

35th Pacific Asia Conference on Language, Information and Computation (PACLIC35)

Shanghai, China
5-7 November 2021

ISBN: 978-1-7138-9188-8

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2021) by PACLIC 35 Organizing Committee, PACLIC Steering Committee, and the authors
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Regular Papers

| | |
|--|-----|
| Under the Microscope: Interpreting Readability Assessment Models for Filipino | 1 |
| <i>Joseph Marvin Imperial and Ethel Ong</i> | |
| SEPP: Similarity Estimation of Predicted Probabilities for Defending and Detecting Adversarial Text | 11 |
| <i>Hoang-Quoc Nguyen-Son, Seira Hidano, Kazuhide Fukushima and Shinsaku Kiyomoto</i> | |
| A Comprehensive Analysis of PMI-based Models for Measuring Semantic Differences | 21 |
| <i>Taichi Aida, Mamoru Komachi, Toshinobu Ogiso, Hiroya Takamura and Daichi Mochihashi</i> | |
| Exploring sentiment constructions: connecting deep learning models with linguistic construction . | 32 |
| <i>Shu-Kai Hsieh and Yu-Hsiang Tseng</i> | |
| A Neuro-Symbolic Approach for Question Answering on Research Articles | 40 |
| <i>Komal Gupta, Tirthankar Ghosal and Asif Ekbal</i> | |
| An Empirical Performance Analysis of State-of-the-Art Summarization Models for Automatic Minuting | 50 |
| <i>Muskaan Singh, Tirthankar Ghosal Ghosal and Ondrej Bojar</i> | |
| Data Augmentation to Address Out of Vocabulary Problem in Low Resource Sinhala English Neural Machine Translation | 61 |
| <i>Aloka Fernando and Surangika Ranathunga</i> | |
| “Prickly Voice” or “Smelly Voice”? Comprehending novel synaesthetic metaphors | 71 |
| <i>Yin Zhong and Kathleen Ahrens</i> | |
| Translation for professionals: corpus-based study of translation universals in computing | 80 |
| <i>Huang Baorong</i> | |
| Island effects in Vietnamese relativization: A formal judgment study | 89 |
| <i>Quyên Nguyen</i> | |
| Lexical Meanings of Ditransitive Verbs in Japanese | 99 |
| <i>Hiroko Kimura and Chigusa Morita</i> | |
| Examine persuasion strategies in Chinese on social media | 108 |
| <i>Yu-Yun Chang, Po-Ya Angela Wang, Han-Tang Hung, Ka-Sîng Khóo and Shu-Kai Hsieh</i> | |
| Exploring Conditional Text Generation for Aspect-Based Sentiment Analysis | 119 |
| <i>Siva Uday Sampreeth Chebolu, Franck Deroncourt, Nedim Lipka and Thamar Solorio</i> | |
| Aspect or Manner? A Study of Reduplicated Adverbials in Mandarin Chinese | 130 |
| <i>Siaw-Fong Chung and Chu-Ren Huang</i> | |
| Inter-clausal Anaphora in Chinese Conditionals: a Multi-factorial Analysis | 140 |
| <i>Shunting Chen, Pascal Amsili and Yiming Liang</i> | |
| A multidimensional analysis: speaking style of learner speech across proficiency levels | 149 |
| <i>Tingting Cheng</i> | |
| Variation across Everyday Conversations: Factor Analysis of Conversations using Semantic Cate- gories of Functional Expressions | 160 |
| <i>Yuya Chiba and Ryuichiro Higashinaka</i> | |

