Proceedings of the 33rd Pacific Asia Conference on Language, Information and Computation

Edited by
Ryo Otoguro, Mamoru Komachi and Tomoko Ohkuma

September 13–15, 2019
Future University Hakodate
Foreword

It is our great pleasure and honor to hold the 33rd Pacific Asia Conference on Language, Information and Computation (PACLIC 33) at Future University Hakodate, Japan. “Open space, Open mind” is the underlying philosophy of Future University Hakodate, which makes it exactly the right place to host PACLIC. Following the long tradition of PACLIC, PACLIC 33 also emphasizes the synergy of theoretical analysis and processing of natural language. PACLIC 33 aims to enhance the interaction between researchers working in different fields of language study in the Asia-Pacific region as well as around the world.

We received 132 submissions from around the world including Algeria, Brazil, Chile, China, Czechia, France, Hong Kong, India, Indonesia, Iran, Italy, Japan, Macao, Mexico, Philippines, Singapore, South Korea, Taiwan, Thailand, Tunisia, Ukraine, the UK, the US and Vietnam. Out of 132 papers, 37 were accepted for oral presentations and 26 for poster presentations. The acceptance rate for oral presentations and poster presentations are 28% and 20% respectively.

In addition to oral and poster presentations, the conference highlights four keynote talks, one special invited talk and one satellite workshop. We are grateful to Justine Cassell from Carnegie Mellon University, Mary Dalrymple from University of Oxford, Yuji Matsumoto from Nara Institute of Science and Technology/Riken AIP and Junichi Tsujii from National Institute of Advanced Industrial Science and Technology for accepting to give a keynote talk. Jong-Bok Kim from Kyung Hee University has kindly agreed to give an invited talk in commemoration of the Humboldt Research Award given to him in 2019. We also thank Yasunari Harada, Chu-Ren Huang, Jong-Bok Kim, Yasuhiro Katagiri and Miwa Morishita for organizing the 27th Joint Workshop on Linguistics and Language Processing during the conference and Rachel Edita O. Roxas and Manolito V. Octaviano for giving an invited talk in the workshop.

PACLIC 33 would not be made possible without the support from many people. We would like to express our sincere gratitude toward program committee members and sub-reviewers whose professional reviews allowed us to maintain the high quality standard of PACLIC. We are also deeply indebted to the local organizing committee at Future University Hakodate: Yasuhiro Katagiri, Hitoshi Matsubara, Hajime Murai, Asuka Terai, Misako Nambu, Kaoru Sumi and Ayahiko Niimi as well as student staff members. We would also like to thank Waseda University Comprehensive Research Organization, The Association for Natural Language Processing, The Japan Society for Artificial Intelligence and Hakodate City for their generous financial support for the conference.

Ryo Otoguro  
Mamoru Komachi  
Tomoko Ohkuma  
PACLIC 33 Program Committee Chairs
Organizers

Steering Committee Standing Members
Chu-Ren Huang, The Hong Kong Polytechnic University, Hong Kong
Jong-Bok Kim, Kyung Hee University, Seoul
Ryo Otoguro, Waseda University, Tokyo
Rachel Edita O. Roxas, National University, Manila
Maosong Sun, Tsinghua University, Beijing
Benjamin T'sou, City University of Hong Kong, Hong Kong
Min Zhang, Soochow University, Suzhou

Organizing Committee
Ryo Otoguro, Waseda University (Chair)
Yasunari Harada, Waseda University (Co-chair)
Yasuhiro Katagiri, Future University Hakodate (Honorary chair)

Local Organizing Committee
Yasuhiro Katagiri, Future University Hakodate (Chair)
Hitoshi Matsubara, Future University Hakodate (Co-chair)
Hajime Murai, Future University Hakodate
Asuka Terai, Future University Hakodate
Misako Nambu, Future University Hakodate
Kaoru Sumi, Future University Hakodate
Ayahiko Niimi, Future University Hakodate

Program Committee

Chairs
Ryo Otoguro, Waseda University
Mamoru Komachi, Tokyo Metropolitan University
Tomoko Ohkuma, Fuji Xerox Co.
Members

