

PACLIC 36 (2022)

**Proceedings of the 36th Pacific Asia Conference on
Language, Information and Computation**

20-22 October, 2022
De La Salle University
Manila, Philippines

©2022 PACLIC 36 (2022) 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 De La Salle University

2401 Taft Ave, Malate, Manila

1004 Metro Manila

Acknowledgement

PACLIC 36 (2022) is jointly hosted by the Department of English and Applied Linguistics (DEAL), Br. Andrew Gonzalez FSC College of Education (BAGCED), De La Salle University, Manila, Philippines, the College of Computing and Information Technologies (CCIT), National University, Manila, Philippines, and the Lasallian Institute for Development and Education Research (LIDER), De La Salle University, Manila, Philippines.

Foreword

The 36th Pacific Asia Conference on Language, Information, and Computation, PACLIC 36 (2022) is organized by the Department of English and Applied Linguistics, Br. Andrew Gonzalez FSC College of Education, De La Salle University, the National University (Philippines), and the Lasallian Institute for Development and Educational Research (LIDER), De La Salle University. This edition of the PACLIC series of conferences, as its long tradition, continues to emphasize the synergy of theoretical analysis and processing of natural language, aiming to strengthen the communication between researchers working in different fields of language study in the Asia-Pacific region as well as around the world.

Like its predecessors in 2020 and 2021, the 36th PACLIC Conference was held fully online due to the COVID-19 pandemic's continued restriction on mobility in the Philippines. We fervently hope that the local and global situation on COVID eases in the future, such that we could resume a face-to-face format of the conference. We received 164 submissions out of 25 countries in total, tallying a 62.19 percent acceptance rate. In the 102 papers accepted, 84 of which were for oral presentations and 18 for poster presentations. In addition to excellent oral and poster presentations, the conference also highlighted three plenary talks and three invited talks delivered by notable scholars in the field. We thank Kathleen Ahrens, Hanjung Lee, Stefanie Shamila Pillai, Wilkinson Daniel Wong Gonzales, and Nguyen Thi Minh Huyen.

We would also like to thank the PACLIC Organizing and Steering Committee for being the backbone of this conference, especially in light of logistical difficulties and unique challenges of the conference's online setup. We express our sincerest gratitudes to the Local Organizing Committee, Reviewers, and the Secretariat for ensuring that the conference's preparation and execution were done efficiently and smoothly. Lastly, we would like to thank Department of English and Applied Linguistics, the Br. Andrew Gonzalez FSC College of Education and the Office of the President and the Office of the Vice President Research and Innovation of the De La Salle University for all the support to this endeavor.

Shirley N. Dita

Jong-Bok Kim

Rachel Roxas

36th PACLIC Program Committee Chairs
(on behalf of the Organizing Committee)

Organizers

PACLIC 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

Local Organizing Committee

Conference Chair: Shirley N. Dita
Conference Co-Chairs: Arlene O. Trillanes
Rochelle Irene G. Lucas

Program Committee

Chairs: Jong-Bok Kim
Rachel Editha Roxas
Shirley N. Dita
Co-Chairs: Rochelle Irene Lucas
Leah Gustilo
Raymund Vitorio
Members: Sharon Joy Bulalang
Grace Cortez
Aileen Bautista

Publicity Committee

Chair: Mideth B. Abisado
Members: Angelique D. Lacasandile
Ryan Richard H. Guadaña

Registration Committee

Chair: Marilou Santos
Member: Ma. Angelica Gumangan

Logistics Committee

Chair: Frederick Talaue
Members: Philip Rentillo
Aileen Bautista

Communications Committee

Chair: Jean Purpura
Member: Eden Grace Canopio

Secretariat

Lovely Liz De Luna
Ma. Nilda Perilla
Ronnie Sabado

Program Sub-committee/Reviewers:

Muhammad Afzaal	Ramsey Ferrer	Vladimir Mariano	Shu-Ing Shyu
Ericson Alieto	Ana Cristina Fortes	Dalos Miguel	Melanie Siegel
Alice Mae Arbon	Cecilia Genuino	Eusebio Jr. Mique	Pornsiri Shinghapreecha
Aireen Arnuco	Ryn Jean Fe Gonzales	Alen Mateo Muñoz	Sanghoun Song
Chirbet Ayunon	Wilkinson Gonzales	Naonori Nagaya	Elineth Suarez
Judith Azcarraga	Yanhui Gu	Thi Minh Nguyen	Christian Sy
Nguyen Bach	Nicanor Guinto	Nathaniel Oco	Jennifer Tan-De Ramos
Aileen Bautista	Munpyo Hong	Ethel Ong	Michael Tanangkingsing
Alejandro Bernardo	Joseph Imperial	Chautamane Onsuwan	Zhivang Teng
Annie Mae Berowa	Rowland Imperial	Jong C. Park	Gina Ugalingan
Sharon Bulalang	Jeffrey ingosan	Rafael Michael Paz	Paolo Valdez
Mary Joy Canon	Jong-Bok Kim	Charmaine Ponay	Raymund Vitorio
Marvin Casalan	Valia Kordoni	Ariel Robert Ponce	Tak-Sum Wong
Jasper Kyle Catapang	Huei-Ling Lai	Nattama Pongpairoj	Hongzhi Xu
Angie Ceniza	Nguyen Le Minh	Jeanne Purpura	Cheng-Zen Yang
Maria Art Clariño	Phuong Le-Hong	Edward Jay Quinto	Aiden Yeh
Romina Grace Cortez	Yong-Hun Lee	Rodolfo Jr. Rafa	Satoru Yokoyama
Kristine de Leon	Xiaoqian Li	Reginald Neil Recario	Liang-Chih Yu
Aprillette Devanadera	Joanna Marie Lim	Philip Rentillo	
Nimfa Dimaculangan	Rochelle Irene Lucas	Rachel Roxas	
Chenee Dino-Aparicio	Romualdo Mabuan	Byong-Rae Ryu	
Shirley Dita	Alvin Malicdem	Ria Sagum	
Jennibelle Ella	Erlyn Manguilimotan	Julie Ann Salido	
Ma. Regina Estuar	Wenwei Mao	Sheena Sapuay	

Invited Talks

Keynote Speaker

Breaking through glass ceilings or opening glass doors? Addressing the challenges of equality and inclusion via metaphors

Kathleen Ahrens (The Hong Kong Polytechnic University)

Women, but not men, are faced with metaphorical glass ceilings, and in some cases, glass cliffs. These novel constructed metaphors reflect the difficulties women have in advancing in their career, particularly in terms of advancement in leadership roles in politics and in the professions. However, what is less clear is if there are gendered associations for source domains associated with conventionalized conceptual metaphors. This talk will examine this issue and present recent advances in conceptual metaphor theory that use ontologies and collocational patterns to verify source domains (Ahrens & Jiang 2020) as a prior step to examining gendered associations. Drawing on data from business and politics, I will also explain how women use these conventionalized metaphors to position themselves as leaders. I will close my talk with a discussion of how we, as scholars, can influence the creation and interpretation of metaphor use so as to engender positive social change.

Plenary Speakers

Cue reliability and motivation in grammar and language use: A new look at differential subject marking

Hanjung Lee (Sung Kyun Kwan University, South Korea)

In many languages, arguments such as subjects and objects enjoy substantial freedom in terms of the form in which they are realized: a noun or pronoun with or without a following or preceding functional particle, a bare noun, a null-form argument, etc. This talk is concerned with caseless subjects in Korean, that is, those subjects that occur without functional particles signaling case or discourse function. An interesting and challenging problem for theoretical approaches to case is that caseless and case-marked subjects are in systematic contrast as to their interpretation. This makes Korean a Differential Subject Marking (DSM) language in Aissen's (2003) terms, wherein some subjects are marked with formal particles while others are not, depending on the semantic and pragmatic features of the subject.

