Proceedings of the 30th Pacific Asia Conference on Language, Information and Computation

edited by

Jong C. Park,
KAIST

Jin-Woo Chung,
KAIST

© The PACLIC 30
October 28-30, 2016
Acknowledgments

PACLIC 30 is organized by Korean Society for Language and Information (KSLI), Korea Advanced Institute of Science and Technology (KAIST), and Institute for the Study of Language and Information at Kyung Hee University (ISLI), under the auspices of PACLIC Steering Committee.

This work was supported by the NRF (National Research Foundation of Korea) Grant funded by the MOE (Ministry of Education).
Welcome & Acknowledgments

Welcome to the 30th Pacific Asia Conference on Language, Information and Computation (PACLIC30)!

We are really happy to have you all here at Kyung Hee University. Ever since 1995, the PACLIC has been a meeting place for researchers working on language and related fields. We have tried to share our research results in the field of theoretical and computational linguistics. It has been a favorite conference for many theoretical and computational linguists in the Pacific-Asia region as well as around the world. This year is no exception. We have researchers from various regions including Europe, the United States, and Middle East.

In addition to regular paper and poster presentations, we have three exceptional plenary talks and five invited papers on ‘interesting issues of language and computing’, respectively. The three plenary talks will be given by Stefan Müller from Freie Universität Berlin, Goran Nenadic from University of Manchester, and Anoop Sarkar from Simon Fraser University. The five invited papers are delivered by Jae-Woong Choe from Korea University, Shirley Dita from De La Salle University, Yasunari Harada from Waseda University, Chu-Ren Huang from The Hong Kong Polytechnic University, and Hongyong Liu from University of Macau. We thank all these plenary speakers and invited papers for the contribution. The selection process of regular papers and posters has been highly competitive with the acceptance rate of only 26.9% (29 papers) for oral presentation and 28.7% (31 posters) for poster presentation from the total 108 submitted papers. This can ensure the high quality of the oral presentations as well as posters.

Together with the distinguished plenary speakers and invited papers and high quality papers and posters, I have a strong belief that all the participants will greatly benefit from each other, sharing stimulating ideas, discussing provoking suggestions, lightening the direction of our linguistic research in future, and so forth.

This conference would not be possible without the support and efforts from many people, including program members and reviewers, organizers, volunteers, and sponsors. I would especially like to thank the Program Committee Chair, Prof. Jong C. Park, who tackled the many tasks associated with planning a program for the conference with enthusiasm and efficiency. In particular, many thanks go to Sanghoun Song for designing and maintaining the conference website and helping to coordinate the efforts of the Organizing Committee. I also thank student staff members Rok Sim, Okgi Kim, SeulKee Park, and graduate students at Kyung Hee University for their time and efforts to make everything for this conference run smoothly. We also thank the National Research Foundation of Korea, Institute for the Study of Language and Information at Kyung Hee University, Korean Society for Language and Information, and Korea Advanced Institute of Science and Technology for the generous financial and logistics support for the conference.

We hope you enjoy the conference and your stay in Kyung Hee, Seoul!

Jong-Bok Kim

PACLIC 30 Organization Committee Chair
Preface

The 30th Pacific Asia Conference on Language, Information and Computation (PACLIC 30) is held at Kyung Hee University in Seoul on (Friday) 28 October – (Sunday) 30 October 2016. The conference is co-hosted by the Korea Society of Language and Information, Institute for the Study of Language and Information at KHU, and Korea Advanced Institute of Science and Technology. The PACLIC series of conferences emphasize the synergy of theoretical analysis and processing of language, and provide a forum for researchers in different fields of language study in the Pacific-Asia region to share their findings and interests in the formal and empirical study of languages. Organized under the auspices of the PACLIC Steering Committee, PACLIC 30 continues our long standing collaborative efforts among theoretical and computational linguists in the Pacific-Asia region.

This year, we received 108 paper submissions that represent healthy diversity, with authors from 22 countries, which include Australia, Cambodia, China, Czech Republic, Denmark, France, Hong Kong, India, Indonesia, Ireland, Israel, Italy, Japan, Korea, New Zealand, Philippines, Russia, Taiwan, Thailand, United Kingdom, United States, and Vietnam (80.2% from 13 countries in Asia Pacific, 11.7% from 7 countries in Europe, 6.1% from the United States, and 2.0% from the Middle East). We wish to extend our deep thanks to all the authors for submitting papers to PACLIC 30 and for their contributions.

