of the Pronominal System of Philippine Languages
Shirley N. Dita ............................................................... 45

Statistical Translation Model Based On Source Syntax Structure
Qun Liu, Yang Liu and Haitai Mi ......................................... 61
Regular Papers

Lexical Semantics-Syntactic Model for Defining and Subcategorizing Attribute Noun Class
Xiaopeng Bai and Hui Wang ................................................................. 65

Benglish Verbs: A Case of Code-mixing in Bengali
Shishir Bhattacharja ............................................................................. 75

Enhanced Genre Classification through Linguistically Fine-Grained POS Tags
Alex Chengyu Fang and Jing Cao ....................................................... 85

Identifying Emotional Expressions, Intensities and Sentence Level Emotion Tags Using a Supervised Framework
Dipankar Das and Sivaji Bandyopadhyay .......................................... 95

Decision Theory and Discourse Particles: A Case Study from a Large Japanese Sentiment Corpus
Christopher Davis ........................................................................... 105

Finding Appropriate Subset of Votes Per Classifier Using Multiobjective Optimization: Application to Named Entity Recognition
Asif Ekbal, Sriparna Saha and Md. Hasanuzzaman ............................. 115

Experimental Approach to Reference in Discourse: Working Memory Capacity and Language Comprehension in Russian
Olga Fedorova, Ekaterina Delikishkina and Anna Uspenskaya .......... 125

A Computational Model of Language Generation Applied to Japanese Wh-questions
Jason Ginsburg .................................................................................. 133

Chains in Syntax and Morphology
Thomas Gross ..................................................................................... 143

Feature Subset Selection Using Genetic Algorithm for Named Entity Recognition
Md. Hasanuzzaman, Sriparna Saha and Asif Ekbal ............................ 153

Developing Punjabi Morphology, Corpus and Lexicon
Muhammad Humayoun and Aarne Ranta ........................................ 163

Effective Use of Linguistic Features for Sentiment Analysis of Korean
Hayeon Jang and Hyopil Shin ............................................................. 173

Word Order and NP Structure in Korean: A Constraint Based Approach
Jong-Bok Kim, Nam-Guen Lee and Yae-shiek Lee ............................. 183

Evidentials and Epistemic Modal in Korean: Evidence from Their Interactions
Chungmin Lee .................................................................................... 193
Implementation of Korean Syllable Structures in the Typed Feature Structure Formalism
Gyu-hyung Lee, Ye-seul Park and Yong-hun Lee ........................................... 203

Focus Types and Subject-Object Asymmetry in Korean Case Ellipsis:
A New Look at Focus Effects
Hanjung Lee and Haejeong Choi ................................................................. 213

eSpatialML: An Event-Driven Spatial Annotation Framework
Kiyong Lee, Jonathan Webster and Alex Chengyu Fang ............................... 223

Domain-Independent Novel Event Discovery and Semi-Automatic Event Annotation
Hao Li, Xiang Li, Heng Ji and Yuval Marton .................................................. 233

Developing an Online Indonesian Corpora Repository
Ruli Manurung, Bayu Distiawan and Desmond Darma Putra ......................... 243

Topicalization and Truth Conditions: A Categorial Account
Hiroaki Nakamura ............................................................................................ 251

Unsupervised Classification of Biomedical Abstracts Using Lexical Association
Jonathon Read, Jonathan Webster and Alex Chengyu Fang ............................. 261

Licensing Nominals in the Multiple Nominative Constructions in Korean:
A Mereological Perspective
Byong-Rae Ryu .................................................................................................. 271

Combination of 3 Types of Speech Recognizers for Anaphora Resolution
Kazutaka Shimada, Noriko Tanamachi and Tsutomu Endo .............................. 281

Syntactic and Semantic Discrepancies among the Verbs for ‘kill’
in English, Chinese and Thai
Kingkarn Thepkanjana and Satoshi Uehara .................................................. 291

Expanding Chinese Sentiment Dictionaries from Large Scale Unlabeled Corpus
Hongzhi Xu, Kai Zhao, Likun Qiu and Changjian Hu ....................................... 301

Representing Events in Japanese Complex Predicates
Yoshie Yamamori ............................................................................................... 311

Fault-Tolerant Learning for Term Extraction
Yuhang Yang, Hao Yu, Yao Meng, Yingliang Lu and Yingju Xia ....................... 321

Syntactically Complex Demonstratives and Sortal Inherency
Richard Zuber .................................................................................................... 331

Poster Papers

On Automated Hyperonym Hierarchy Construction
Using an Internet Search Engine
Sergey Afonin .................................................................................................... 341
An Ontological Analysis of Japanese and Chinese Kinship Terms
  Songiy Baik and Hee-Rahk Chae .................................................. 349

GRASP: Grammar- and Syntax-based Pattern-Finder for Collocation and Phrase Learning
  Mei-hua Chen, Chung-chi Huang, Shih-ting Huang and Jason S. Chang ........ 357