This talk will focus on a hitherto unexplained property of caseless-subject clauses triggering a direct perception interpretation. I will first present evidence from conversation data demonstrating that caseless subjects predominantly occur in clause types that have an agent directly identifiable in the here and now, whereas nominative-marked subjects are most productively used in clause types wherein the identification of an agent cannot be grounded in the here and now. Based on this evidence, I will propose a new account of DSM in terms of cue reliability (Rosch & Mervis 1975; Levshina 2021), arguing that the association of caseless subjects with seemingly unrelated features such as direct perception in the here and now, agentivity, definiteness, tense deficiency and a simple,thetic interpretation (the most preferred information structure associated with caseless-subject clauses) follows from an economical use of formal particles motivated by the reliability of cues for identifying or predicting the grammatical, referential, or discourse status of an argument NP or a clause containing it.

When such cues are weak, sentences are likely to convey information that is less predictable and redundant; speakers prefer to mark the subject by overt particles in such sentences, as using case markers would lead to more uniform information density than leaving them unmarked.

These preliminary results support efficiency-based accounts of patterns of grammar, and underscore the importance of communicative efficiency in explaining and motivating grammatical rules and language use (Hawkins 2004, Haspelmath 2008, Jaeger 2010, Lee 2010, 2016, 2021, Lestrade & de Hoop 2016, Levshina 2021).

Building the next generation WordNet for Filipino with sense embeddings and network science

Briane Paul V. Samson (De La Salle University, Manila, Philippines)

The low-resource and highly morphological setting of Philippine languages is a challenge in developing a word representation or a language model. The current linguistic resources lack in rich semantic data that is crucial in most NLP tasks, and the fast-paced evolution and adaptation of Philippine languages make things even more difficult in creating a well defined language model. Thus, building a Filipino WordNet is crucial in advancing the landscape of Filipino NLP. With the vast amount of data in many digital platforms that can represent different domains and varieties of words through time including changes in its semantic and syntactic forms, we aim to create word representations of the Filipino language that is temporal and context aware and store them in the expanded Filipino WordNet. Given that languages continue to evolve and adapt, we also investigate the diachronic emergence and semantic shifts of word senses across different contexts and media, especially for the low-resource Filipino and Philippine English languages.

Invited Lectures

BibePortMal: A mobile app dictionary for Melaka Portuguese

Stefanie Shamila Pillai (University of Malaya, Malaysia)

Melaka Portuguese (also known as Papiá Cristang) has its roots in the arrival of the Portuguese in Melaka in the 16th Century. The language is still spoken by the descendants of unions between the Portuguese and locals especially in Melaka, but the dwindling number of fluent speakers and the lack of intergenerational transmission has led to this language being classified as one of the many endangered languages in Malaysia (Pillai, Soh, & Kajita, 2014). Several community engagement projects have been undertaken to translate research into efforts to revitalise Melaka Portuguese (Pillai, Phillip, & Soh, 2017). Among the efforts to encourage the use of the language and to assist the teaching of the language by a community member, a Melaka Portuguese-English dictionary in the form of a mobile application, BibePortMal, was developed. In this session I will talk about the rationale and the process of developing the application with community representatives. Some of the challenges in its development will also be discussed along with plans to improve the application.

Harnessing the power of social media in linguistic analysis: A diachronic and sociolinguistic study of Philippine English(es) using the Twitter Corpus of Philippine Englishes

Wilkinson Daniel Wong Gonzales (The Chinese University of Hong Kong)

