

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

© 2019 The PACLIC 33 Organizing Committee and PACLIC Steering Committee

All rights reserved. Except as otherwise expressly permitted under copyright law, no part of this publication may be reproduced, digitized, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, Internet or otherwise, without the prior permission of the publisher.

Copyright of contributed papers reserved by respective authors

ISSN 2619-7782

Published by Waseda Institute for the Study of Language and Information, Waseda University, Tokyo, Japan

Acknowledgments

PACLIC 33 is hosted by Future University Hakodate in conjunction with The Japan Association for the Study of Logic, Language and Information and supported by Waseda University Comprehensive Research Organization, The Association for Natural Language Processing, The Japan Society for Artificial Intelligence and Hakodate City.

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 Ootoguro, 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 Ootoguro, 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 Ootoguro, Waseda University
Mamoru Komachi, Tokyo Metropolitan University
Tomoko Ohkuma, Fuji Xerox Co.

Members

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

Additional reviewers

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

Table of Contents

Foreword	iii
Organizers	iv

Regular Papers

A Gold Standard Dependency Treebank for Indonesian Language <i>Ika Alfina, Arawinda Dinakaramani, Mohamad Ivan Fanany and Heru Suhartanto</i> 1	
Investigating an Effective Character-level Embedding in Korean Sentence Classification <i>Won Ik Cho, Seok Min Kim and Nam Soo Kim</i>	10
Incorporating Chains of Reasoning over Knowledge Graph for Distantly Supervised Biomedical Knowledge Acquisition <i>Qin Dai, Naoya Inoue, Paul Reisert, Ryo Takahashi and Kentaro Inui</i>	19
Epistemic marker, event type and factivity in emotion expressions <i>Xuefeng Gao, Chu-Ren Huang and Sophia Yat-Mei Lee</i>	29
AMR Normalization for Fairer Evaluation <i>Michael Wayne Goodman</i>	37
A CCG-based Compositional Semantics and Inference System for Comparatives <i>Izumi Haruta, Koji Mineshima and Daisuke Bekki</i>	47
A Type-Theoretical Approach to Register Classification <i>Renkui Hou and Chu-Ren Huang</i>	57
Modeling the Idiomaticity of Chinese Quadra-syllabic Idiomatic Expressions <i>Shu-Kai Hsieh, Yu-Hsiang Tseng and Chiung-Yu Chiang</i>	68
V- <i>gei</i> Double Object Construction and Extra Argument in Mandarin <i>Yu-Yin Hsu and Teng Qu</i>	76
Re-examining Syntactic, Semantic and Pragmatic Properties of Long-Distance Bound <i>Cakicasin</i> in Korean: An Experimental Study <i>Ji-Hye Kim and Yong-Hun Lee</i>	85
The persuade-construction in Korean controls nothing <i>Juwon Lee and Sanghoun Song</i>	95
Pretrained language model transfer on neural named entity recognition in Indonesian conversational texts <i>Rezka Leonandya and Fariz Ikhwantri</i>	104