| | |
|---|-----|
| Dependency Enhanced Contextual Representations for Japanese Temporal Relation Classification <i>Chenjing Geng, Fei Cheng, Masayuki Asahara, Lis Kanashiro Pereira and Ichiro Kobayashi</i> | 170 |
| Applying Masked Language Models to Search for Suitable Verbs Used in Academic Writing <i>Chooi Ling Goh</i> | 180 |
| Using Distributional Principles for the Semantic Study of Contextual Language Models <i>Olivier Ferret</i> | 189 |
| A Comparative Study of Collocation Extraction Methods from the Perspectives of Vocabulary and Grammar: A Case Study in the Field of Journalism <i>Lulu Gu, Yue Pan and Pengyuan Liu</i> | 201 |
| A Corpus-based Lexical Semantic Study of Mandarin Verbs of “Tui” and “La” <i>Sun Han and Huang Chu-Ren</i> | 211 |
| Unpredictable Attributes in Market Comment Generation <i>Yumi Hamazono, Tatsuya Ishigaki, Yusuke Miyao, Hiroya Takamura and Ichiro Kobayashi</i> | 217 |
| Grammatical Error Correction via Supervised Attention in the Vicinity of Errors <i>Hikomichi Ishii, Akihiro Tamura and Takashi Ninomiya</i> | 227 |
| Can Monolingual Pre-trained Encoder-Decoder Improve NMT for Distant Language Pairs? <i>Hwichan Kim and Mamoru Komachi</i> | 235 |
| Language change in Chinese political discourse based on the relationship between sentence and clause <i>Renkui Hou, Chu-Ren Huang and Kathleen Ahrens</i> | 244 |
| Various Errors Improve Neural Grammatical Error Correction <i>Shota Koyama, Hiroya Takamura and Naoaki Okazaki</i> | 251 |
| Construction of MBTI Personality Estimation Model Considering Emotional Information <i>Ryota Kishima, Kazuyuki Matsumoto, Minoru Yoshida and Kenji Kita</i> | 262 |
| Analyzing Semantic Changes in Japanese Words Using BERT <i>Kazuma Kobayashi, Taichi Aida and Mamoru Komachi</i> | 270 |
| Automatic Analysis of Linguistic Features in Journal Articles of Different Academic Impacts with Feature Engineering Techniques <i>Siyu Lei, Ruiying Yang and Chu-Ren Huang</i> | 281 |
| HPSG/MRS-Based Natural Language Generation Using Transformer <i>Gyu-Min Lee, Sanghoun Song</i> | 291 |
| Improving Adversarial Text Generation with n-Gram Matching <i>Shijie Li and Massimo Piccardi</i> | 301 |
| Spatial-temporal attributes in verbal semantics: A corpus-based lexical semantic study of discrim- inating Mandarin near synonyms of “tui” and “la” <i>Qiangmei Liang and Chu-Ren Huang</i> | 310 |
| Span Detection for Vietnamese Aspect-Based Sentiment Analysis <i>Kim Nguyen Thi Thanh, Sieu Huynh Khai, Phuc Pham Huynh, Luong Phan Luc, Duc-Vu Nguyen and Kiet Nguyen Van</i> | 318 |
| The Paradox of Subject-Oriented Resultative Compounds in Taiwan Southern Min | 329 |