Research on contemporary Philippine English remain relatively scarce and inadequate in comparison to research on other varieties such as American English and Singapore English, partially due to the lack of large-scale, organized, publicly available data sets that allow comprehensive and in-depth investigations of the variety. Responding to this demand, I introduce the Twitter Corpus of Philippine Englishes (TCOPE) – a 135-million-word corpus created from roughly 27 million tweets sampled from 29 major cities in the Philippine archipelago. In the first part of the talk, I provide an overview: I discuss the considerations that went into TCOPE’s design, the compilation procedure, the format, and how interested individuals can access the corpus. Then, I illustrate the utility of the corpus by showcasing how it can be used to insightfully examine the linguistic features of Philippine English as well as the relationship between these features and socio-temporal factors (e.g., ethno-geographic region, time, age, sex), focusing on four documented Philippine English features: (1) the use of irregular past tense morpheme -t, (2) double comparatives, (3) subjunctive were in subordinate counterfactual clauses, and (4) the phrasal verb base from. My initial explorations confirm patterns observed in previous research but go further to show the multifaceted and dynamic nature of Philippine English, providing empirical support for the theory that Philippine English is at the final stage of Schneider’s dynamic model. Because of its large size, sampling distribution, and its availability in different corpus formats, TCOPE can be used to investigate features in ‘general’ contemporary Philippine English as well as different types of variation, particularly diachronic and ethno-geographic variation – a feat that might not be possible with other Philippine English corpora. In combination with other existing corpora, TCOPE has the potential to broaden horizons in the diachronic and sociolinguistic study of Philippine English(es).

Towards universal syntactic-semantic resources for Vietnamese

Nguyen Thi Minh Huyen (VNU University of Science, Hanoi, Vietnam)

Despite the fact that Vietnamese is spoken by around 100 million persons all over the world, it remains a low resource language in terms of gold datasets for natural language processing (NLP). Although many NLP applications have been developed in recent years thanks to the emergence of deep learning and embedding methods, it is important to build sustainable linguistic resources, using sophisticated linguistic annotation frameworks for the Vietnamese language, in harmony with universal frameworks. In this talk, I will present our work on the construction of Vietnamese syntactic-semantic resources, including a VerbNet-based lexicon and annotated corpora for syntactic and semantic parsing.

Table of Contents

Integrating Label Attention into CRF-based Vietnamese Constituency Parser	1
<i>Duy Vu-Tran, Phu-Thinh Pham, Duc Do, An-Vinh Luong and Dien Dinh</i>	
Question-Answer Pairing from IM Conversations via Message Merging and Reply-to Prediction	10
<i>Thamolwan Poopradubsil and Chia-Hui Chang</i>	
The Representation of Discontinuity and the Correspondence Principle	20
<i>Ratna Nirupama and Prakash Mondal</i>	
Analysis of Well-being in Some Visayan Languages	30
<i>Randolph Catungal</i>	
The Information Packaging of the Do-Constructions in Chinese, Russian, and Czech	35
<i>Kawai Chui, Hsiang-Lin Yeh and Shih-Hui Lin</i>	
Exploring metaphorical polysemy with Multiple Correspondence analysis: A corpus-based study on the predicative hei ‘black’ in Chinese	45
<i>Jinmeng Dou and Meichun Liu</i>	
Sentiment Analysis in Code-Mixed Vietnamese-English Sentence-level Hotel Reviews	54
<i>Thin Dang Van, Hao Duong Ngoc and Ngan Nguyen Luu-Thuy</i>	
A Contrastive Corpus-based Analysis of Rhetoric in Asian, American, and European Low-Cost Carriers’ Slogans	62
<i>Ramsey Ferrer</i>	
Cross-strait Variations on Two Near-synonymous Loanwords xie2shang1 and tan2pan4: A Corpus-based Comparative Study	73
<i>Yueyue Huang and Chu-Ren Huang</i>	
How syntactic analysis influences the calculation of mean dependency distance: Evidence from the enhanced dependency representation	83
<i>Tsy Yih, Jianwei Yan and Haitao Liu</i>	
Chat-log Disentanglement via Same-Thread Classification and Direct-Reply Prediction	93
<i>Chia-Hui Chang, Zhi-Xian Liu, Thamolwan Poopradubsil, Yu-Ching Liao and Yu-Hao Wu</i>	
Japanese Named Entity Recognition from Automatic Speech Recognition Using Pre-trained Models	102
<i>Seiichiro Kondo, Naoya Ueda, Teruaki Oka, Masakazu Sugiyama, Asahi Hentona and Mamoru Komachi</i>	
Improving Automatic Evaluation of Acceptability Based on Language Models with a Coarse Sentence Representation	109
<i>Vijay Daultani and Naoaki Okazaki</i>	
Text Classification Using a Graph Based on Relationships Between Documents	119
<i>Hiromu Nakajima and Minoru Sasaki</i>	
Speaker Identification of Quotes in Japanese Novels based on Gender Classification Model by BERT	126
<i>Yuki Zenimoto and Takehito Utsuro</i>	
Developing and Evaluating a Dataset for How-to Tip Machine Reading at Scale	137
<i>Fuzhu Zhu, Shuting Bai, Tingxuan Li and Takehito Utsuro</i>	
Incorporating Linguistic Knowledge for Abstractive Multi-document Summarization	147