Long-distance dependencies in continuation grammar <i>Cara Su-Yi Leong and Michael Yoshitaka Erlewine</i>	114
On Null Clausal Complements in Taiwan Southern Min <i>Huei-Ling Lin</i>	123
A Community Detection Method Towards Analysis of Xi Feng Parties in the Northern Song Dynasty <i>Qianying Liu, Qiyao Wang, Wending Chen and Daisuke Kawahara</i>	129
Analysis of Reply-Tweets for Buzz Tweet Detection <i>Kazuyuki Matsumoto, Yuta Hada, Minoru Yoshida and Kenji Kita</i>	138
Evaluating the suitability of human-oriented text simplification for machine translation <i>Rei Miyata and Midori Tatsumi</i>	147
Building Cendana: a Treebank for Informal Indonesian <i>David Moeljadi, Aditya Kurniawan and Debaditya Goswami</i>	156
Simulating Segmentation by Simultaneous Interpreters for Simultaneous Machine Translation <i>Akiko Nakabayashi and Tsuneaki Kato</i>	165
Attention mechanism for recommender systems <i>Xuan-Huy Nguyen and Le-Minh Nguyen</i>	174
Identifying Adversarial Sentences by Analyzing Text Complexity <i>Hoang-Quoc Nguyen-Son, Tran Phuong Thao, Seira Hidano and Shinsaku Kiyomoto</i>	182
Phi-Agreement by C in Japanese: Evidence from Person Restriction on the Subject <i>Miki Obata and Mina Sugimura</i>	191
Towards the Non-predicate Modification Analysis of the Expressive Small Clause in Japanese <i>Kenji Oda</i>	196
Syntax and Semantics of Numeral Classifiers in Japanese <i>Atsushi Oho</i>	203
An emoticon is well worth a few empathetic words <i>Juan Pablo Rodriguez Gomez, Tomoko Iizuka, Edson T. Miyamoto, Changyun Moon and Kaoruko Ouchi</i>	212
Utilization of histories by country in question-answering system to solve world history essay type questions <i>Kotaro Sakamoto, Yuta Fukuhara, Madoka Ishioroshi, Kosuke Ohya, Keigo Iwasaki, Hideyuki Shibuki and Tatsunori Mori</i>	219

Over-sampling Methods for Polarity Classification of Imbalanced Microblog Texts <i>Kiyooki Shirai and Yunmin Xiang</i>	228
Thai Learners of English are Sensitive to Number-Agreement Violations <i>Teeranoot Siriwittayakorn and Edson T. Miyamoto</i>	237
<i>May</i> and <i>Can</i> Constructions in Spoken Corpus: A Constructionist Approach <i>Tsi-Chuen Tsai and Huei-Ling Lai</i>	244
On the Effectiveness of Low Rank Matrix Factorization for LSTM Model Compression <i>Genta Indra Winata, Andrea Madotto, Jamin Shin, Elham J. Barezi and Pascale Fung</i>	253
Prospective Result of Causative Predicates: A Uniform Analysis <i>Yusuke Yagi</i>	263
Probabilistic Measures for Diffusion of Linguistic Innovation: As Seen in the Usage of Verbal “Nok” in Thai Twitter <i>Nozomi Yamada and Pittayawat Pittayaporn</i>	271
Thai Legal Term Correction using Random Forests with Outside-the-sentence Features <i>Takahiro Yamakoshi, Vee Satayamas, Hutchatai Chanlekha, Yasuhiro Ogawa, Takahiro Komamizu, Asanee Kawtrakul and Katsuhiko Toyama</i>	279
A Corpus of Sentence-level Annotations of Local Acceptability with Reasons <i>Wonsuk Yang, Jung-Ho Kim, Seungwon Yoon, Chaehun Park and Jong C. Park</i>	288
Explicit Contextual Semantics for Text Comprehension <i>Zhuosheng Zhang, Yuwei Wu, Zuchao Li and Hai Zhao</i>	298
Chinese–Japanese Unsupervised Neural Machine Translation Using Sub-character Level In- formation <i>Longtu Zhang and Mamoru Komachi</i>	309
FTA: a novel feature training approach for classification <i>Wanwan Zheng and Mingzhe Jin</i>	316

Poster papers

Bi-directional Decoder Model with Efficient Fine-tuning of Embedding for Named Entity Recognition <i>Panuwat Assawinjaipetch, Kiyooki Shirai, Virach Sornlertlamvanich and Sanparith Marukatat</i>	324
Making Metaphors: A Quantitative Analysis of Metaphor Production and Interpretation in Japanese Using a Multimodal Task <i>Brian Birdsell, Natsuko Tatsuta and Hiroaki Nakamura</i>	334

