

# **2019 International Conference on Asian Language Processing (IALP 2019)**

**Shanghai, Singapore  
15 – 17 November 2019**



**IEEE Catalog Number: CFP1944I-POD  
ISBN: 978-1-7281-5015-4**

**Copyright © 2019 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:	CFP1944I-POD
ISBN (Print-On-Demand):	978-1-7281-5015-4
ISBN (Online):	978-1-7281-5014-7

**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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

## Contents

Welcome Message -----	I
Conference Committees -----	III
Keynote Speeches-----	VII
Conference Program -----	IX
Author Index -----	XVII

### Papers

<b>Quantifying the Use of English Words in Urdu News-Stories</b> <i>Mehtab Alam Syed, Arif Ur Rahman, Muzammil Khan</i> -----	1
<b>A New Method of Tonal Determination for Chinese Dialects</b> <i>Yan Li, Zhiyi Wu</i> -----	7
<b>Duplicate Question Detection based on Neural Networks and Multi-head Attention</b> <i>Heng Zhang, Liangyu Chen</i> -----	13
<b>Acoustic Scene Classification Using Deep Convolutional Neural Network via Transfer Learning</b> <i>Min Ye, Hong Zhong, Xiao Song, Shilei Huang, Gang Cheng</i> -----	19
<b>Japanese Particle Error Correction employing Classification Model</b> <i>Youichiro Ogawa, Kazuhide Yamamoto</i> -----	23
<b>Effect of Preprocessing for Distributed Representations: Case Study of Japanese Radiology Reports</b> <i>Taro Tada, Kazuhide Yamamoto</i> -----	29
<b>Character Decomposition for Japanese-Chinese Character-Level Neural Machine Translation</b> <i>Jinyi Zhang, Tadahiro Matsumoto</i> -----	35
<b>Japanese grammatical simplification with simplified Corpus</b> <i>Yumeto Inaoka and Kazuhide Yamamoto</i> -----	41
<b>Effective Data Augmentation Approaches to End-to-End Task-Oriented Dialogue</b> <i>Jun Quan, Deyi Xiong</i> -----	47
<b>Extremely Low Resource Text simplification with Pre-trained Transformer Language Model</b> <i>Takumi Maruyama and Kazuhide Yamamoto</i> -----	53
<b>Aspect-based Opinion Mining for Code-Mixed Restaurant Reviews in Indonesia</b>	

<i>Andi Suciati, Indra Budi</i> -----	59
<b>Using Convolutional Neural Network with BERT for Intent Determination</b>	
<i>Changai He, Sibao Chen</i> -----	65
<b>What affects the difficulty of Chinese syntax?</b>	
<i>Yueming Du, Lijiao Yang</i> -----	71
<b>Classified Description and Application of Chinese Constitutive Role</b>	
<i>Mengxiang Wang, Cuiyan Ma</i> -----	75
<b>Design and Implementation of Myanmar Speech Synthesis System Based on HMM-DNN</b>	
<i>Mengyuan Liu and Jian Yang</i> -----	79
<b>Exploring Letter's Differences between Partial Indonesian Branch Language and English</b>	
<i>Nankai Lin, Sihui Fu, Jiawen Huang, Shengyi Jiang</i> -----	84
<b>CIEA: A Corpus for Chinese Implicit Emotion Analysis</b>	
<i>Dawei Li, Jin Wang, Xuejie Zhang</i> -----	90
<b>Learning Deep Matching-Aware Network for Text Recommendation using Clickthrough Data</b>	
<i>Haonan Liu, Nankai Lin, Shengyi Jiang, Zitao Chen and Ke Li</i> -----	96
<b>Comprehension correlates of the occurrence and deletion of “de” in Mandarin “N1 (de) N2” structures</b>	
<i>Junyuan Zhao, Junru Wu</i> -----	102
<b>Effects of English Capitals On Reading Performance of Chinese Learners: Evidence from Eye Tracking</b>	
<i>Yang Wei, Fu Xinyu</i> -----	108
<b>A Machine Learning Model for the Dating of Ancient Chinese Texts</b>	
<i>Xuejin Yu, Wei Huangfu</i> -----	115
<b>Correlational Neural Network Based Feature Adaptation in L2 Mispronunciation Detection</b>	
<i>Wenwei Dong and Yanlu Xie</i> -----	121
<b>Acoustic Cues for Sensory Modality and Referential Vagueness of Personal Pronouns: Me or Not? You or Not?</b>	
<i>Aibao Zhou, Yanbing Hu, Xiaoyong Lu, Tao Pan</i> -----	126
<b>Automated Prediction of Item Difficulty in Reading Comprehension Using Long Short-Term Memory</b>	
<i>Li-Huai Lin, Tao-Hsing Chang, Fu-Yuan Hsu</i> -----	132
<b>Automatic Meta-evaluation of Low-Resource Machine Translation Evaluation Metrics</b>	
<i>Junting Yu, Wuying Liu, Hongye He, Lin Wang</i> -----	136
<b>Combination of Semantic Relatedness with Supervised Method for Word Sense Disambiguation</b>	
<i>Qiaoli Zhou, Yuguang Meng</i> -----	142
<b>Acquisition of Knowledge with Time Information from Twitter</b>	