<i>Congbo Ma, Wei Emma Zhang, Hu Wang, Shubham Gupta and Mingyu Guo</i>	
VART: Vocabulary Adapted BERT Model for Multi-label Document Classification	157
<i>Zhongguang Zheng, Lu Fang and Yiling Cao</i>	
Are Emoji, Sentiment, and Emotion Friends? A Multi-task Learning for Emoji, Sentiment, and Emotion Analysis	166
<i>Gopendra Vikram Singh, Dushyant Singh Chauhan, Mauajama Firdaus, Asif Ekbal and Pushpak Bhat-tacharyya</i>	
Annotation and Multi-modal Methods for Quality Assessment of Multi-party Discussion	175
<i>Tsukasa Shiota and Kazutaka Shimada</i>	
Impact of Distance Measures on Urdu Document Clustering	183
<i>Zarmeen Nasim and Sajjad Haider</i>	
Zero-shot and few-shot approaches for tokenization, tagging, and dependency parsing of Tagalog text	190
<i>Angelina Aquino and Franz de Leon</i>	
Syntactic Analysis of Particle -kah in Malay Polar Questions	203
<i>Serena Pik Wun Jong and Lawrence Yam Leung Cheung</i>	
A distinctive collexeme analysis of near-synonym constructions -dang/ying-gai + verb	210
<i>Zhuo Zhang, Meichun Liu and Dingxuan Zhou</i>	
A Critical Discourse Analysis of Philippine Political Advertising	220
<i>Heherson Valdez and Jayson Mark Colcol</i>	
Morphological Paradigm of Nouns and Verbs: Meaning and Functions in Bisakol, a Philippine-type language	234
<i>Ana Cristina Fortes</i>	
Can the Translation Memory Principle Benefit Neural Machine Translation? A Series of Extensive Experiments with Input Sentence Annotation	243
<i>Yaling Wang and Yves Lepage</i>	
To Post or Not to Post: Exploring Students' Writing Strategies in a COVID-stricken Classroom . .	253
<i>Jennifer Diamante, Romualdo Mabuan and Emmalyn Venturillo</i>	
Voices in Cyberspace: A Move-Step Analysis of Politics-centered Online News Editorials	263
<i>Melojie Lauron and Yva Dy</i>	
Probable Human Knowledge Elicitation for Efficient Human-in-the-Loop Text Classification with-out Any Label-Related Data	272
<i>Osamu Saisho</i>	
SMTCE: A Social Media Text Classification Evaluation Benchmark and BERTology Models for Vietnamese	282
<i>Luan Nguyen, Kiet Nguyen and Ngan Nguyen</i>	
Word Formation Processes in Masbatenyo	292
<i>Romualdo Mabuan, Shirley Dita and Michael Tanangkingsing</i>	
Arabic Named Entity Recognition Using Variant Deep Neural Network Architectures and Combi-natorial Feature Embedding Based on CNN, LSTM and BERT	302
<i>Manel Affi and Chiraz Latiri</i>	

