

2020 International Conference on Asian Language Processing (IALP 2020)

**Kuala Lumpur, Malaysia
4-6 December 2020**



IEEE Catalog Number: CFP2044I-POD
ISBN: 978-1-7281-7690-1

**Copyright © 2020 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP2044I-POD
ISBN (Print-On-Demand):	978-1-7281-7690-1
ISBN (Online):	978-1-7281-7689-5

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
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

Table of Content

About the Conference	VIII
Message from the Conference Chairs	IX
Message from the Program Chairs	X
Conference Committees	XI
Keynote Speeches	XIV
Author Index	XVII
Papers	
1. Conversation and Others	
P1-1 Detect Turn-takings in Subtitle Streams with Semantic Recall Transformer Encoder <i>Yuhai Liang and Qiang Zhou</i>	1
P1-2 Leader Identification Using Multimodal Information in Multi-party Conversations <i>Tsukasa Shiota, Kouki Honda, Kazutaka Shimada and Takeshi Saitoh.....</i>	7
P1-3 Speaker Identification and Its Application to Social Network Construction for Chinese Novels <i>Yuxiang Jia, Huayi Dou, Shuai Cao and Hongying Zan.....</i>	13
P1-4 Enhancing Attention Models via Multi-head Collaboration <i>Huadong Wang and Mei Tu.....</i>	19
P1-5 Grammatical Error Correction: More Data with More Context <i>Kevin Parnow, Zuchao Li and Hai Zhao.....</i>	24
P1-6 Online Handwritten Mongolian Character Recognition using CMA-MOHR and Coordinate Processing <i>Fan Yang, Feilong Bao and Guanglai Gao.....</i>	30
2. Information Retrieval	
P2-1 Semantic-Based Query Expansion for Academic Expert Finding <i>Theresia V. Rampisela and Evi Yulianti</i>	34
P2-2 Sentences Similarity Based on Deep Structured Semantic Model and Semantic Role Labeling <i>Qiaoli Zhou, zhiwen Jiang and Fengling Yang.....</i>	40
P2-3 Improved Cross-Lingual Document Similarity Measurement <i>Udhan Isuranga, Janaka Sandaruwan, Udesch Athukorala and Gihan Dias.....</i>	45
P2-4 Related Article Extraction Using Learning to Rank <i>Tsukasa Shiota, Kazutaka Shimada, Shinji Nogami and Shuhei Fukuyama.....</i>	50

P2-5 Mongolian Questions Classification in the Law Domain <i>Guangyi Wang, Feilong Bao and Weihua Wang</i>	56
P2-6 Comparative Analysis of Nine Arabic Stemmers on Microblog Information Retrieval <i>Amal Almazrui, Manal Almazrui and Hend Alkhalifa</i>	60
3. Syntax and Semantics Processing	
P3-1 Stylistic Syntactic Structure Extraction and Semantic Clustering for Different Registers <i>Haiyan Wu, Ying Liu and Qingfeng Wu</i>	66
P3-2 Adjusting Indonesian Multiword Expression Annotation to the Penn Treebank Format <i>Jessica Naraiswari Arwidarasti, Ika Alfina and Adila Alfa Krisnadhi</i>	75
P3-3 Feedforward Approach to Sequential Morphological Analysis in the Tagalog Language <i>Arian N. Yambao and Charibeth K. Cheng</i>	81
P3-4 Aksara: An Indonesian Morphological Analyzer that Conforms to the UD v2 Annotation Guidelines <i>Muhammad Yudistira Hanifmuti and Ika Alfina</i>	86
P3-5 A Burmese Dependency Parsing Method Based on Transfer Learning <i>Cunli Mao, Zhibo Man, Zhengtao Yu, Zhenhan Wang, Shengxiang Gao and Yafei Zhang</i>	92
P3-6 Using Annotation Projection for Semantic Role Labelling of Low-Resourced Language: Sinhala <i>Sandun Gunasekara, Dulanjaya Chathura, Chamoda Jeewantha and Gihan Dias</i>	98
P3-7 Selecting the UD v2 Morphological Features for Indonesian Dependency Treebank <i>Ika Alfina, Daniel Zeman, Arawinda Dinakaramani, Indra Budi and Heru Suhartanto</i>	104
4. Linguistics I: Language Learning and Others	
P4-1 Production of L2 Mandarin Mono- and Di-syllabic Tones by Kazakh Learners of Different Levels <i>Zhuanzhuan Ji, Yanlu Xie and Jinsong Zhang</i>	110
P4-2 Effect of Perceptual Training on the Production of English Tense-Lax Vowels by Native Chinese Speakers <i>Jian Gong, Zhenzhen Yang, William Bellamy, Feng Wang and Xiaoli Ji</i>	115
P4-3 The Effect of Vowel Contexts on Voice Onset Time of Mandarin Word-Initial Stops <i>Dan Du and Jinsong Zhang</i>	121
P4-4 Exploring Correlation Between Perception of Postvocalic Nasals and Production of Vowel Nasalization in English by Chinese Learners <i>Ran Meng, Ying Chen, Qimiao Zhou and Hongkun Fu</i>	125
P4-5 Study on Semantic Transparency of Chinese Compounds Based on Word Embedding <i>Xuemei Tang and Shichen Liang</i>	130
P4-6 What Affects the Word Order of Target Language in Simultaneous Interpretation <i>Zhongxi Cai, Koichiro Ryu and Shigeki Matsubara</i>	135

