Proceedings of the 31st Pacific Asia Conference on Language, Information and Computation

November 16-18, 2017
University of the Philippines Cebu
Cebu City, Philippines

Rachel Edita Roxas
Editor

©2017 PACLIC 31 (2017) 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 the National University (Philippines)
551 M.F. Jhocson St., Sampaloc, Manila
1008 Metro Manila

Acknowledgement
PACLIC 31 (2017) is organized by the University of the Philippines Cebu, National University (Philippines), Computing Society of the Philippines – Special Interest Group on Natural Language Processing, and the Linguistic Society of the Philippines.

Foreword

It is with great pride that the Philippines hosted the 31st Pacific Asia Conference on Language, Information and Communication (or PACLIC) from November 16 to 18, 2017 held at the University of the Philippines Cebu.

A warm welcome to all the participants from over the world who came to exchange knowledge and ideas on language. Following the long tradition of PACLIC conferences, PACLIC 31 (2017) emphasizes the synergy of theoretical frameworks and processing of natural language, providing a forum for researchers from different fields to share and discuss progress in scientific studies, development and application of the topics related to the study of languages. Included in the proceedings are 2 keynote speeches, 3 invited talks, 26 oral papers, and 22 poster papers.

Thank you to our organizers. Our local organizer, University of the Philippines Cebu headed by Chancellor Atty Liza D. Corro and her able team. And our other organizers namely: National University (Philippines), the Computing Society of the Philippines – Special Interest Group on Natural Language Processing, and the Linguistic Society of the Philippines.

Gratitude is also expressed to our conference co-chairs who have labored towards the successful implementation of this conference: Robert Roxas of University of the Philippines Cebu, Nathaniel Oco of National University, and Shirley Dita of De La Salle University. We express our thanks to the members of our Program Committee who reviewed the many papers that were submitted to us.

Warm gratitude is also extended to our sponsors who willingly gave their support to this event.

To those who worked behind the scenes to make this conference possible, I would also like to extend my heartfelt thanks and until we host PACLIC in the Philippines again.

Rachel Edita Roxas
Conference Chair, PACLIC 31 (2017)

Steering Committee

Hee-Rahk Chae, Hankuk University of Foreign Studies
Chu-Ren Huang, Hong Kong Polytechnic University
Rachel Edita O. Roxas, National University (Philippines)
Maosong Sun, Tsinghua University
Benjamin T'sou, City University of Hong Kong
Kei Yoshimoto, Tohoku University
Min Zhang, Soochow University

Conference Chair

Rachel Edita Roxas, National University (Philippines)

Co-Chairs

Robert Roxas, University of the Philippines Cebu
Nathaniel Oco, National University (Philippines)
Shirley Dita, De La Salle University

Local Organizing Committee

Dhong Fhel Gom-os, University of the Philippines Cebu
Rubelito Abella, University of the Philippines Cebu
Maricris Camatang, University of the Philippines Cebu
Michelle Rodriguez, University of the Philippines Cebu

Program Committee

Shirley Dita, De La Salle University (co-chair)
Qin Lu, Hong Kong Polytechnic University (co-chair)
Alice Mae Arbon, Silliman University
Aireen Arnuco, De La Salle University
Stephane Bressan, National University of Singapore
Hee-Rahk Chae, Hankuk University of Foreign Studies
Kuang-Hua Chen, National Taiwan University
Hsin-Hsi Chen, National Taiwan University
Jin-Woo Chung, Korea Advanced Institute of Science and Technology
Siaw-Fong Chung, National Chengchi University
Guohong Fu, Heilongjiang University
Helena Hong Gao, Nanyang Technological University
Wei Gao, Qatar Computing Research Institute
Yasunari Harada, Waseda University
Choochart Haruechaiyasak, National Electronics and Computer Technology Center
Munpyo Hong, Sungkyunkwan University
Jong-Bok Kim, Kyung Hee University
Mac Kim, Commonwealth Scientific and Industrial Research Organisation's Data61
Oi Yee Kwong, Chinese University of Hong Kong
Yong-Hun Lee, Chungnam National University
Dongsik Lim, Hongik University
Erlyn Manguilimotan, Weathernews, Inc.
Miren Morales, De La Salle University
Yoshiki Mori, Universiy of Tokyo
Natchanan Natpratan, Kasetsart University
Ponrudee Netisopakul, King Mongkut's Institute of Technology Ladkrabang
Robertus Nugroho, Commonwealth Scientific and Industrial Research Organisation's Data61
Nathaniel Oco, National University (Philippines)
Chutamanee Onsuwan, Thammasat University
Ryo Otoguro, Waseda University
Cecile Paris, Commonwealth Scientific and Industrial Research Organisation
Jong Park, Korea Advanced Institute of Science and Technology
Rodolfo Raga, Jose Rizal University
Robert Roxas, University of the Philippines Cebu
Rachel Edita Roxas, National University (Philippines)
Samira Shaikh, University of North Carolina at Charlotte
Virach Sornlertlamvanich, Thammasat University
Keh-Yih Su, Academia Sinica
Michael Tanangkingsing, National Taiwan University
Aline Villavicencio, Federal University of Rio Grande do Sul
Chang Xu, Commonwealth Scientific and Industrial Research Organisation's Data61
Cheng-Zen Yang, Yuan Ze University
Satoru Yokoyama, Chiba Institute of Science
Liang-Chih Yu, Yuan Ze University