Reputation Analysis Using Key Phrases and Sentiment Scores Extracted from Reviews	313
<i>Yipu Huang, Minoru Sasaki and Kanako Komiya</i>	
Perceptual Overlap in Classification of L2 Vowels: Australian English Vowels Perceived by Experienced Mandarin Listeners	317
<i>Yizhou Wang, Rikke L. Bundgaard-Nielsen, Brett J. Baker and Olga Maxwell</i>	
Sinhala Sentence Embedding: A Two-Tiered Structure for Low-Resource Languages	325
<i>Gihan Weeraprameshwara, Vihanga Jayawickrama, Nisansa de Silva and Yudhanjaya Wijeratne</i>	
V-collocates with will and be going to: A Corpus-based Analysis	337
<i>Jarry Chia-Wei Chuang</i>	
Empirical Evaluation of Language Agnostic Filtering of Parallel Data for Low Resource Languages	346
<i>Praveen Dakwale, Talaat Khalil and Brandon Denis</i>	
Construction of Japanese BERT with Fixed Token Embeddings	356
<i>Arata Suganami and Hiroyuki Shinnou</i>	
Crosslinguistic Influence on VOT Spectrum: A comparative study on English, Mandarin and Min	362
<i>Jarry Chia-Wei Chuang</i>	
A Study of Re-generating Sentences Given Similar Sentences that Cover Them on the Level of Form and Meaning	369
<i>Hsuan-Wei Lo, Yifei Zhou, Rashel Fam and Yves Lepage</i>	
Vocabulary expansion of compound words for domain adaptation of BERT	379
<i>Hirotaaka Tanaka and Hiroyuki Shinnou</i>	
A comparison of the validity of measurement methods of the general English proficiency by dictation and read-aloud performance	388
<i>Katsunori Kotani and Takehiko Yoshimi</i>	
Curve-fitting of frequency distributions of dependency distances in a multi-lingual parallel corpus	396
<i>Masanori Oya</i>	
Measuring public opinion on the import of US pork in Taiwan	405
<i>Yu-Lun Huang, Yu-Yun Chang, Jyi-Shane Liu and Chang-Yi Lin</i>	
Synthesis and Evaluation of a Domain-specific Large Data Set for Dungeons & Dragons	415
<i>Akila Peiris and Nisansa de Silva</i>	
From Frying to Speculating: Google Ngram evidence to the meaning development of ‘?’ in Mandarin Chinese	425
<i>Jing Chen and Chu-Ren Huang</i>	
A Model-Theoretic Formalization of Natural Language Inference Using Neural Network and Tableau Method	430
<i>Ayahito Saji, Yoshihide Kato and Shigeki Matsubara</i>	
Word Sense Disambiguation of Corpus of Historical Japanese Using Japanese BERT Trained with Contemporary Texts	438
<i>Kanako Komiya, Nagi Oki and Masayuki Asahara</i>	
Gain-framed Buying or Loss-framed Selling? The Analysis of Near Synonyms in Mandarin in Prospect Theory	447
<i>Xin Luo and Chu-Ren Huang</i>	