We would also like to thank 53 Program Committee members for writing over 327 reviews and for making the final paper selection possible among 108 submissions in total. After receiving acceptance letters for either oral or poster presentations, however, some authors of accepted papers chose to withdraw their submissions or to change their presentation modes afterwards. As a result, we have 29 papers for oral presentation (26.9% acceptance rate) and 31 papers for poster presentation (55.6% acceptance rate). Each submission was reviewed by at least three reviewers to ensure that all accepted papers meet the high quality standards of the PACLIC conference. We are extremely grateful to the Program Committee members for all their hard work, without which the preparation of this program would not have been possible at all.

We are delighted to have three plenary keynote speeches and five invited papers addressing different aspects of language and computing in PACLIC 30. The three keynote speeches are given by Stefan Müller from Freie Universität Berlin, Goran Nenadic from University of Manchester, and Anoop Sarkar from Simon Fraser University. The five invited papers are delivered by Jae-Woong Choe from Korea University, Shirley Dita from De La Salle University, Yasunari Harada from Waseda University, Churen Huang from The Hong Kong Polytechnic University, and Hongyong Liu from University of Macau. These plenary and invited papers will not be only informative but also enlightening to the audience, leading to many innovative research ideas in the future. We would like to thank General Chair Hee-Rahk Chae and the Steering Committee for their valuable help and advice. We also wish to extend our appreciation to Organization Chair Jong-Bok Kim and the Organization Committee members, for their exceptional dedication and coordination in their work.

We hope that you enjoy the conference!

Jong C. Park
PACLIC 30 Program Committee Chair
PACLIC 30 Organizers

Steering Committee:
Hee-Rahk Chae, Hankuk University of Foreign Studies, Seoul
Chu-Ren Huang, The Hong Kong Polytechnic University, Hong Kong
Rachel Edita O. Roxas, National University, Manila
Maosong Sun, Tsinghua University, Beijing
Benjamin T'sou, City University of Hong Kong, Hong Kong
Kei Yoshimoto, Tohoku University, Sendai
Min Zhang, Soochow University, Suzhou

Conference Chair:
Hee-Rahk Chae, Hankuk University of Foreign Studies

Organization Committee Chair:
Jong-Bok Kim, Kyung Hee University

Organization Committee:
Sang-gu Kang, Hansung University
Dongsik Lim, Hongik University
Sanghoun Song, Incheon National University

Program Committee Chair:
Jong C. Park, KAIST

Program Committee Members:
Wirote Aroonmanakun, Chulalongkorn University
Stephane Bressan, National University of Singapore
Kuang-Hua Chen, National Taiwan University
Hsin-Hsi Chen, National Taiwan University
Eng Siong Chng, Nanyang Technological University
Sae Youn Cho, Kangwon National University
Sung-Kwon Choi, ETRI
Jin-Woo Chung, KAIST
Siaw-Fong Chung, National Chengchi University
Beatrice Daille, Laboratoire d'Informatique de Nantes Atlantique
Shirley Dita, De La Salle University
Minghui Dong, Institute for Infocomm Research
Guohong Fu, Hefei University
Wei Gao, Qatar Computing Research Institute
Program Committee Coordinator:

Jin-Woo Chung, KAIST
# Table of Contents

## Keynote Speeches

*The CoreGram Project: Theoretical Linguistics, Theory Development and Verification*
Stefan Müller ................................................................. 3

*Inferring Methodological Meta-knowledge from Large Biomedical Corpora*
Goran Nenadic ..................................................................... 5

*The Challenge of Simultaneous Speech Translation*
Anoop Sarkar ......................................................................... 7

## Invited Papers

*The Significance of Background Information in Acceptability Judgements of Korean Sentences*
Jae-Woong Choe .................................................................... 11

*Measuring Diversified Proficiency of Japanese Learners of English*
Yasunari Harada .................................................................... 13

*Endurant vs Perdurant: Ontological Motivation for Language Variations*
Chu-Ren Huang .................................................................... 15

*The Syntax of the Chinese Excessive Resultative Construction*
Hongyong Liu ....................................................................... 27

*The grammar and semantics of disjuncts in World Englishes*
Shirley Dita .......................................................................... 35