<i>Kohei Yamamoto, Kazutaka Shimada</i> -----	148
<b>Research on Chinese Text Error Correction Based on Sequence Model</b>	
<i>Jianyong Duan Yang Yuan, Hao Wang, Xiaopeng Wei, Zheng Tan</i> -----	154
<b>Integrating Knowledge in End-to-End Automatic Speech Recognition for Mandarin-English Code-Switching</b>	
<i>Chia-Yu Li, Ngoc Thang Vu</i> -----	160
<b>Improving Question Classification with Hybrid Networks</b>	
<i>Yichao Cao, Miao Li, Tao Feng, Ruijing Wang, Yue Wu</i> -----	166
<b>A Study on the Robustness of Pitch Range Estimation from Brief Speech Segments</b>	
<i>Wenjie Peng, Kaiqi Fu, Wei Zhang, Yanlu Xie, Jinsong Zhang</i> -----	172
<b>Are Scoring Feedback of CAPT Systems Helpful for Pronunciation Correction? -- An Exception of Mandarin Nasal Finals</b>	
<i>Rui Cai, Wei Wei, Jinsong Zhang</i> -----	177
<b>Examination-Style Reading Comprehension with Neural augmented Retrieval</b>	
<i>Yiqing Zhang, Hai Zhao and Zhuosheng Zhang</i> -----	182
<b>Automatic Extraction and Quantitative Evaluation of the Character Relationship Networks from Children's Literature works</b>	
<i>Kun Ma, Lijiao Yang</i> -----	188
<b>A Chinese word segment model for energy literature based on Neural Networks with Electricity User Dictionary</b>	
<i>Bochuan Song, Bo Chai, Qiang Zhang, Quanye Jia</i> -----	194
<b>Fusion of Image-text attention for Transformer-based Multimodal Machine Translation</b>	
<i>Junteng Ma, Shihao Qin, Lan Su, Xia Li, Lixian Xiao</i> -----	199
<b>Speech Command Classification System for Sinhala Language based on Automatic Speech Recognition</b>	
<i>Thilini Ranagalage, Lakshika Athapaththu, Pamoda Abeyawardana, Uthayasanker Thayasivam and Sanath Jayasena</i> -----	205
<b>An Enhancement of Malay Social Media Text Normalization for Lexicon-Based Sentiment Analysis</b>	
<i>Muhammad Fakhur Razi Abu Bakar, Norisma Idris, Liyana Shuib</i> -----	211
<b>Improving text simplification by corpus expansion with unsupervised Learning</b>	
<i>Akihiro Katsuta and Kazuhide Yamamoto</i> -----	216
<b>Cross Language Information Retrieval Using Parallel Corpus with Bilingual Mapping Method</b>	
<i>Rinaldi Andrian Rahmanda, Mirna Adriani and Dipta Tanaya</i> -----	222
<b>Language Detection in Sinhala-English Code-mixed Data</b>	
<i>Ian Smith, Uthayasanker Thayasivam</i> -----	228
<b>Sinhala and Tamil Speech Intent Identification From English Phoneme Based ASR</b>	
<i>Yohan Karunanayake, Uthayasanker Thayasivam, Surangika Ranathunga</i> -----	234
<b>An experimental Tagalog Finite State Automata spellchecker with Levenshtein</b>	