Laurence Anthony  
Alice Mae Arbon  
Masayuki Asahara  
Qian Chen  
Doris Chen  
Charibeth Cheng  
Emmanuele Chersoni  
Sung-Kwon Choi  
Jin-Woo Chung  
Li Dong  
Helena Gao  
Yasunari Harada  
Hitomi Hirayama  
Jeffrey J. Holliday  
Munpyo Hong  
Shu-Kai Hsieh  
Jiangping Huang  
Suyeon Im  
Tomoyuki Kajiwara  
Daisuke Kawahara  
Jong-Bok Kim  
Ji-Hye Kim  
Kanako Komiyama  
Valia Kordoni  
Yusuke Kubota  
Oi Yee Kwong  
Huei-Ling Lai  
Olivia Lam  
Yong-Hun Lee  
Albert Lee  
Sang-Im Lee-Kim  
Baoli Li  
Wei-Wen Liao  
Dongsik Lim  
Jingxia Lin  
Te-Hsin Liu  
Bingquan Liu  
Yunfei Long  
Lu Lu  
Chen Lyu  
Erlyn Manguilimotan  
Yuji Matsumoto  
Koji Mineshima  
Yasuhide Miura  
Porrudee Netisopakul  
Takashi Ninomiya  
Hitoshi Nishikawa  
Nathaniel Oco  
Kenji Oda  
Ethel Ong  
Chutamanee Onsuwan  
Yohei Oseki  
David Yoshikazu Oshima  
Jong C. Park  
Nattama Pongpairoj  
Haoliang Qi  
Tao Qian  
Rodolfo Jr Raga  
Yafeng Ren  
Hiroyuki Shinnou  
Shu-Ing Shyu  
Melanie Siegel  
Pornsiri Singhapreecha  
Leif Romeritch Syliongka  
Zhiyang Teng  
Yuen-Hsien Tseng  
Takehito Utsuro  
Zhongqing Wang  
Tak-Sum Wong  
Hongzi Xu  
Yun Xue  
Jie Yang  
Cheng-Zen Yang  
Satoru Yokoyama  
Minoru Yoshida  
Liang-Chih Yu  
Longtu Zhang  
Meishan Zhang  
Jiajun Zhang  
Chengzhi Zhang

Additional reviewers

Kristine Mae Adlaon  
Yoshihiko Asao  
Eun Jin Chun  
Masahiro Kaneko  
Wakako Kashino  
Yuan Ling  
Rui Liu  
Rutian Ma  
Hiroki Narita  
Mizuki Sango  
Helen Villanueva  
Hiroaki Yamada  
Hayahide Yamagishi  
Qingqing Zhou
# Table of Contents

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

## Regular Papers

A Gold Standard Dependency Treebank for Indonesian Language  
*Ika Alfina, Arawinda Dinakaramani, Mohamad Ivan Fanany and Heru Suhartanto* .......... 1

Investigating an Effective Character-level Embedding in Korean Sentence Classification  
*Won Ik Cho, Seok Min Kim and Nam Soo Kim* ................................. 10

Incorporating Chains of Reasoning over Knowledge Graph for Distantly Supervised Biomedical Knowledge Acquisition  
*Qin Dai, Naoya Inoue, Paul Reisert, Ryo Takahashi and Kentaro Inui* .............. 19

Epistemic marker, event type and factivity in emotion expressions  
*Xuefeng Gao, Chu-Ren Huang and Sophia Yat-Mei Lee* .......................... 29

AMR Normalization for Fairer Evaluation  
*Michael Wayne Goodman* ................................................. 37

A CCG-based Compositional Semantics and Inference System for Comparatives  
*Izumi Haruta, Koji Mineshima and Daisuke Bekki* ............................ 47

A Type-Theoretical Approach to Register Classification  
*Renkui Hou and Chu-Ren Huang* ......................................... 57

Modeling the Idiomaticity of Chinese Quadra-syllabic Idiomatic Expressions  
*Shu-Kai Hsieh, Yu-Hsiang Tseng and Chiung-Yu Chiang* ....................... 68

V-gei Double Object Construction and Extra Argument in Mandarin  
*Yu-Yin Hsu and Teng Qu* .................................................. 76

Re-examining Syntactic, Semantic and Pragmatic Properties of Long-Distance Bound Cakicasin in Korean: An Experimental Study  
*Ji-Hye Kim and Yong-Hun Lee* ........................................... 85

The persuade-construction in Korean controls nothing  
*Juwon Lee and Sanghoun Song* .............................................. 95

Pretrained language model transfer on neural named entity recognition in Indonesian conversational texts  
*Rezka Leonandya and Fariz Ikhwantri* ...................................... 104
Long-distance dependencies in continuation grammar
Cara Su-Yi Leong and Michael Yoshitaka Erlewine ........................................ 114