Huei-Ling Lin

| | |
|---|-----|
| Joint Chinese Word Segmentation and Part-of-speech Tagging via Two-stage Span Labeling | 336 |
| <i>Duc-Vu Nguyen, Linh-Bao Vo, Ngoc-Linh Tran, Kiet Nguyen and Ngan Nguyen</i> | |
| Machine Learning Approach for Depression Detection in Japanese | 346 |
| <i>Yuka Niimi, Yutaka Miyaji</i> | |
| KC4Align: Improving Sentence Alignment method for Low-resource Language Pairs | 354 |
| <i>Ha Nguyen Tien, Dat Nguyen Huu, Huong Le Thanh, Vinh Nguyen Van and Minh Nguyen Quang</i> | |
| Native Language Identification and Reconstruction of Native Language Relationship Using Japanese Learner Corpus | 364 |
| <i>Mitsuhiro Nishijima and Ying Liu</i> | |
| ALICE++: Adversarial Training for Robust and Effective Temporal Reasoning | 373 |
| <i>Lis Pereira, Fei Cheng, Masayuki Asahara and Ichiro Kobayashi</i> | |
| Unknown Intent Detection using Multi-Objective Optimization on Deep Learning Classifiers | 383 |
| <i>Prerna Prem, Zishan Ahmad, Asif Ekbal, Shubhashis Sengupta, Sakshi Jain and Roshini Rammani</i> | |
| Schizophrenia Through the Lens of Chinese Classifier: A Preliminary Study | 392 |
| <i>Edward Ruoyang Shi</i> | |
| Natural Language Inference using Neural Network and Tableau Method | 402 |
| <i>Ayahito Saji, Daiki Takao, Yoshihide Kato and Shigeki Matsubara</i> | |
| Talking with the Theorem Prover to Interactively Solve Natural Language Inference | 411 |
| <i>Atsushi Sumita, Yusuke Miyao and Koji Mineshima</i> | |
| Knowledge Grounded Multimodal Dialog Generation in Task-oriented Settings | 421 |
| <i>Deeksha Varshney and Anushkha Singh, Asif Ekbal</i> | |
| Study of Near Synonymous Mental-State Verbs: A MARVS Perspective | 432 |
| <i>Yuan Tao and Yuan Tao</i> | |
| Deep Speaker Verification Model for Low-Resource Languages and Vietnamese Dataset | 442 |
| <i>Dat Vi Thanh, Thanh Pham Viet and Trang Nguyen Thi Thu</i> | |
| Understanding Advanced Level Academic Writing on Syntactic Complexity | 452 |
| <i>Miao Wang and Wander Lowie</i> | |
| Investigating effects of gender identity on interpreters' styles: A case study of multidimensional analysis | 463 |
| <i>Xiaoshu Yuan, Yunzhang Shi and Hongyu Wan</i> | |
| A quantitative investigation of English adnominal modifiers | 476 |
| <i>Tsy Yih and Haitao Liu</i> | |

Poster Papers

| | |
|---|-----|
| A MARVS Analysis of Two Chinese Near-Synonymous Verbs of Jumping Based on Chinese Corpora | 483 |
| <i>Yingying Peng</i> | |
| Learning Sense-Specific Static Embeddings using Contextualised Word Embeddings as a Proxy | 493 |
| <i>Yi Zhou, Danushka Bollegala</i> | |
| Readability and Linearity | 503 |

| | |
|--|-----|
| <i>Yo Ehara</i> | |
| Covering a sentence in form and meaning with fewer retrieved sentences | 513 |
| <i>Yuan Liu and Yves LePage</i> | |
| Stronger Baseline for Robust Results in Multimodal Sentiment Analysis | 523 |
| <i>Wei Yang and Jun Ogata</i> | |
| ParaCotta: Synthetic Multilingual Paraphrase Corpora from the Most Diverse Translation Sample Pair | 533 |
| <i>Alham Fikri Aji, Tirana Noor Fatyanosa, Radityo Eko Prasajo, Philip Arthur, Suci Fitriany, Salma Qonitah, Nadhifa Zulfa, Tomi Santoso and Mahendra Data</i> | |
| Holistic interpretation in locative alternation – Evidence from self-paced reading | 543 |
| <i>Rui Cao</i> | |
| Multi-tasking Dialogue Comprehension with Discourse Parsing | 551 |
| <i>Yuchen He, Zhuosheng Zhang and Hai Zhao</i> | |
| A Corpus-based Gender Study of Hedges in Spoken British English | 562 |
| <i>Ruofan Du</i> | |
| The Annotation of Antonym Information in the ‘Word List by Semantic Principles’ | 572 |
| <i>Sachi Kato, Masayuki Asahara, Nanami Moriyama and Asami Ogiwara, Makoto Yamazaki</i> | |
| Incorporating Semantic Textual Similarity and Lexical Matching for Information Retrieval | 582 |
| <i>Hiroki Iida and Naoaki Okazaki</i> | |
| Supervised Word Sense Disambiguation on Taiwan Hakka Polysemy with Neural Network Models: A Case Study of BUN, TUNG and LAU | 592 |
| <i>Huei-Ling Lai, Hsiao-Ling Hsu, Jyi-Shane Liu, Chia-Hung Lin and Yanhong Chen</i> | |
| A Sentiment Analysis of Men’s and Women’s Speech in the BNC64 | 603 |
| <i>Yong-Hun Lee and Ji-Hye Kim</i> | |
| Optimizing Domain Specificity of Transformer-based Language Models for Extractive Summariza- tion of Financial News Articles in Korean | 611 |
| <i>Huije Lee, Wonsuk Yang, Chaehun Park, Hoyun Song, Eugene Jang and Jong C. Park</i> | |
| Evaluating a How-to Tip Machine Comprehension Model with QA Examples collected from a Community QA Site | 622 |
| <i>Tingxuan Li, Shuting Bai and Fuzhu Zhu, Takehito Utsuro</i> | |
| Syntactic Distribution of the Semantic Classes of Dative Verbs in English and Cantonese: A Crosslinguistic Perspective | 628 |
| <i>Ziying Li and Hanjung Lee</i> | |
| A Corpus-based Analysis of Prosodic Pauses in <i>bă</i> , <i>gěi</i> and <i>ràng</i> Constructions in Taiwan Mandarin | 640 |
| <i>Chihkai Lin</i> | |
| A Corpus-based Lexical Semantic Study of Mandarin Verbs of <i>zhidao</i> and <i>liaojie</i> | 646 |
| <i>Yi Liu</i> | |
| Three Types of Average Dependency Distances of Sentences in a Multilingual Parallel Corpus . . . | 652 |
| <i>Masanori Oya</i> | |
| PS-GAN: Feature augmented text generation in Telugu | 662 |