GPT-2 Contextual Data Augmentation for Word Sense Disambiguation	455
<i>Rakia Saidi, Fethi Jarray, Jeongwoo Jay Kang and Didier Schwab</i>	
Theorizing on Meaning-Making in Classroom Interaction Using Presuppositions-To-Theory (C-PRETTY) Research	463
<i>Janeson Miranda and Rempson Gipaya</i>	
Taste of Wine or Taste of a Person: (Synesthetic) Metaphors in Wine Reviews	475
<i>Yin Zhong and Wing Sum Tse</i>	
Contextual-Boosted Deep Neural Collaborative Filtering Approach for Arabic Textual Documents Recommendation	485
<i>Ons Meddeb, Mohsen Maraoui and Mounir Zrigui</i>	
Learning Sentence Embeddings In The Legal Domain with Low Resource Settings	494
<i>Sahan Jayasinghe, Lakith Rambukkanage, Ashan Silva, Nisansa de Silva, Shehan Perera and Madhavi Perera</i>	
How Does Structure Affect Surface Rule Application?: Accent and Verbal Morphology in Japanese Deverbal (Noun) Compounds	503
<i>Yu Tomita</i>	
Collection Methods and Data Characteristics of the PagkataoKo Dataset	513
<i>Edward Tighe, Luigi Acorda, Alexander Ii Agno, Jesah Gano, Timothy Go, Gabriel Santiago and Claude Sedillo</i>	
Novelty Detection in Community Question Answering Forums	525
<i>Tirthankar Ghosal, Vignesh Edithal, Tanik Saikh, Saprativa Bhattacharjee, Asif Ekbal and Pushpak Bhat-tacharyya</i>	
A Graph-Based Approach Leveraging Posts and Reactions for Detecting Rumors on Online Social Media	533
<i>Asimul Haque and Muhammad Abulaish</i>	
Siamese AraBERT-LSTM Model based Approach for Arabic Paraphrase Detection	545
<i>Adnen Mahmoud and Mounir Zrigui</i>	
HMIST: Hierarchical Multilingual Isometric Speech Translation using Multi-Task Learning Framework and it's influence on Automatic Dubbing	554
<i>Nidhir Bhavsar, Aakash Bhatnagar and Muskaan Singh</i>	
Bio-Medical Multi-label Scientific Literature Classification using LWAN and Dual-attention module	564
<i>Deepanshu Khanna, Aakash Bhatnagar, Nidhir Bhavsar, Muskaan Singh and Petr MotliceK</i>	
An Empirical Comparison of off-the-shelve Semantic Similarity methods for down-streaming Meeting Similarity	572
<i>Aditya Upadhyay, Aakash Bhatnagar, Nidhir Bhavsar and Muskaan Singh</i>	
An End-to-End Multilingual System for Automatic Minuting of Multi-Party Dialogues	582
<i>Aakash Bhatnagar, Nidhir Bhavsar, Muskaan Singh and Petr MotliceK</i>	
Law Retrieval with Supervised Contrastive Learning Using the Hierarchical Structure of Law . . .	590
<i>Jungmin Choi, Ukyo Honda, Taro Watanabe, Hiroki Ouchi and Kentaro Inui</i>	
Approaching Neural Chinese Word Segmentation as a Low-Resource Machine Translation Task .	600
<i>Pinzhen Chen and Kenneth Heafield</i>	

Attitudes towards pedagogical code-switching: A verbal guise approach	607
<i>Andrea Althea G. Aballa, Chyme Fresca M. Colinares, Danielle Majella C. Milanese, John Chael D. Wang, Edward Jay M. Quinto and John Christopher D. Castillo</i>	
Syllabified Sequence-to-sequence attention for transliteration	617
<i>G. Vyshnavi, G. Sridevi, P. Krishna Reddy and M. Ritesh</i>	
The Relationship between L2 Grit and Intrinsic Reading Motivation of Filipino Pre-service Teachers in Central Mindanao	626
<i>Jannie Faye Nacario, King Arman Calingasan, Omar Pendi and Bai Sittie Maguing</i>	
Pronunciation of English Words with /th/ Sounds among Senior High School Learners	637
<i>Jeremiah Saavedra</i>	
SEND: A Simple and Efficient Noise Detection Algorithm for Vietnamese Real Estate Posts	647
<i>Khanh Quoc Tran, An Tran-Hoai Le, An Trong Nguyen, Tung Tran Nguyen Doan, Son Huynh Thanh, Bao Le Hoang, Hoang Nguyen Minh, Triet Minh Thai, Hoang Le Huy, Dang T. Huynh, Binh T. Nguyen, Nhi Y. T. Ho and Trung T. Nguyen</i>	
Gunita: Visualizing the evolution of the Philippines' languages with a historical Philippine text corpora	657
<i>Amiel Bornales, Jonn Yuvallos and Courtney Ngo</i>	
Multi-Word Verbs in Senior High School Academic Papers: A Corpus-Based Study	664
<i>Ma. Angelica Gumangan and Shirley Dita</i>	
Early writing of a bilingual child: A content analysis of his YouTube video comments	674
<i>Bernard Barruga</i>	
Metalinguistic Corrective Feedback and Students' Response to Feedback in L2 Writing	682
<i>Pilar Caparas</i>	
Automatic Minuting: A Pipeline Method for Generating Minutes from Multi-Party Meeting Proceedings	691
<i>Kartik Shinde, Tirthankar Ghosal, Muskaan Singh and Ondrej Bojar</i>	
Vanilla Recurrent Neural Networks for Interpretable Semantic Textual Similarity	703
<i>Piotr Andruskiewicz and Barbara Rychalska</i>	
Filipino Undergraduates' Perceptions of Translanguaging in a Linguistically Diverse Context	713
<i>King Arman Calingasan, Earl Emunahlyn Joice Erlano, Haina Salik and Abdullaken Sinagandal</i>	
Zamboanga Chavacano Verbal Aspects: Superstrate and Substrate Influences in Morphosyntactic Behavior	723
<i>Abee Eijansantos, Jeric Ventoza, Rochelle Irene Lucas and Ericson Alieto</i>	
Philippine English Proficiency of the K12 students: Basis for the improvement of the English Curriculum	733
<i>Lemuel Fontillas, Sherrilyn Quintos and Cynthia Ronquillo</i>	
Verify: Breakthrough accuracy in the Urdu fake news detection using Text classification	740
<i>Santosh Na, Parkash Na, Toto Na and Dr. Zarmeem Nasim</i>	
Extractive Text Summarization with Latent Topics using Heterogeneous Graph Neural Network	749
<i>Tuan Anh Phan, Ngoc Dung Nguyen and Khac-Hoai Nam Bui</i>	
Tweet Review Mining focusing on Celebrities by Machine Reading Comprehension based on BERT	757