Mitigating Problems in Analogy-based EBMT with SMT and vice versa: A Case Study with Named Entity Transliteration
  Sandipan Dandapat, Sara Morrissey, Sudip Kumar Naskar and Harold Somers .... 365

Why the English Easiest Type Became the Hardest in Russian, or Russian Adults’ Comprehension of before and after sentences
  Olga Fedorova ............................................................................... 373

Simpler Is Better: Re-evaluation of Default Word Alignment Models in Statistical MT
  Mark Fishel .................................................................................. 381

The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme
  Serge Heiden ................................................................. 389

Using Corpus-based Linguistic Approaches in Sense Prediction Study
  Jia-Fei Hong, Sue-Jin Ker, Chu-Ren Huang and Kathleen Ahrens .................. 399

Synthetic Compounds in Mandarin
  Shuang Hong ................................................................. 409

  Darren Hsin-hung Lin and Shelley Ching-yu Hsieh .......................... 417

The KOLON System: Tools for Ontological Natural Language Processing in Korean
  Juliano Paiva Junho, Yumi Jo and Hyopil Shin .................................. 425

An Approach toward Register Classification of Book Samples in the Balanced Corpus of Contemporary Written Japanese
  Wakako Kashino and Manabu Okumura ....................................... 433

A Computational Treatment of Korean Serial Verb Constructions
  Jong-Bok Kim, Jaehyung Yang and Sanghoun Song ............................. 439

A Supervised Machine Learning Approach for Event-Event Relation Identification
  Anup Kumar Kolya, Asif Ekbal and Sivaji Bandyopadhyay ...................... 447

Arguments for Parallel Distributed Parsing: Toward the Integration of Lexical and Sublexical (Semantic) Parsings
  Kow Kuroda ............................................................................... 455
Scrambling of Wh-phrases in Japanese
Paul Law ................................................................. 463

A Comparative Study on the Aspectual Classification of Korean and Japanese Verbs
in Relation to ‘-ess’ ‘-essess’ in Korean and ‘suden’ in Japanese
Younghee Lee and Byong-Rae Ryu ................................ 471

A Modular Architecture for the Wide-Coverage Translation of Natural Language Texts
into Predicate Logic Formulas
Yusuke Miyao, Alastair Butler, Kei Yoshimoto and Jun’ichi Tsujii .......... 481

Toward a Unified Computational Model of Quantificational Scope Readings
Prakash Mondal .......................................................... 489

CCG of Japanese Sentence-final Particles
Sumiyo Nishiguchi ..................................................... 497

A Multi-Dimensional Analysis of Japanese Benefactives:
The Case of the Yaru-Construction
Akira Otani and Mark Steedman .................................... 503

Typed Dependency Relations for Syntactic Analysis of Thai Sentences
Siripong Potisuk .......................................................... 511

Exploiting a Multilingual Web-based Encyclopedia
for Bilingual Terminology Extraction
Fatiha Sadat .............................................................. 519

Language Homogeneity in the Japanese Wikipedia
Karl-André Skevik ....................................................... 527

Incorporate Credibility into Context for the Best Social Media Answers
Qi Su, Helen Kai-yun Chen and Chu-Ren Huang ..................... 535

Student Award Papers

A Query Focused Multi Document Automatic Summarization
Pinaki Bhaskar and Sivaji Bandyopadhyay .......................... 545

Through Low-Cost Annotation to Reliable Parsing Evaluation
Marek Gráč, Miloš Jakubíček and Vojtěch Kovář ..................... 555

A Look into the Acquisition of English Motion Event Conflation
by Native Speakers of Chinese and Japanese
Ryan Spring .............................................................. 563

Stage/Individual-level Predicates, Topics and Indefinite Subjects
Shuiying Yao .............................................................. 573

How Well Conditional Random Fields Can be Used in Novel Term Recognition
Xing Zhang, Yan Song and Alex Chengyu Fang ....................... 583
Student Papers

Clitics in Arabic Language: A Statistical Study
Fahad Alotaiby, Salah Foda and Ibrahim Alkharashi .......................... 595

AUTOLEX: An Automatic Lexicon Builder for Minority Languages
Using an Open Corpus
Evan Liz C. Buhay, Marie Joy P. Evardone, Hansel B. Nocon,
Davis Muhajereen D. Dimalen and Rachel Edita O. Roxas ................. 603

Improving Information Extraction Using Knowledge Model
Yue Chen .................................................................................. 613

Finding Emotion Holder from Bengali Blog Texts: An Unsupervised Syntactic Approach
Dipankar Das and Sivaji Bandyopadhyay .................................... 621

Terminological Ontology and Cognitive Processes in Translation
Fumiko Kano Glückstad ................................................................. 629

Detection of Users Suspected of Pretending to Be Other Users in a Community Site
by Using Messages Submitted to Non-Target Categories
Naoki Ishikawa, Ryo Nishimura, Yasuhiro Watanabe,
Masaki Murata and Yoshihiro Okada ....................................... 637

