

# Table of Contents

## Invited Speeches

Virtual Linked Lexical Knowledge Base for Causality Reasoning <i>Choi Key-Sun</i>	1
Focus-Marking in Chinese and Malay: A Comparative Perspective <i>Jie Xu</i>	2

## Paper Presentation

### Linguistics-1

A Constraint-Based Analysis of Association With Focus in Japanese <i>Yusuke Kubota</i>	16
On the Event Structure of Indirect Passives in Japanese <i>Naoyuki Ono</i>	28
An Event-Based Interpretation of Japanese Honorific Constructions Using RRG Operators <i>Akira Ishikawa</i>	38
Multiple Nominative Constructions in Japanese and Their Theoretical Implications <i>Masahiro Akiyama</i>	50

### NLP-1

Automatic Learning of Stemming Rules for the Indonesian Language <i>Lily Suryana Indradjaja and Stephane Bressan</i>	62
A Simplified Latent Semantic Indexing Approach for Multi-Linguistic Information Retrieval <i>Liu Yi, Lu Haiming, Lu Zengxiang and Wang Pu</i>	69

An integrated approach for Chinese word segmentation <i>Guohong Fu and K.K. Luke</i>	80
Korean Phrase Structure Grammar and Its Implementations into the LKB system <i>Jong-Bok Kim and Jaehyung Yang</i>	88
Porting Grammars between Typologically Similar Languages: Japanese to Korean <i>Roger Kim, Mary Dalrymple, Ronald M. Kaplan and Tracy Holloway King</i>	98
 <b>Linguistics-2</b>	
Mandarin Chinese <i>Shenme</i> in Interaction <i>Fuhui Hsieh</i>	106
The Semantics of Shapes: A Study based on Mandarin <i>Quan1zi5</i> (圈子) <i>Cui-xia Weng and Chu-Ren Huang</i>	115
Stock Markets as Ocean Water: A Corpus-based, Comparative Study of Mandarin Chinese, English and Spanish <i>Siaw-Fong Chung, Kathleen Ahrens and Ya-hui Sung</i>	124
Subject Positions and Derivational Scope Calculation in Minimalist Syntax: A <i>Phase-Based</i> approach <i>Yukiko Ueda</i>	134
 <b>NLP-2</b>	
Context-rule Model for POS tagging <i>Yu-Fang Tsai and Keh-Jiann Chen</i>	146
Chinese Word Segmentation based on Contextual Entropy <i>Jin Hu Huang and David Powers</i>	152
Topic Segmentation for Short Texts <i>Tao-Hsing Chang and Chia-Honag Lee</i>	159
Cross-Lingual Text Filtering Based on Text Concepts and KNN <i>Li Shaozi, Su Weifeng, Li Tangqiu and Chen Huowang</i>	166

### Linguistics-3

- The Semantics of Onomatopoeic Speech Act Verbs 174  
*I-Ni Tsai and Chu-Ren Huang*
- Mandarin Adverbial *Jiu* in Discourse 182  
*Fuhui Hsieh*
- A Synchronous Corpus-Based Study of Verb-Noun Fluidity in Chinese 194  
*Oi Yee Kwong and Benjamin K. Tsou*
- Non-Monotonic Negativity 204  
*Sumiyo Nishiguchi*

### NLP-3

- The SVM With Uneven Margins and Chinese Document Categorization 216  
*Yaoyong Li and John Shawe-Taylor*
- The development of Tagged Uyghur Corpus 228  
*Yusup Aibaidula and Kim-Teng Lua*
- A Vector-Based Algorithm for Chinese Text Classification 235  
*Luo Changri and He Tingting*
- A Large-scale Lexical Semantic Knowledge-Base of Chinese 243  
*Wang Hui and Yu Shiwen*

### NLP-4

- An Effective Combination of Different Order N-grams 251  
*Sen Zhang and Na Dong*
- Efficient Methods for Multigram Compound Discovery 257  
*Wu Horng Jyh Paul, Ng Hong I and Gong Ruibin*
- Translation Template Learning Based on Hidden Markov Modeling 269  
*Nguyen Minh Le, Akari Shimazu and Susumu Horiguchi*
- News-Oriented Keyword Indexing with Maximum Entropy Principle 277  
*Li Sujian, Wang Houfeng, Yu Shiwen and Xin Chengsheng*

## NLP-5

- Extracting Chinese Multi-Word Units from Large-Scale Balanced Corpus 282  
*Liu Jianzhou, He Tingting and Liu Xiaohua*
- A New Sentence Reduction based on Decision Tree Model 290  
*Nguyen Minh Le and Susumu Horiguchi*
- Japanese Parser on the basis of the Lexical-FG Formalism and its Evaluation 298  
*Hiroshi Masuichi, Tomoko Ohkuma, Hiroki Yoshimura and Yasunari Harada*
- A Statistical Approach to Chinese-to-English Back-Transliteration 310  
*Chun-Jen Lee, Jason S. Chang and Jyh-Shing Roger Jang*
- On the Sentence Category Transfer of Action-effect Sentences in Chinese-English Machine Translation 319  
*Keliang Zhang*

## Linguistics-4

- Modeling verb Order in Complex Multi-Verbal Predicate Constructions 328  
*Olivia S.-C. Lam and Adams B. Bodo*
- The Structure of Spatial Expressions in Saisiyat 339  
*Pei-Shu Tsai*
- A Constraint-Based Grammar of Case: To Correctly Predict Case Phrases Occurring without Their Head Verb 351  
*Hiroki Koga*
- The Floating of Negative Factors and the Recognition of Semantic Patterns of *Huaiyi* Sentences in Mandarin 362  
*Xiao Guozheng and Guo Tingting*

## Poster

- Using Mutual Information to Identify New Features for Text documents of Various Domains 372  
*Guo Zhi Li*
- Applicability Analysis of Corpus-derived Paraphrases toward Example-based Paraphrasing 380  
*Kiyonori Ohtake and Kazuhide Yamamoto*

A Word Selection Model Based on Lexical Semantic Knowledge in English Generation <i>Chen Yi-Dong, Li Tang-Qiu and Zheng Xu-Ling</i>	392
Corpus-Based Ontology Learning for Word Sense Disambiguation <i>Sin Jae Kang</i>	399
On Intra-page and Inter-page Semantic Analysis of Web Pages <i>Jun Wang, Jicheng Wang, Gangshan Wu and Hiroshi Tsuda</i>	408
Towards a Multi-Objective Corpus for Vietnamese Language <i>Vu Hai Quan, Pham Nam Trung, Nguyen Duc Hoang Ha, Huynh Bao Toan, Le Hoai Bac and Hoang Kiem</i>	416
Using Zero Anaphora Resolution to Improve Text Categorization <i>Ching-Long Yeh and Yi-Chun Chen</i>	423
Dependency of Long-Distance Reflexives <i>Hyeran Lee</i>	431
Voicing Constraint and Segmental-Tonal Neighborhood Effects on Clusters in Thai <i>Rattima Nitisaroj</i>	441
The treatment of Japanese Focus Particles based on Lexical-Functional Grammar <i>Tomoko Ohkuma, Hiroshi Masuichi, Hiroki Yoshimura and Yasunari Harada</i>	448
Empty Category and the effect of Teaching in Sentence Processing <i>Takako Kawasaki and Kiyoshi Ishikawa</i>	456