<i>Yuta Nozaki, Kotoe Sugawara, Yuki Zenimoto and Takehito Utsuro</i>	
Aspect-based Sentiment Analysis for Vietnamese Reviews about Beauty Product on E-commerce Websites	767
<i>Quang-Linh Tran, Phan Thanh Dat Le and Trong-Hop Do</i>	
Intent Detection and Slot Filling from Dependency Parsing Perspective: A Case Study in Vietnamese	777
<i>Phu-Thinh Pham, Duy Vu-Tran, Duc Do, An-Vinh Luong and Dien Dinh</i>	
Annotating Entity and Causal Relationships on Japanese Vehicle Recall Information	783
<i>Hsuan-Yu Kuo, Youmi Ma and Naoaki Okazaki</i>	
A Deep Learning based Framework for Image Paragraph Generation in Hindi	792
<i>Santosh Kumar Mishra, Sushant Sinha, Sriparna Saha and Pushpak Bhattacharyya</i>	
UIT-ViCoV19QA: A Dataset for COVID-19 Community-based Question Answering on Vietnamese Language	801
<i>Triet Thai, Ngan Chu Thao-Ha, Anh Vo and Son Luu</i>	
Vowel sequences in Old Japanese: from a corpus-based approach	811
<i>Chihkai Lin</i>	
Effectiveness Analysis of Word Sense Disambiguation Using Example of Word Senses from WordNet	818
<i>Hiroshi Sekiya and Minoru Sasaki</i>	
Effective Use of Japanese Dictionary Definition Sentences in Learning Hierarchical Embedding of Dictionaries	826
<i>Yuki Ishii and Minoru Sasaki</i>	
Bi-directional Cross-Attention Network on Vietnamese Visual Question Answering	834
<i>Duy-Minh Nguyen-Tran, Tung Le, Minh Le Nguyen and Huy Tien Nguyen</i>	
The Development and Assessment of Pattern Matching Algorithms Used by ZEE: A Filipino Sign Language (FSL) Dictionary and English-Learning App	842
<i>Maria Doinique Daculan and Asherrie Jaye Tan</i>	
Hate speech detection: a comparison of mono and multilingual transformer model with cross-language evaluation	853
<i>Koyel Ghosh and Dr. Apurbalal Senapati</i>	