P4-7 When Dialects Disappear: A Report of PPLRC in Henan Province <i>Lei Wang and Huibin Zhuang</i>	141
5. Language Understanding I	
P5-1 HisBERT for Conversational Reading Comprehension <i>Chuang Liu, Deyi Xiong, Yuxiang Jia, Hongying Zan and Changjian Hu</i>	147
P5-2 Effective Strategies for Low-Resource Reading Comprehension <i>Yimin Jing and Deyi Xiong</i>	153
P5-3 Building Chinese Word Knowledge Base for Children's Leveled Reading <i>Zhiying Liu, Lijiao Yang, Jiaomei Zhou and Lu Zhang</i>	158
P5-4 Colloquial Arabic Tweets: Collection, Automatic Annotation, and Classification <i>Yasmin Khalifa and Ashraf Elnagar</i>	163
P5-5 Construction of a Descriptive Framework for Grammatical Attributes of Relational Mark of Korean Compound Sentence Based on Knowledge Representation <i>Yude Bi, Huijie Zhang and Yalin Dong</i>	169
P5-6 Exploring Hybrid Linguistic Feature Sets to Measure Filipino Text Readability <i>Joseph Marvin Imperial and Ethel Ong</i>	175
6. Spoken Language Processing	
P6-1 A Bilingual Speech Synthesis System of Standard Malay and Indonesian Based on HMM- DNN <i>Feng Chen, Jian Yang and Lixuan Zhao</i>	181
P6-2 Detecting Mild Pathological Voice based on Multi-modality Features <i>Quanyi Wang, Feng Zou, Manwa L. Ng, Nan Yan and Lan Wang</i>	187
P6-3 Transformer-based Arabic Dialect Identification <i>Wanqiu Lin, Maulik Madhavi, Rohan Kumar Das and Haizhou Li</i>	192
P6-4 Agent/Client Speech Identification for Mixed-Channel Conversation in Customer Service Call Centers <i>Van Hai Do and Van Tuan Mai</i>	197
P6-5 Prosodic Comparisons of Two Kinds of Contrastive Focus in Mandarin Speech <i>Jieqiong Luo, Yanlu Xie and Jinsong Zhang</i>	201
P6-6 Gated Bilinear Networks for Vowel Formant Estimation <i>Wang Dai, Zheng Hua, Jinsong Zhang, Yanlu Xie and Binghuai Lin</i>	205
P6-7 Self-speech Recognition in High Depression Inclination Individuals <i>Xiaoyong Lu, Daimin Shi, Yang Liu, Jingyi Yuan, Tao Pan and Yanqin Li</i>	210
7. Language Understanding II	
P7-1 An End-to-End Approach for Document-level Event Factuality Identification in Chinese <i>Xiaojia Li, Yun Zhang and Zhong Qian</i>	215