Multiple Pivots in Statistical Machine Translation for Low Resource Languages <i>Sari Dewi Budiwati and Masayoshi Arimitsu</i>	345
Semi-supervised learning for all-words WSD using self-learning and fine-tuning <i>Rui Cao, Jing Bai, Wen Ma and Hiroyuki Shinnou</i>	356
A Reinforced Improved Attention Model for Abstractive Text Summarization <i>Yu Chang, Hang Lei, Xiaoyu Li and Yiming Huang</i>	362
Semantic Distance and Creativity in Linguistic Synaesthesia <i>Emmanuele Chersoni, Francesca Strik Lievers and Chu-Ren Huang</i>	370
Investigating Mandarin Negative Terms: An Evaluation of Semantic-Pragmatic Meanings and Metaphorical Mechanisms <i>Siaw-Fong Chung, Yi-Ling Tseng, Heng-Chia Liao and Man-Hua Huang</i>	379
Mapping distributional to model-theoretic semantic spaces: a baseline <i>Franck Dernoncourt</i>	388
Intrinsic Evaluation of Grammatical Information within Word Embeddings <i>Daniel Edmiston and Taek Kim</i>	395
A Continuation-based Analysis of Contrastive <i>Wa</i> in Japanese <i>Hitomi Hirayama</i>	405
Effects of Prosodic Focus on Voice Onset Time (VOT) in Chongming Chinese <i>Yitian Hong, Si Chen, Yike Yang and Bei Li</i>	414
Web Page Segmentation for Non Visual Skimming <i>Judith Jeyafreeda, Stéphane Ferrari, Fabrice Maurel, Gaël Dias and Emmanuel Giguet</i>	423
Automatic Speech Act Classification of Korean Dialogue based on the Hierarchical Structure of Speech Act Categories <i>Youngeun Koo, Jiyouon Kim and Munpyo Hong</i>	432
Investigation of Mandarin Clickbait Headlines: A Case Study of <i>Biàn Zhèyàng</i> <i>Chi-Ling Lee, Siaw-Fong Chung and Hui-Wen Liu</i>	442
On the “Easy” Task of Evaluating Chinese Irony Detection <i>An-Ran Li, Emmanuele Chersoni, Rong Xiang, Chu-Ren Huang and Qin Lu</i>	452
Towards Better Ad Experience: Click Prediction Leveraging Sequential Networks Derived Specifically From User Search Behaviors <i>Shengzhe Li, Tomoko Izumi, Yu Kuratake, Jiali Yao, Jerry Turner, Daisuke Kawahara and Sadao Kurohashi</i>	461

Cantonese turn-initial particles: annotation of discourse-interactional functions in dialog corpora <i>Andreas Liesenfeld</i>	471
Are TERRORISM and kongbu zhuyi translation equivalents? A corpus-based investigation of meaning, structure and alternative translations <i>Lily Lim</i>	480
L1 and L2 Processing of Chinese Separable VO Compounds <i>Junghwan Maeng</i>	488
Syntax and Semantics of Adjectives in Cape Verdean Creole: A View from Markedness <i>Chigusa Morita and Miki Obata</i>	496
Japanese Daily Utterance Styles: A Factor Analysis based on Balanced Corpus <i>Hajime Murai</i>	503
A Speaker Accent Recognition System for Filipino Language <i>Batman Odulio, Justin Raphael Ariaso, Karl Adrian Cruz, Mico Ian Orjalo, Ramon Rodriguez, Angelica Dela Cruz and Manolito Octaviano Jr</i>	511
A corpus-based investigation of collexemes for active-passive alternation in the English part of an English-Japanese parallel corpus <i>Masanori Oya</i>	516
Korean-to-Chinese Machine Translation using Chinese Character as Pivot Clue <i>Jeonghyeok Park and Hai Zhao</i>	522
Adapting Neural Machine Translation for English-Vietnamese using Google Translate system for Back-translation <i>Nghia Luan Pham and Van Vinh Nguyen</i>	531
Re-unifying Floating Numeral Quantifiers and Secondary Predicates in Japanese <i>Hideaki Yamashita</i>	540