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:

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


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

*Brianne 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
Duy Vu-Tran, Phu-Thinh Pham, Duc Do, An-Vinh Luong and Dien Dinh

Question-Answer Pairing from IM Conversations via Message Merging and Reply-to Prediction 10
Thamolwan Poopradubsil and Chia-Hui Chang

The Representation of Discontinuity and the Correspondence Principle ............................ 20
Ratna Nirupama and Prakash Mondal

Analysis of Well-being in Some Visayan Languages ......................................................... 30
Randolph Catungal

The Information Packaging of the Do-Constructions in Chinese, Russian, and Czech ................. 35
Kawai Chui, Hsiang-Lin Yeh and Shih-Hui Lin

Exploring metaphorical polysemy with Multiple Correspondence analysis: A corpus-based study
on the predicative hei ‘black’ in Chinese ................................................................. 45
Jinmeng Dou and Meichun Liu

Sentiment Analysis in Code-Mixed Vietnamese-English Sentence-level Hotel Reviews ............. 54
Thin Dang Van, Hao Duong Ngoc and Ngan Nguyen Luu-Thuy

A Contrastive Corpus-based Analysis of Rhetoric in Asian, American, and European Low-Cost
Carriers’ Slogans ................................................................. 62
Ramsey Ferrer

Cross-strait Variations on Two Near-synonymous Loanwords xie2shang1 and tan2pan4: A Corpus-
based Comparative Study ................................................................. 73
Yueyue Huang and Chu-Ren Huang

How syntactic analysis influences the calculation of mean dependency distance: Evidence from the
enhanced dependency representation ................................................................. 83
Tsy Yih, Jianwei Yan and Haitao Liu

Chat-log Disentanglement via Same-Thread Classification and Direct-Reply Prediction ............. 93
Chia-Hui Chang, Zhi-Xian Liu, Thamolwan Poopradubsil, Yu-Ching Liao and Yu-Hao Wu

Japanese Named Entity Recognition from Automatic Speech Recognition Using Pre-trained Models 102
Seiichiro Kondo, Naoya Ueda, Teruaki Oka, Masakazu Sugiyama, Asahi Hentona and Mamoru Komachi

Improving Automatic Evaluation of Acceptability Based on Language Models with a Coarse Sen-
tence Representation ................................................................. 109
Vijay Daultani and Naoaki Okazaki

Text Classification Using a Graph Based on Relationships Between Documents ......................... 119
Hiromu Nakajima and Minoru Sasaki

Speaker Identification of Quotes in Japanese Novels based on Gender Classification Model by BERT 126
Yuki Zenimoto and Takehito Utsuro

Developing and Evaluating a Dataset for How-to Tip Machine Reading at Scale ..................... 137
Fuzhu Zhu, Shuting Bai, Tingxuan Li and Takehito Utsuro

Incorporating Linguistic Knowledge for Abstractive Multi-document Summarization ................ 147
Congbo Ma, Wei Emma Zhang, Hu Wang, Shubham Gupta and Mingyu Guo

VART: Vocabulary Adapted BERT Model for Multi-label Document Classification ........................................ 157

Zhongguang Zheng, Lu Fang and Yiling Cao

Are Emoji, Sentiment, and Emotion Friends? A Multi-task Learning for Emoji, Sentiment, and Emotion Analysis ................................................................. 166

Gopendra Vikram Singh, Dushyant Singh Chauhan, Mauajama Firdaus, Asif Ekbal and Pushpak Bhat-tacharyya

Annotation and Multi-modal Methods for Quality Assessment of Multi-party Discussion .......................... 175

Tsukasa Shiota and Kazutaka Shimada

Impact of Distance Measures on Urdu Document Clustering ................................................................. 183

Zarmeen Nasim and Sajjad Haider

Zero-shot and few-shot approaches for tokenization, tagging, and dependency parsing of Tagalog text 190

Angelina Aquino and Franz de Leon

Syntactic Analysis of Particle -kah in Malay Polar Questions ............................................................ 203

Serena Pik Wun Jong and Lawrence Yam Leung Cheung

A distinctive collexeme analysis of near-synonym constructions -dang/ying-gai + verb .......................... 210

Zhuo Zhang, Meichun Liu and Dingxuan Zhou

A Critical Discourse Analysis of Philippine Political Advertising .................................................. 220

Heherson Valdez and Jayson Mark Colcol

Morphological Paradigm of Nouns and Verbs: Meaning and Functions in Bisakol, a Philippine-type language ................................................................. 234

Ana Cristina Fortes

Can the Translation Memory Principle Benefit Neural Machine Translation? A Series of Extensive Experiments with Input Sentence Annotation .................................. 243

Yaling Wang and Yves Lepage

To Post or Not to Post: Exploring Students’ Writing Strategies in a COVID-stricken Classroom . 253

Jennifer Diamante, Romualdo Mabuan and Emmalyn Venturillo

Voices in Cyberspace: A Move-Step Analysis of Politics-centered Online News Editorials .................. 263