Detecting Nasty Comments from BBS Posts
Tatsuya Ishisaka and Kazuhide Yamamoto ................................. 645

Using Various Features in Machine Learning to Obtain High Levels of Performance
for Recognition of Japanese Notational Variants
Masahiro Kojima, Masaki Murata, Jun’ichi Kazama, Kow Kuroda, Atsushi Fujita,
Eiji Aramaki, Masaaki Tsuchida, Yasuhiro Watanabe and Kentaro Torisawa ...... 653

Korean Kes-Relative Clauses as Stages
Sang-Hee Park ............................................................................ 661

Finite State Morphology and Sindhi Noun Inflections
Mutee U Rahman Muslim and Mohammad Iqbal Bhatti .................. 669

The Acquisition of Imperfective Aspect Marking in Korean as a Second language
by Japanese Learners
Ju-Yeon Ryu, Kaoru Horie and Yasuhiro Shirai .............................. 677

Multi-aspects Review Summarization Based on Identification of Important Opinions
and their Similarity
Ryosuke Tadano, Kazutaka Shimada and Tsutomu Endo ................. 685

Generation of Summaries that Appropriately and Adequately Express the Contents
of Original Documents Using Word-Association Knowledge
Kazuki Takigawa, Masaki Murata, Masaaki Tsuchida, Stijn De Saeger,
Kazuhide Yamamoto and Kentaro Torisawa ............................... 693
Adjectival Modification to Nouns in Mandarin Chinese:
Case Studies on “cháng+noun” and “adjective+tú shǔ gu n”
Shan Wang and Chu-Ren Huang ......................................................... 701

Compositional Operations of Mandarin Chinese Perception Verb “kàn”:
A Generative Lexicon Approach
Shan Wang and Chu-Ren Huang ......................................................... 707

Workshop on Advanced Corpus Solutions

Workshop on Advanced Corpus Solutions
Janne Bondi Johannessen ................................................................. 717

Degrees of Orality in Speech-like Corpora:
Comparative Annotation of Chat and E-mail Corpora
Eckhard Bick .................................................................................. 721

Parallel Suffix Arrays for Corpus Exploration
Johannes Goller .............................................................................. 731

Fast Syntactic Searching in Very Large Corpora for Many Languages
Miloš Jakubíček, Adam Kilgariff, Diana McCarthy and Pavel Rychlík ............ 741

A Multilingual Speech Resource: The Nordic Dialect Corpus
Janne Bondi Johannessen, Joel Priestley and Anders Nøklestad .................... 749

Dialect Corpora Taken Further:
The DynaSAND Corpus and Its Application in Newer Tools
Jan Pieter Kunst and Franca Wesseling .............................................. 759

Advanced Corpus Solutions for Humanities Researchers
James Wilson, Anthony Hartley, Serge Sharoff and Paul Stephenson ............... 769

Entity Coherence in Comparable Learner Corpora: Seeking Pedagogical Insights
Mitsuko Yamura-Takei, Miho Fujiwara and Etsuko Yoshida ...................... 779

Workshop on Model and Measurement of Meaning (M3)

The Use of a Cultural Protocol for Quantifying Cultural Variations in Verb Semantic between Chinese and French
Hintat Cheung, Yann Desalle, Karine Duvignau, Bruno Gaume, Chunhan Chang and Pierre Magistry ........................................ 791

Towards the Global SentiWordNet
Amitava Das and Sivaji Bandyopadhyay ............................................ 799

Towards an Automatic Measurement of Verbal Lexicon Acquisition:
The Case for a Young Children-versus-Adults Classification in French and Mandarin
Yann Desalle, Shu-Kai Hsieh, Bruno Gaume and Hintat Cheung .................... 809
Graph Representation of Synonymy and Translation Resources for Crosslinguistic Modelisation of Meaning
Benoît Gaillard, Yannick Chudy, Pierre Magistry, Shu-Kai Hsieh and Emmanuel Navarro ................................................................. 819

Exploring Chinese Verbal Lexicon Developmental Trend with Semantic Space
Ching-Fen Pan .................................................................................. 831

Computational Modeling of Verb Acquisition, from a Monolingual to a Bilingual Study
Laurent Prévol, Chun-Han Chang and Yann Desalle ............................... 841

Cross-sortal Predication and Polysemy
Petr Šimon and Chu-Ren Huang .......................................................... 853

Verb Use in Chinese Children: Extensibility of Instrument
Te-Hsin Wu ....................................................................................... 863

Workshop on Semantics and Logical Flavor: In Honor of Professor Akira Ikeya

Natural Language Production in Database Semantics
Roland Hausser .............................................................................. 875

Change of Location and Change of State: How Telicity is Attained
Chungmin Lee .................................................................................. 885

Various Evidentials in Korean
Kyung An Song .............................................................................. 895

A Note on Pseudo-comparatives like “John is rich like X!” and “Like X, John is rich!”
Benjamin T’sou ............................................................................. 907