P7-2 Multi-domain Language Understanding of Task-Oriented Dialogue Based on Intent Enhancement <i>Feng Yu, Dequan Zheng and Xiaotian Zhao</i>	221
P7-3 Acquisition of Periodic Events with Person Attributes <i>Kohei Yamamoto and Kazutaka Shimada</i>	229
P7-4 Sarcasm Detection Using Role Pairs: Bootstrapping Role Pairs and Weighting <i>Satoshi Hiai and Kazutaka Shimada</i>	235
P7-5 IndoAbbr: A New Benchmark Dataset for Indonesian Abbreviation Identification <i>Shanshan Li, Nankai Lin, Lixian Xiao and Shengyi Jiang</i>	241
P7-6 Chinese Chemical Named Entity Recognition Based on Morpheme <i>Guirong Wang, Bo Xia, Ye Xiao, Gaoqi Rao and Endong Xun</i>	247
8. Sentiment Classification	
P8-1 Incorporating Multi-Type External Information for Document-Level Sentiment Classification <i>Pengyuan Liu and Chenghao Zhu</i>	253
P8-2 Estimating Evaluation of Cosmetics Reviews with Machine Learning Methods <i>Qing Ma, Miran Tsukagoshi and Masaki Murata</i>	259
P8-3 Joint Embedding of Words and Labels for Sentiment Classification <i>Yingwei Sheng and Inui Takashi</i>	264
P8-4 Sentiment classification with syntactic relationship and attention for teaching evaluation texts <i>Lingling Mu, Yadi Li and Hongying Zan</i>	270
P8-5 Exploiting Vietnamese Social Media Characteristics for Textual Emotion Recognition in Vietnamese <i>Khang Phuoc-Quy Nguyen and Kiet Van Nguyen</i>	276
P8-6 Structurally Enhanced Interactive Attention Network for Aspect-Level Sentiment Classification <i>Chunfeng Liang, Yumeng Fu and Chengguo Lv</i>	282
9. Machine Translation	
P9-1 Statistical Machine Translation Approach for Lexical Normalization on Indonesian Text <i>Ajmal Kurnia and Evi Yulianti</i>	288
P9-2 Korean Neural Machine Translation Using Hierarchical Word Structure <i>Jeonghyeok Park and Hai Zhao</i>	294
P9-3 Transfer Building of Multiword Expression Resource from Indonesian to Malay <i>Wuying Liu and Lin Wang</i>	299
P9-4 English to Sinhala Neural Machine Translation	

<i>Thilakshi Fonseka, Rashmini Naranpanawa, Ravinga Perera and Uthayasanker Thayasivam</i>	305
P9-5 Semi-Supervised Low-Resource Style Transfer of Indonesian Informal to Formal Language with Iterative Forward-Translation <i>Haryo Akbarianto Wibowo, Tatag Aziz Prawiro, Muhammad Ihsan, Alham Fikri Aji, Radityo Eko Prasajo, Rahmad Mahendra and Suci Fitriany</i>	310
P9-6 A Neural Machine Translation Approach for Translating Malay Parliament Hansard to English Text <i>Yu-Zane Low, Lay-Ki Soon and Shageenderan Sapai</i>	316
10. Linguistics II: Syntax	
P10-1 The Intersubjectivity of the Sentence-Final Ba in Discourse <i>Minfeng Wang</i>	321
P10-2 The Survey on the Distribution of MC Fei and Xiao Initial Groups in Chinese Dialects <i>Yan Li and Xiaochuan Song</i>	326
P10-3 A Study on the Synergistic Relationship between Lexical Structure and Word Frequency, Polysemy and Synonymy in Modern Chinese <i>Xinpei Hong and Yuelong Wang</i>	336
P10-4 Study for the Processing of the Chinese Verb-Noun Anomalous Collocations with Metaphors <i>Mengxiang Wang, Hongcheng Ma and Junjie Hong</i>	342
P10-5 The Pragmatic Function and Syntactic Mechanism of the Co-occurrence of Two Negative Components <i>Hongcheng Ma</i>	346
P10-6 The Syntactic Analysis of Unaccusative Verbs in Archaic Chinese <i>Mengbin Liu</i>	351