Jiajun Zhang, Institute of Automation Chinese Academy of Sciences
Hai Zhao, Shanghai Jiao Tong University
Michael Zock, Centre Nationnal de la Recherche Scientifique

Table of Contents

Modeling Answering Strategies for the Polar Questions across Languages
Jong-Bok Kim 2

International Speech Communication Association Distinguished Lecture: Principles and Design of a System for Academic Information Retrieval based on Human-Machine Dialogue
Hiroya Fujisaki 10

Assessing Authenticity in Media Englishes and the Englishes of Popular Culture
Andrew Moody 11

Corpus Linguistic Analysis for Language Planning
Joel Ilao 12

Modality Markers in Cebuano and Tagalog
Michael Tanangkingsing 13

Between Reading Time and Information Structure
Masayuki Asahara 15

Doubt, incredulity, and particles in Japanese falling interrogatives
Lukas Rieser 25

The Phrasal-Prepositional Verbs in Philippine English: A Corpus-based Analysis
Jennibelle Ella and Shirley Dita 34

A Type-Logical Approach to Potential Consructions in Japanese
Hiroaki Nakamura 42

Standard and nonstandard lexicon in aviation English: A corpus linguistic study
Ramsey Ferrer, Jollene Empinado, Eloisa Marie Calico and Jan Yharie Floro 50

Stylometric Studies based on Tone and Word Length Motifs
Renkui Hou and Chu-Ren Huang 56

Ensemble Technique Utilization for Indonesian Dependency Parser
Arief Rahman, Kuncoro Adhiguna and Ayu Purwarianti 64

Raising to Object in Japanese: An HPSG Analysis
Akira Ohtani 72

Using Stanford Part-of-Speech Tagger for the Morphologically-rich Filipino Language
Matthew Phillip Go and Nicco Nocon 81

Unsupervised Bilingual Segmentation using MDL for Machine Translation
Bin Shan, Hao Wang and Yves Lepage 89

The Importance of Automatic Syntactic Features in Vietnamese Named Entity Recognition
Hoang Pham and Phuong Le-Hong 97

Multiple Nominative Constructions in Japanese: An Incremental Grammar Perspective
Tohru Seraku 104

BTG-based Machine Translation with Simple Reordering Model using Structured Perceptron
Hao Wang and Yves Lepage 114

Arabic-English Text Translation Leveraging Hybrid NER
Emna Hkiri, Souheyl Mallat and Mounir Zriguii 124

#ActuallyDepressed: Characterization of Depressed Tumblr Users' Online Behavior from Rules Generation Machine Learning Technique
Czarina Rae Cahutay and Aileen Joan Vicente 132

A Parallel Recurrent Neural Network for Language Modeling with POS Tags
Chao Su, Heyan Huang, Shumin Shi, Yuhang Guo and Hao Wu 140

Identifying Deception in Indonesian Transcribed Interviews through Lexical-based Approach
Tifani Warnita and Dessi Puji Lestari 148

Foreign Influence and Sound Change: A Case Study of Cantonese Alveolar Affricates
Yizhou Lan 155