<b>edit-distance feature</b>	
<i>Joseph Marvin R. Imperial Czerittonnie Gail V. Ya-On, Jennifer C. Ureta</i> -----	240
<b>Statistical Machine Learning for Transliteration: Transliterating names between Sinhala, Tamil and English</b>	
<i>H.S. Priyadarshani, M.D.W. Rajapaksha, M.M.S.P. Ranasinghe, K. Sarveswaran, G.V. Dias</i> -----	244
<b>Automatic Recognition of Allusions in Tang Poetry Based on BERT</b>	
<i>Xuemei Tang, Shichen Liang, Jianyu Zheng, Renfen Hu, Zhiying Liu</i> -----	255
<b>Exploring Characteristics of Word Co-occurrence Network in Translated Chinese</b>	
<i>Jianyu Zheng, Kun Ma, Xuemei Tang, Shichen Liang</i> -----	261
<b>Diachronic Synonymy and Polysemy: Exploring Dynamic Relation Between Forms and Meanings of Words Based on Word Embeddings</b>	
<i>Shichen Liang, Jianyu Zheng, Xuemei Tang, Renfen Hu, Zhiying Liu</i> -----	267
<b>Automatic answer ranking based on sememe vector in KBQA</b>	
<i>Yadi Li, Lingling Mu, Hao Li, Hongying Zan</i> -----	273
<b>Improving Japanese-English Bilingual Mapping of Word Embeddings based on Language Specificity</b>	
<i>Yuting Song, Biligsaikhan Batjargal, Akira Maeda</i> -----	279
<b>The Initial Research of Mongolian Literary Corpus-Take the Text of Da.Nachugdorji's Work for Instance</b>	
<i>YinhuaHai</i> -----	284
<b>Towards Robust Neural Machine Reading Comprehension via Question Paraphrases</b>	
<i>Ying Li, Hongyu Li and Jing Liu</i> -----	290
<b>Employing Gated Attention and Multi-similarities to Resolve Document-level Chinese Event Coreference</b>	
<i>Haoyi Cheng, Peifeng Li, Qiaoming Zhu</i> -----	296
<b>A Multi-stage Strategy for Chinese Discourse Tree Construction</b>	
<i>Tishuang Wang, Peifeng Li, Qiaoming Zhu</i> -----	302
<b>Articulatory Features Based TDNN Model for Spoken Language Recognition</b>	
<i>Jiawei Yu, Minghao Guo, Yanlu Xie and Jinsong Zhang</i> -----	308
<b>A Study on Syntactic Complexity and Text Readability of ASEAN English News</b>	
<i>Yusha Zhang, Nankai Lin, Shengyi Jiang</i> -----	313
<b>Tibetan word segmentation method based on CNN-BiLSTM-CRF model</b>	
<i>Lili Wang, Hongwu Yang, Xiaotian Xing, Yajing Yan</i> -----	319
<b>Separate Answer Decoding for Multi-class Question Generation</b>	
<i>Kaili Wu, Yu Hong, Mengmeng Zhu, Hongxuan Tang, Min Zhang</i> -----	325
<b>Converting an Indonesian Constituency Treebank to the Penn Treebank Format</b>	
<i>Jessica Naraiswari Arwidarasti, Ika Alfina, Adila Alfa Krisnadhi</i> -----	331
<b>How to Answer Comparison Questions</b>	
<i>Hongxuan Tang, Yu Hong, Xin Chen, Kaili Wu, Min Zhang</i> -----	337