On Null Clausal Complements in Taiwan Southern Min
Huei-Ling Lin ........................................................................................................ 123

A Community Detection Method Towards Analysis of Xi Feng Parties in the Northern Song
Dynasty
Qianying Liu, Qiyao Wang, Wending Chen and Daisuke Kawahara ....................... 129

Analysis of Reply-Tweets for Buzz Tweet Detection
Kazuyuki Matsumoto, Yuta Hada, Minoru Yoshida and Kenji Kita .......................... 138

Evaluating the suitability of human-oriented text simplification for machine translation
Rei Miyata and Midori Tatsumi ............................................................................. 147

Building Cendana: a Treebank for Informal Indonesian
David Moeljadi, Aditya Kurniawan and Debaditya Goswami ............................... 156

Simulating Segmentation by Simultaneous Interpreters for Simultaneous Machine Translation
Akiko Nakabayashi and Tsuneaki Kato .............................................................. 165

Attention mechanism for recommender systems
Xuan-Huy Nguyen and Le-Minh Nguyen ............................................................ 174

Identifying Adversarial Sentences by Analyzing Text Complexity
Hoang-Quoc Nguyen-Son, Tran Phuong Thao, Seira Hidano and Shinsaku Kiyomoto 182

Phi-Agreement by C in Japanese: Evidence from Person Restriction on the Subject
Miki Obata and Mina Sugimura .......................................................................... 191

Towards the Non-predicate Modification Analysis of the Expressive Small Clause in Japanese
Kenji Oda ............................................................................................................ 196

Syntax and Semantics of Numeral Classifiers in Japanese
Atsushi Oho ......................................................................................................... 203

An emoticon is well worth a few empathetic words
Juan Pablo Rodriguez Gomez, Tomoko Iizuka, Edson T. Miyamoto, Changyun Moon and Kaoruko Ouchi ................................................................. 212

Utilization of histories by country in question-answering system to solve world history essay type questions
Kotaro Sakamoto, Yuta Fukuhara, Madoka Ishioroshi, Kosuke Ohya, Keigo Iwasaki, Hideyuki Shibuki and Tatsunori Mori ....................................................... 219
Over-sampling Methods for Polarity Classification of Imbalanced Microblog Texts
Kiyoki Shirai and Yunmin Xiang .................................................. 228

Thai Learners of English are Sensitive to Number-Agreement Violations
Teeranoot Siriwittayakorn and Edson T. Miyamoto ............................ 237

May and Can Constructions in Spoken Corpus: A Constructionist Approach
Tsii-Chuen Tsai and Huei-Ling Lai .................................................. 244

On the Effectiveness of Low Rank Matrix Factorization for LSTM Model Compression
Genta Indra Winata, Andrea Madotto, Jamin Shin, Elham J. Barezi and Pascale Fung 253

Prospective Result of Causative Predicates: A Uniform Analysis
Yusuke Yagi .................................................................................. 263

Probabilistic Measures for Diffusion of Linguistic Innovation: As Seen in the Usage of Verbal “Nok” in Thai Twitter
Nozomi Yamada and Pittayawat Pittayaporn ...................................... 271

Thai Legal Term Correction using Random Forests with Outside-the-sentence Features
Takahiro Yamakoshi, Vee Satayamas, Hutchatai Chanlekha, Yasuhiro Ogawa, Takahiro Komamizu, Asanee Kawtrakul and Katsuhiko Toyama ............................... 279

A Corpus of Sentence-level Annotations of Local Acceptability with Reasons
Wonsuk Yang, Jung-Ho Kim, Seungwon Yoon, Chaehun Park and Jong C. Park ........ 288

Explicit Contextual Semantics for Text Comprehension
Zhuosheng Zhang, Yuwei Wu, Zuchao Li and Hai Zhao ........................ 298

Chinese–Japanese Unsupervised Neural Machine Translation Using Sub-character Level Information
Longtu Zhang and Mamoru Komachi ............................................. 309

FTA: a novel feature training approach for classification
Wanwan Zheng and Mingzhe Jin ................................................. 316

Poster papers

Bi-directional Decoder Model with Efficient Fine-tuning of Embedding for Named Entity Recognition
Panuwat Assawinjaipech, Kiyoki Shirai, Virach Sornlertlamvanich and Sanparith Marukatat .............................................................. 324