Unsupervised Method for Improving Arabic Speech Recognition Systems
Mohamed Labidi, Mohsen Maraoui and Mounir Zrigui 161

Remarks on epistemically biased questions
David Yoshikazu Oshima 169

The blocking effect and Korean caki
Hyunjun Park and Haihua Pan 178

Expressing prediction and epistemicity with Korean -(ul) kes i and Mandarin Chinese hui
Eunson Yoo 187

Remarks on Denominal -Ed Adjectives
Tomokazu Takehisa 196

Subjecthood and Grammatical Relations in Korean: An Experimental Study with Honorific Agreement and Plural Copying
Ji-Hye Kim, Yong-Hun Lee and James Hye-Suk Yoon 206

A Stylistic Analysis of a Philippine Essay, "The Will of the River"
Pilar Caparas 214

A Corpus-based Analysis of Near-Synonymous Sentence-final Particles in Mandarin Chinese: "bale" and "eryi"
Xuefeng Gao and Yat-Mei Lee 222

Extracting a Lexicon of Discourse Connectives in Czech from an Annotated Corpus
Pavlína Synková, Magdaléna Rysová, Lucie Poláková and Jiří Mírovský 232

Word Learning by Young Bilinguals: Understanding the Denotation and Connotation Differences of "Cut" Verbs in English and Chinese
Keng Hwee Neo and Helena Gao 241

A corpus-based study on synesthesia in Korean ordinary language
Charmhun Jo 249

Intrusions of Masbate Lexicon in Local Bilingual Tabloid
Cecilia Genuino and Romualdo Mabuan 255

Facebook Integration into University Classes: Opportunities and Challenges
Romualdo Mabuan and Gregorio Ebron Jr. 265

Semantic Similarity Analysis for Paraphrase Identification in Arabic Texts
Adnen Mahmoud and Mounir Zrigui 274

An Empirical Study of Language Relatedness for Transfer Learning in Neural Machine Translation
Raj Dabre, Tetsuji Nakagawa and Hideto Kazawa 282

Rule-based Reordering and Post-Processing for Indonesian-Korean Statistical Machine Translation
Candy Olivia Mawalim, Dessi Puji Lestari and Ayu Purwarianti 287

Sentence Complexity Estimation for Chinese-speaking Learners of Japanese
Jun Liu and Yuji Matsumoto 296

Conceptualizing EDUCATION in Hong Kong and China (1984-2014)
Kathleen Ahrens and Huiheng Zeng 303

Multi-dimensional Meanings of Subjective Adverbs - Case Study of Mandarin Chinese Adverb Pianpian
Mi Zhou, Yao Yao and Chu-Ren Huang 312

Morphosyntactic Analysis of the Pronominal System of Southern Alta
Marvin Abreu 319

Wh-island Effects in Korean Scrambling Constructions
Juyeon Cho 329

A Crowdsourcing Approach for Annotating Causal Relation Instances in Wikipedia
Kazuaki Hanawa, Akira Sasaki, Naoaki Okazaki and Kentaro Inui 336

Automatic Categorization of Tagalog Documents Using Support Vector Machines
April Dae Bation, Aileen Joan Vicente and Erlyn Manguilimotan 346

Lexicalization, Separation and transitivity: A comparative study of Mandarin VO compound Variations
Menghan Jiang and Chu-Ren Huang 354

Extracting Important Tweets for News Writers using Recurrent Neural Network with Attention Mechanism and Multi-task Learning
Taro Miyazaki, Shin Toriumi, Yuka Takei, Ichiro Yamada and Jun Goto 363

Tweet Extraction for News Production Considering Unreality
Yuka Takei, Taro Miyazaki, Ichiro Yamada and Jun Goto 370

Discovering Conversation Spaces in the Public Discourse of Gender Violence: a Comparative Between Two Different Contexts
Meliza De La Paz, Maria Regina Estuar and John Noel Victorino 376

Investigating Phrase-Based and Neural-Based Machine Translation on Low-Resource Settings
Hai Long Trieu, Vu Tran and Nguyen Le Minh 384

Japanese all-words WSD system using the Kyoto Text Analysis ToolKit
Hiroyuki Shinnou, Kanako Komiya, Minoru Sasaki and Shinsuke Mori 392

Revisiting Tones in Twic East Dinka
Yu-Leng Lin 400