<b>Improved DNN-HMM English Acoustic Model Specially For Phonotactic Language Recognition</b>	
<i>Wei-wei Liu, Guo-Chun Li Li, Cun-Xue Zhang, Hai-Feng Yan, Jing He, Ying-Xin Gan, Yan-Miao Song, Jian-Hua Zhou, Jian-Zhong Liu, Ying Yin, Ya-nan Li, Yu-Bin Huang, Ting Ruan, Wei Liu, Rui-Li Du, Hua-ying Bai, Wei Li and Sheng-Ge Zhang</i>	343
<b>Confidence Modeling for Neural Machine Translation</b>	
<i>Taichi Aida, Kazuhide Yamamoto</i>	349
<b>Using WHY-type Question-Answer Pairs to Improve Implicit Causal Relation Recognition</b>	
<i>Huibin Ruan, Yu Hong, Yu Sun, Yang Xu and Min Zhang</i>	355
<b>On the Etymology of the 'river' in Chinese</b>	
<i>Huibin Zhuang, Zhanting Bu</i>	361
<b>Using Mention Segmentation to Improve Event Detection with Multi-head Attention</b>	
<i>Jiali Chen, Yu Hong, Jingli Zhang, Jianmin Yao</i>	367
<b>Chinese Spelling Check based on Sequence Labeling</b>	
<i>Zijia Han, Zonglin Liu, Qiansheng Wang, Guohong Fu</i>	373
<b>Research on Tibetan Text Classification Method Based on Neural Network</b>	
<i>Zhensong Li, Jie Zhu, Zhixiang Luo, Saihu Liu</i>	379
<b>BERT with Enhanced Layer for Assistant Diagnosis Based on Chinese Obstetric EMRs</b>	
<i>Kunli Zhang, Chuang Liu, Xuemin Duan, Lijuan Zhou, Yueshu Zhao, Hongying Zan</i>	384
<b>A Systematic Investigation of Neural Models for Chinese Implicit Discourse Relationship Recognition</b>	
<i>Dejian Li, Man Lan, Yuanbin Wu</i>	390
<b>Syntax-aware Transformer Encoder for Neural Machine Translation</b>	
<i>Sufeng Duan, Hai Zhao, Junru Zhou, Rui Wang</i>	396
<b>An End-to-End Model Based on TDNN-BiGRU for Keyword Spotting</b>	
<i>Shuzhou Chai, Zhenye Yang, Changsheng Lv and Wei-Qiang Zhang</i>	402
<b>Coarse-to-Fine Document Ranking for Multi-Document Reading Comprehension with Answer-Completion</b>	
<i>Hongyu Liu, Shumin Shi, Heyan Huang</i>	407
<b>Developing a machine learning-based grade level classifier for Filipino children's literature</b>	
<i>Joseph Marvin Imperial, Rachel Edita Roxas, Erica Mae Campos, Ani Rosa Almario</i>	413
<b>Exploring Context's Diversity to Improve Neural Language Model</b>	
<i>Yanchun Zhang, Xingyuan Chen, Peng Jin, Yajun Du</i>	419
<b>Phrase-Based Tibetan-Chinese Statistical Machine Translation</b>	
<i>YONG Cuo, Xiaodong SHI, NYIMA Tashi, Yidong CHEN</i>	424
<b>A General Procedure for Improving Language Models in Low-Resource Speech Recognition</b>	
<i>Qian Liu, Wei-Qiang Zhang, Jia Liu, Yao Liu</i>	428

<b>Effect of Music Training on the Production of English Lexical Stress by Chinese English Learners</b>	
<i>Jie Lian, hui feng and Yingjie Zhao</i> -----	434
<b>Carrier Sentence Selection with Word and Context Embeddings</b>	
<i>Chak Yan Yeung, John Lee, Benjamin Tsou</i> -----	439
<b>Multiple-source Entity Linking with Incomplete Sources</b>	
<i>Qinguo Liu, Shui Liu, Lemao Liu, Bo Xiao</i> -----	445
<b>Neural Machine Translation Strategies for Generating Honorific-style Korean</b>	
<i>Lijie Wang, Mei Tu, Mengxia Zhai, Huadong Wang, Song Liu, Sang Ha Kim</i> -----	450
<b>Ranking Like Human: Global-View Matching via Reinforcement Learning for Answer Selection</b>	
<i>Yingxue Zhang, Ping Jian, Ruiying Geng and Fandong Meng</i> -----	456
<b>Development of a Filipino Speaker Diarization in Meeting Room Conversations</b>	
<i>Angelica H. De La Cruz, Rodolfo C. Raga Jr</i> -----	462
<b>Statistical Analysis of Syllable Duration of Uyghur Language</b>	
<i>Askar Hamdulla, Guzalnur Dilmurat, Gulnur Arkin, Mijit Ablimit</i> -----	468
<b>A Comparative Analysis of Acoustic Characteristics between Kazak &amp; Uyghur Mandarin Learners and Standard Mandarin Speakers</b>	
<i>Gulnur Arkin, Gvljan Alijan, Askar Hamdulla, Mijit ablimit</i> -----	474
<b>Construction of Quantitative Index System of Vocabulary Difficulty in Chinese Grade Reading</b>	
<i>Huiping Wang, Lijiao Yang, Huimin Xiao</i> -----	480
<b>An Measurement Method of Ancient Poetry Difficulty for Adaptive Testing</b>	
<i>Huiping Wang, Bihua Wang</i> -----	487
<b>Prosodic Realization of Focus in Changchun Mandarin and Nanjing Mandarin</b>	
<i>Ying Chen, Jiajing Zhang, Bingying Ye, Chenfang Zhou</i> -----	492
<b>Research on New Event Detection Methods for Mongolian News</b>	
<i>Shijie Wang, Feilong Bao, Guanglai Gao</i> -----	498
<b>Improving Mandarin Prosody Boundary Detection by Using Phonetic Information and Deep LSTM Model</b>	
<i>Ju Lin, Zhuanzhuan Ji, Wenwei Dong, Yanlu Xie, Jinsong Zhang</i> -----	504