Melojie Lauron and Yva Dy

Probable Human Knowledge Elicitation for Efficient Human-in-the-Loop Text Classification without Any Label-Related Data .................................................. 272

Osamu Saisho

SMTCE: A Social Media Text Classification Evaluation Benchmark and BERTology Models for Vietnamese ................................................................. 282

Luan Nguyen, Kiet Nguyen and Ngan Nguyen

Word Formation Processes in Masbateño ................................................................. 292

Romualdo Mabuan, Shirley Dita and Michael Tanangkingsing

Arabic Named Entity Recognition Using Variant Deep Neural Network Architectures and Combinatorial Feature Embedding Based on CNN, LSTM and BERT ........... 302

Manel Affi and Chiraz Latiri
Reputation Analysis Using Key Phrases and Sentiment Scores Extracted from Reviews
Yipu Huang, Minoru Sasaki and Kanako Komiya

Perceptual Overlap in Classification of L2 Vowels: Australian English Vowels Perceived by Experienced Mandarin Listeners
Yizhou Wang, Rikke L. Bundgaard-Nielsen, Brett J. Baker and Olga Maxwell

Sinhala Sentence Embedding: A Two-Tiered Structure for Low-Resource Languages
Gihan Weeraprameshwara, Vihanga Jayawickrama, Nisansa de Silva and Yudhanjaya Wijeratne

V-collocates with will and be going to: A Corpus-based Analysis
Jarry Chia-Wei Chuang

Empirical Evaluation of Language Agnostic Filtering of Parallel Data for Low Resource Languages
Praveen Dakwale, Talaat Khalil and Brandon Denis

Construction of Japanese BERT with Fixed Token Embeddings
Arata Suganami and Hiroyuki Shinno

Crosslinguistic Influence on VOT Spectrum: A comparative study on English, Mandarin and Min
Jarry Chia-Wei Chuang

A Study of Re-generating Sentences Given Similar Sentences that Cover Them on the Level of Form and Meaning
Hsuan-Wei Lo, Yifei Zhou, Rashel Fam and Yves Lepage

Vocabulary expansion of compound words for domain adaptation of BERT
Hirotaka Tanaka and Hiroyuki Shinno

A comparison of the validity of measurement methods of the general English proficiency by dictation and read-aloud performance
Katsunori Kotani and Takehiko Yoshimi

Curve-fitting of frequency distributions of dependency distances in a multi-lingual parallel corpus
Masanori Oya

Measuring public opinion on the import of US pork in Taiwan
Yu-Lun Huang, Yu-Yun Chang, Jyi-Shane Liu and Chang-Yi Lin

Synthesis and Evaluation of a Domain-specific Large Data Set for Dungeons & Dragons
Akila Peiris and Nisansa de Silva

From Frying to Speculating: Google Ngram evidence to the meaning development of ‘?’ in Mandarin Chinese
Jing Chen and Chu-Ren Huang

A Model-Theoretic Formalization of Natural Language Inference Using Neural Network and Tableau Method
Ayahito Saji, Yoshihide Kato and Shigeki Matsubara

Word Sense Disambiguation of Corpus of Historical Japanese Using Japanese BERT Trained with Contemporary Texts
Kanako Komiya, Nagi Oki and Masayuki Asahara

Gain-framed Buying or Loss-framed Selling? The Analysis of Near Synonyms in Mandarin in Prospect Theory
Xin Luo and Chu-Ren Huang
GPT-2 Contextual Data Augmentation for Word Sense Disambiguation ........................................... 455
Rakia Saidi, Fethi Jarray, Jeongwoo Jay Kang and Didier Schwab

Theorizing on Meaning-Making in Classroom Interaction Using Presuppositions-To-Theory (C-PRETTY) Research ................................................................. 463
Janeson Miranda and Rempson Gipaya

Taste of Wine or Taste of a Person: (Synesthetic) Metaphors in Wine Reviews ................................. 475
Yin Zhong and Wing Sum Tse

Contextual-Boosted Deep Neural Collaborative Filtering Approach for Arabic Textual Documents Recommendation ................................................................. 485
Ons Meddeb, Mohsen Maraoui and Mounir Zrigui

Learning Sentence Embeddings In The Legal Domain with Low Resource Settings .......................... 494
Sahan Jayasinghe, Lakith Rambukkanage, Ashan Silva, Nisansa de Silva, Shehan Perera and Madhavi Perera

How Does Structure Affect Surface Rule Application?: Accent and Verbal Morphology in Japanese Deverbal (Noun) Compounds ............................................. 503
Yu Tomita

Collection Methods and Data Characteristics of the PagkataoKo Dataset ..................................... 513
Edward Tighe, Luigi Acorda, Alexander Il Agno, Jesah Gano, Timothy Go, Gabriel Santiago and Claude Sedillo