Pratyusha Musunuru and Varshit Battu

Whose change of state do resultative adverbialised adjectives describe in Japanese locative-alternating constructions? 672

Masataka Ogawa and Takeshi Kishiyama

ViVQA: Vietnamese Visual Question Answering 683

Khanh Quoc Tran, An Trong Nguyen, An Tran-Hoai Le and Kiet Van Nguyen

Monolingual vs multilingual BERTology for Vietnamese extractive multi-document summarization 692

Huy To Quoc, Kiet Nguyen Van, Ngan Nguyen Luu-Thuy and Anh Nguyen Gia-Tuan

Animosity and suffering: Metaphors of BITTERNESS in English and Chinese 700

Gabor Parti, Andreas Liesenfeld and Chu-Ren Huang

Analysis of Text-Semantics via Efficient Word Embedding using Variational Mode Decomposition 711

Rohith Ramakrishnan, Anirudh Vadakedath and U Vamsi Krishna, Premjith B and Kp Soman

Science Mapping of Publications in Natural Language Processing in the Philippines: 2006 to 2020 721

Rachel Edita O. Roxas, Joseph Marvin R. Imperial and Angelica H. De La Cruz

Construction and Evaluation of Japanese Sentence-BERT Models 731

Naoki Shibayama and Hiroyuki Shinnou

Verbal Semantics as Eventive Information: The Case of Verb “制服”and “制約”in Chinese Gigaword Corpus 739

Zhan Wang and Ming Ming Chiu

Study of Similarity Measures as Features in Classification for Answer Sentence Selection Task in Hindi Question Answering: Language-Specific v/s Other Measures 747

Devika Verma, Ramprasad Joshi, Shubhamkar Joshi and Onkar Susladkar

A Hybrid PSO-GA for Extractive Text Summarization 757

Bui Thi-Mai-Anh, Thi Thu Trang Nguyen, Tran Thi Dinh and Nguyen Thi Hoai

From Near-synonyms to Divergent Viewpoint Foci: A Corpus-based MARVS Driven Account of Two Verbs of Attention 767

Ziqian Wang and Chu-Ren Huang

Metaphor Development in Public Discourse Using an ARIMA Time Series Analysis Approach . . 776

Winnie Huiheng Zeng, Dennis Tay and Kathleen Ahrens