## Oral Presentation Papers

*A “Maximal Exclusion” Approach to Structural Uncertainty in Dynamic Syntax*
Tohru Seraku ......................................................................... 39

*Korean Language Resources for Everyone*
Jungyeul Park, Jeen-Pyo Hong and Jeong-Won Cha ......................... 49

*Secondary Predicates in Native and Nonnative Grammars*
Enchao Shi ........................................................................... 59

*A Generalized Framework for Hierarchical Word Sequence Language Model*
Xiaoyi Wu, Kevin Duh and Yuji Matsumoto .................................. 69

*Processing English Island Sentences by Korean EFL Learners*
Yeonkyung Park and Yong-Hun Lee ........................................... 77

*Multiple Emotions Detection in Conversation Transcripts*
Duc Anh Phan, Hiroyuki Shindo and Yuji Matsumoto ......................... 85

*Long-distance anaphors and the blocking effect revisited-An East Asian perspective*
Hyunjun Park ......................................................................... 95
Developing an Unsupervised Grammar Checker for Filipino Using Hybrid N-grams as Grammar Rules
Matthew Philip Go and Allan Borra ............................................................... 105

Supervised Word Sense Disambiguation with Sentences Similarities from Context Word Embeddings
Shoma Yamaki, Hiroyuki Shinno, Kanako Komiya and Minoru Sasaki ............................... 115

HSSA tree structures for BTG-based preordering in machine translation
Yujia Zhang, Hao Wang and Yves Lepage ..................................................... 123

The Manner/Result Complementarity in Chinese Motion Verbs Revisited
Lei Qiu ....................................................................................................... 133

Yet Another Symmetrical and Real-time Word Alignment Method: Hierarchical Sub-sentential Alignment using F-measure
Hao Wang and Yves Lepage ..................................................................... 143

Event Based Emotion Classification for News Articles
Minglei Li, Da Wang, Qin Lu and Yunfei Long .................................................. 153

The interaction of politeness systems in Korean learners of French
Darcy Sperlich, Jaiho Leem and Eui-Jeen Ahn ................................................. 163

Integrating Word Embedding Offsets into the Espresso System for Part-Whole Relation Extraction
Van-Thuy Phi and Yuji Matsumoto ............................................................... 173

An Experimental Study of Subject Properties in Korean Multiple Subject Constructions (MSCs)
Ji-Hye Kim, Eunah Kim and James Yoon ...................................................... 183

Focal Prominence Underlying Distribution of Mandarin Minimizers
I-Hsuan Chen ............................................................................................ 191

Planting Trees in the Desert: Delexicalized Tagging and Parsing Combined
Daniel Zeman, David Mareček, Zhiwei Yu and Zdeněk Žabokrtský .......................... 199

Recurrent Neural Network Based Loanwords Identification in Uyghur
Chenggang Mi, Yating Yang, Xi Zhou, Lei Wang, Xiao Li and Tonghai Jiang .................. 209

Solving Event Quantification and Free Variable Problems in Semantics for Minimalist Grammars
Yu Tomita .................................................................................................... 219

Testing APSyn against Vector Cosine on Similarity Estimation
Enrico Santus, Emmanuelle Chersoni, Alessandro Lenci, Chu-Ren Huang and Philippe Blache ............................ 229

Recognizing Open-Vocabulary Relations between Objects in Images
Masayasu Muraoka, Sumit Maharjan, Masaki Saito, Kota Yamaguchi, Naoaki Okazaki, Takayuki Okatani and Kentaro Inui ............................................................ 239

Strong Associations Can Be Weak: Some Thoughts on Cross-lingual Word Webs for Translation
Oi Yee Kwong ............................................................................................ 249

Dealing with Out-Of-Vocabulary Problem in Sentence Alignment Using Word Similarity
Hai-Long Trieu, Le-Minh Nguyen and Phuong-Thai Nguyen .................................. 259

A Pipeline Japanese Entity Linking System with Embedding Features
Shuangshuang Zhou .................................................................................... 267

Toward the automatic extraction of knowledge of usable goods
Mei Uemura, Naho Orita, Naoaki Okazaki and Kentaro Inui .............................. 277

A Syntactic Approach to the 1st Person Restriction of Causal Clauses in Korean
Poster Presentation Papers

Retrieval Term Prediction Using Deep Learning Methods
Qing Ma, Ibuki Tanigawa and Masaki Murata ................................................................. 309