Novelty Detection in Community Question Answering Forums .................................................. 525
Tirthankar Ghosal, Vignesh Edithal, Tanik Saikh, Saprativa Bhattacharjee, Asif Ekbal and Pushpak Bhattacharyya

A Graph-Based Approach Leveraging Posts and Reactions for Detecting Rumors on Online Social Media ......................................................................................... 533
Asimul Haque and Muhammad Abulaish

Siamese AraBERT-LSTM Model based Approach for Arabic Paraphrase Detection ....................... 545
Adnen Mahmoud and Mounir Zrigui

Nidhir Bhavsar, Aakash Bhatnagar and Muskaan Singh

Bio-Medical Multi-label Scientific Literature Classification using LWAN and Dual-attention module 564
Deepanshu Khanna, Aakash Bhatnagar, Nidhir Bhavsar, Muskaan Singh and Petr Motlicek

An Empirical Comparison of off-the-shelf Semantic Similarity methods for down-streaming Meeting Similarity ........................................................................... 572
Aditya Upadhyay, Aakash Bhatnagar, Nidhir Bhavsar and Muskaan Singh

An End-to-End Multilingual System for Automatic Minuting of Multi-Party Dialogues ............... 582
Aakash Bhatnagar, Nidhir Bhavsar, Muskaan Singh and Petr Motlicek

Law Retrieval with Supervised Contrastive Learning Using the Hierarchical Structure of Law .... 590
Jungmin Choi, Ukyo Honda, Taro Watanabe, Hiroki Ouchi and Kentaro Inui

Approaching Neural Chinese Word Segmentation as a Low-Resource Machine Translation Task . 600
Pinzhen Chen and Kenneth Heafield
Attitudes towards pedagogical code-switching: A verbal guise approach .............................. 607
Andrea Althea G. Aballa, Chyme Fresca M. Colinares, Danielle Majella C. Milanes, John Chael D. Wang, Edward Jay M. Quinto and John Christopher D. Castillo

Syllabified Sequence-to-sequence attention for transliteration ................................. 617
G. Vyshnavi, G. Sridevi, P. Krishna Reddy and M. Ritesh

The Relationship between L2 Grit and Intrinsic Reading Motivation of Filipino Pre-service Teachers in Central Mindanao ................................................................. 626
Jannie Faye Nacario, King Arman Calingasan, Omar Pendi and Bai Sittie Maguing

Pronunciation of English Words with /th/ Sounds among Senior High School Learners ........ 637
Jeremiah Saavedra

SEND: A Simple and Efficient Noise Detection Algorithm for Vietnamese Real Estate Posts .... 647
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

Gunita: Visualizing the evolution of the Philippines’ languages with a historical Philippine text corpora ................................................................. 657
Amiel Bornales, Jonn Yuvallos and Courtney Ngo

Multi-Word Verbs in Senior High School Academic Papers: A Corpus-Based Study ............ 664
Ma. Angelica Gumangan and Shirley Dita

Early writing of a bilingual child: A content analysis of his YouTube video comments ........... 674
Bernard Barruga

Metalinguistic Corrective Feedback and Students’ Response to Feedback in L2 Writing ........ 682
Pilar Caparas

Automatic Minuting: A Pipeline Method for Generating Minutes from Multi-Party Meeting Proceedings ................................................................. 691
Kartik Shinde, Tirthankar Ghosal, Muskaan Singh and Ondrej Bojar

Vanilla Recurrent Neural Networks for Interpretable Semantic Textual Similarity ................. 703
Piotr Andruszkiewicz and Barbara Rychalska

Filipino Undergraduates’ Perceptions of Translanguaging in a Linguistically Diverse Context ................................................................. 713
King Arman Calingasan, Earl Emunahlyn Joice Erlano, Haina Salik and Abdullaken Sinagandal

Zamboanga Chavacano Verbal Aspects: Superstrate and Substrate Influences in Morphosyntactic Behavior ................................................................. 723
Abee Ejiansantos, Jeric Ventoza, Rochelle Irene Lucas and Ericson Alieto

Filipino English Proficiency of the K12 students: Basis for the improvement of the English Curriculum ................................................................. 733
Lemuel Fontillas, Sherrilyn Quintos and Cynthia Ronquillo

Verify: Breakthrough accuracy in the Urdu fake news detection using Text classification ....... 740
Santosh Na, Parkash Na, Toto Na and Dr. Zarmeen Nasim

Extractive Text Summarization with Latent Topics using Heterogeneous Graph Neural Network ................................................................. 749
Tuan Anh Phan, Ngoc Dung Nguyen and Khac-Hoai Nam Bui

Tweet Review Mining focusing on Celebrities by Machine Reading Comprehension based on BERT 757
Yuta Nozaki, Kotoe Sugawara, Yuki Zenimoto and Takehito Utsuro
Aspect-based Sentiment Analysis for Vietnamese Reviews about Beauty Product on E-commerce Websites .................................................. 767

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