Making Metaphors: A Quantitative Analysis of Metaphor Production and Interpretation in Japanese Using a Multimodal Task
Brian Birdsell, Natsumo Tatsuta and Hiroaki Nakamura ........................ 334
Multiple Pivots in Statistical Machine Translation for Low Resource Languages
Sari Dewi Budiwati and Masayoshi Aritsugi ............................... 345

Semi-supervised learning for all-words WSD using self-learning and fine-tuning
Rui Cao, Jing Bai, Wen Ma and Hiroyuki Shinnou .......................... 356

A Reinforced Improved Attention Model for Abstractive Text Summarization
Yu Chang, Hang Lei, Xiaoyu Li and Yiming Huang ......................... 362

Semantic Distance and Creativity in Linguistic Synaesthesia
Emmanuele Chersoni, Francesca Strik Lievers and Chu-Ren Huang ........ 370

Investigating Mandarin Negative Terms: An Evaluation of Semantic-Pragmatic Meanings
and Metaphorical Mechanisms
Siaw-Fong Chung, Yi-Ling Tseng, Heng-Chia Liao and Man-Hua Huang .... 379

Mapping distributional to model-theoretic semantic spaces: a baseline
Franck Dernoncourt .................................................................. 388

Intrinsic Evaluation of Grammatical Information within Word Embeddings
Daniel Edmiston and Taeuk Kim .............................................. 395

A Continuation-based Analysis of Contrastive Wa in Japanese
Hitomi Hirayama ....................................................................... 405

Effects of Prosodic Focus on Voice Onset Time (VOT) in Chongming Chinese
Yitian Hong, Si Chen, Yike Yang and Bei Li .................................. 414

Web Page Segmentation for Non Visual Skimming
Judith Jeyafreeda, Stéphane Ferrari, Fabrice Maurel, Gaël Dias and Emmanuel Gigué ............................................................... 423

Automatic Speech Act Classification of Korean Dialogue based on the Hierarchical Structure
of Speech Act Categories
Youngeun Koo, Jiyoun Kim and Munpyo Hong ................................. 432

Investigation of Mandarin Clickbait Headlines: A Case Study of Biàn Zhèyàng
Chi-Ling Lee, Siaw-Fong Chung and Hui-Wen Liu ............................ 442

On the “Easy” Task of Evaluating Chinese Irony Detection
An-Ran Li, Emmanuele Chersoni, Rong Xiang, Chu-Ren Huang and Qin Lu .... 452

Towards Better Ad Experience: Click Prediction Leveraging Sequential Networks Derived
Specifically From User Search Behaviors
Shengzhe Li, Tomoko Izumi, Yu Kuratake, Jiali Yao, Jerry Turner, Daisuke Kawahara and
Sadao Kurohashi ........................................................................ 461
Cantonese turn-initial particles: annotation of discourse-interactional functions in dialog corpora

*Andreas Liesenfeld* ............................................................... 471

Are TERRORISM and kongbu zhuyi translation equivalents? A corpus-based investigation of meaning, structure and alternative translations

*Lily Lim* .............................................................................. 480

L1 and L2 Processing of Chinese Separable VO Compounds

*Junghwan Maeng* ................................................................. 488

Syntax and Semantics of Adjectives in Cape Verdean Creole: A View from Markedness

*Chigusa Morita and Miki Obata* .............................................. 496

Japanese Daily Utterance Styles: A Factor Analysis based on Balanced Corpus

*Hajime Murai* ........................................................................ 503

A Speaker Accent Recognition System for Filipino Language

*Batman Odulio, Justin Raphael Ariaso, Karl Adrian Cruz, Mico Ian Orjalo, Ramon Rodriguez, Angelica Dela Cruz and Manolito Octaviano Jr* .................................................. 511

A corpus-based investigation of collexemes for active-passive alternation in the English part of an English-Japanese parallel corpus

*Masanori Oya* ........................................................................ 516

Korean-to-Chinese Machine Translation using Chinese Character as Pivot Clue

*Jeonghyeok Park and Hai Zhao* .............................................. 522

Adapting Neural Machine Translation for English-Vietnamese using Google Translate system for Back-translation

*Nghia Luan Pham and Van Vinh Nguyen* ............................... 531

Re-unifying Floating Numeral Quantifiers and Secondary Predicates in Japanese

*Hideaki Yamashita* ................................................................. 540