Japanese Post-verbal Constructions Revisited
Kohji Kamada .................................................................................................................. 319

Sentence Clustering using PageRank Topic Model
Kenshin Ikegami and Yukio Ohsawa ............................................................................... 329

The Inner Circle vs. the Outer Circle or British English vs. American English
Yong-Hun Lee and Ki-Suk Jun ......................................................................................... 339

A Correlation Analysis of English Particle Placement of Three East Asian EFL Learners Writings
Ha-Eung Kim, Gyu-Hyeong Lee and Yong-Hun Lee ....................................................... 347

The sources of new words and expressions in the Chinese Internet language and the ways by which they enter the Internet language
Aleksandr Sboev .............................................................................................................. 355

Sentiment Clustering with Topic and Temporal Information from Large Email Dataset
Sisi Liu, Ickjai Lee and Guochen Cai ................................................................................ 363

On What an Adnominal Appendix Modifies in Korean Adjunct RDCs
Daeho Chung .................................................................................................................. 373

Automatic Identifying Entity Type in Linked Data
Qingliang Miao, Ruiyu Fang, Shuangyong Song, Zhongguang Zheng, Lu Fang, Yao Meng and Jun Sun .................................................. 383

SMTPOST Using Statistical Machine Translation Approach in Filipino Part-of-Speech Tagging
Nicco Nocon and Allan Borra ......................................................................................... 391

From built examples to attested examples: a syntax-based query for non-specialists
Ilaine Wang, Sylvain Kahane and Isabelle Tellier ............................................................. 397

A Study of Valence & Argument Integration in Chinese Verb-Resultative Complement
Anran Li .......................................................................................................................... 407

Phonological Principles for Automatic Phonetic Transcription of Khmer Orthographic Words
Makara Sok and Larin Adams ......................................................................................... 417

On the Possessor Interpretation of Non-Agentive Subjects
Tomokazu Takehisa ........................................................................................................ 425

Philippine Language Resources: Applications, Issues, and Directions
Nathaniel Oco, Leif Romeritch Syliongka, Tod Allman and Rachel Edita Roxas .................. 433

Generating a Linguistic Model for Requirement Quality Analysis
Juyeon Kang and Jungyeul Park ....................................................................................... 439

Designing CzeDLex – A Lexicon of Czech Discourse Connectives
Jiří Mírovský, Pavlína Jinová, Magdaléna Rysová and Lucie Poláková ......................... 449
Transitivity in Light Verb Variations in Mandarin Chinese -- A Comparable Corpus-based Statistical Approach
Menghan Jiang, Dingxu Shi and Chu-Ren Huang ................................................................. 459

The Interaction between SFP-Ne and SpOAs in Mandarin Chinese--A corpus based approach
Yifan He .................................................................................................................................. 469

On the semantics of Korean modalized question
Arum Kang ............................................................................................................................. 477

The Synaesthetic and Metaphorical Uses of 菜 蔬 ‘taste’ in Chinese Buddhist Suttas
Jiajuan Xiong and Chu-Ren Huang .......................................................................................... 485

L2 Acquisition of Korean locative construction by English L1 speakers: Learnability problem in Korean Figure non-alternating verbs
Sun Hee Park .......................................................................................................................... 493

Towards a unified account of resultative constructions in Korean
Juwon Lee ............................................................................................................................... 501

The use of body part terms in Taiwan and China: Analyzing 血 xue ‘blood’ and 骨 gu ‘bone’ in Chinese
cGigaword v. 2.0
Ren-Feng Duann and Chu-Ren Huang ..................................................................................... 511

A POMDP-based Multimodal Interaction System Using a Humanoid Robot
Sae Iijima and Ichiro Kobayashi .............................................................................................. 519

The Cloud of Knowing: Non-factive al-ta ‘know’ (as a Neg-raiser) in Korean
Chungmin Lee and Seungjin Hong .......................................................................................... 527

Neural Joint Learning for Classifying Wikipedia Articles into Fine-grained Named Entity Types
Masatoshi Suzuki, Koji Matuda, Satoshi Sekine, Naoaki Okazaki and Kentaro Inui ................ 535

MINING CALL CENTER CONVERSATIONS EXHIBITING SIMILAR AFFECTIVE STATES
Rupayan Chakraborty, Meghna Pandharipande and Sunil Kumar Kopparapu ........................ 545