

35th Conference on Computational Linguistics and Speech Processing (ROCLING 2023)

Taipei City, Taiwan
20-21 October 2023

ISBN: 978-1-7138-9035-5

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© (2023) by the Association for Computational Linguistics and Chinese Language Processing (ACLCLP)
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

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

XFEVER: Exploring Fact Verification across Languages	1
<i>Yi-Chen Chang, Canasai Kruengkrai and Junichi Yamagishi</i>	
Story Co-telling Dialogue Generation via Reinforcement Learning and Knowledge Graph	12
<i>Yu-Kai Lee, Chia-Hui Chang</i>	
Improving End-to-end Taiwanese-Speech-to-Chinese-Text Translation by Semi-supervised Learning	21
<i>Yu-Chun Lin, Chung-Che Wang and Jyh-Shing Jang</i>	
Construction of Message Deliver Service Dialog Systems	29
<i>Cheng-Hung Yeh, Chia-Hui Chang</i>	
Auxiliary loss to attention head for end to end speaker diarization	38
<i>Yi-Ting Yang, Jiun-Ting Li and Berlin Chen</i>	
The Pilot Study and Model Construction for Word Segmentation in Taiwan Hakka	44
<i>Chiou-Shing Yeh, Huei-Ling Lai and Jyi-Shane Liu</i>	
Leveraging Dialogue Discourse Parsing in a Two-Stage Framework for Meeting Summarization	54
<i>Yi-Ping Huang, Tien-Hong Lo and Berlin Chen</i>	
Improving Low-Resource Speech Recognition through Multilingual Fine-Tuning with Language Identifiers and Self-Training	63
<i>Karol Nowakowski, Michal Ptaszynski</i>	
AaWLoss: An Artifact-aware Weighted Loss Function for Speech Enhancement	71
<i>En-Lun Yu, Kuan-Hsun Ho and Berlin Chen</i>	
WordRank: A Word Ranking based Training Strategy for Abstractive Document Summarization	79
<i>Hsiao-Wei Chou, Ping-Yen Wu, Jia-Jang Tu and Kuanyu Chen</i>	
Investigating Cross-Institutional Recognition of Cancer Registration Items: A Case Study on Catastrophic Forgetting	89
<i>You Chen Zhang, Chen-Kai Wang, Ming-Ju Tsai and Hong-Jie Dai</i>	
Enhancing Automated English Speaking Assessment for L2 Speakers with BERT and Wav2vec2.0 Fusion	98
<i>Wen-Hsuan Peng, Hsin-Wei Wang, Sally Chen and Berlin Chen</i>	
Impact of Feature Selection Algorithms on Readability Model	106
<i>Tsai-Ning Tai, Hou-Chiang Tseng and Yao-Ting Sung</i>	
Multimodal Speech Training for the Hard of Hearing in Mandarin	116
<i>Ching-Hsiang Ho, Kun-Chuan Tseng and Men-Ching Lei</i>	
Is GPT-4 a Good Islamic Expert for Answering Quran Questions?	124
<i>Sarah Alnefaie, Eric Atwell and Mohammad Ammar Alsalka</i>	
Addressing the issue of Data Imbalance in Multi-granularity Pronunciation Assessment	134
<i>Meng-Shin Lin, Hsin-Wei Wang, Tien-Hong Lo, Berlin Chen and Wei-Cheng Chao</i>	
Category Mapping for Zero-shot Text Classification	141
<i>Qiu-Xia Zhang, Te-Yu Chi, Te-Lun Yang, Yu-Meng Tang, Ta-Lin Chen and Jyh-Shing Jang</i>	
ESC MA-SD Net: Effective Speaker Separation through Convolutional Multi-View Attention and SudoNet	157
<i>Che-Wei Liao, Aye Nyein Aung and Jieih-Weih Hung</i>	

A Comparative Study of Generative Pre-trained Transformer-based Models for Chinese Slogan Generation of Crowdfunding	162
<i>Yu-Cheng Liang, Meng-Heng Zheng and Jheng-Long Wu</i>	
Fine-Tuning and Evaluation of Question Generation for Slovak Language	171
<i>Ondrej Megela, Daniel Hladek, Matus Pleva, Ján Staš, Ming-Hsiang Su and Yuan-Fu Liao</i>	
Phonotactic Constraints on Zhangzhou Onsets	179
<i>Yishan Huang</i>	
Analyzing ChatGPT's Mathematical Deficiencies: Insights and Contributions	188
<i>Vincent Cheng and Yu Zhang</i>	
An Analysis of shi Y''Metaphors in Mandarin Corpora and Learning Materials	194
<i>Yu-Hsiang Shen and Siaw-Fong Chung</i>	
Sentence-level Revision with Neural Reinforcement Learning	202
<i>Zhendong Du, Kenji Hashimoto</i>	
Taiwanese/Mandarin Speech Recognition using OpenAI's Whisper Multilingual Speech Recognition Engine Based on Generative Pretrained Transformer Architecture	210
<i>Yueh-Che Hsieh, Renyuan Lyu and Keming Lyu</i>	
KNOT-MCTS: An Effective Approach to Addressing Hallucinations in Generative Language Modeling for Question Answering	215
<i>Chung-Wen Wu, Guan-Tang Huang and Yue-Yang He, Berlin Chen</i>	
Compact CNNs for End-to-End Keyword Spotting on Resource-Constrained Edge AI Devices	222
<i>Joseph Lin and Renyuan Lyu</i>	
Sound Processing for Cochlear Implants: The Journey of Innovation Toward Artificial Intelligence	227
<i>Enoch Hsin-Ho Huang, Chao-Min Wu and Yu Tsao</i>	
Analyzing Bid-Rigging Related Judicial Cases of Government Procurement Law Using Text Mining Techniques	233
<i>Pei-Zhen Chen, Hsin-Yun Hsu and Jheng-Long Wu</i>	
Fine-Grained Argument Understanding with BERT Ensemble Techniques: A Deep Dive into Financial Sentiment Analysis	242
<i>Eugene Sy, Tzu-Cheng Peng, Shih-Hsuan Huang and Heng-Yu Lin, Yung-Chun Chang</i>	
Application of Deep Learning Technology to Predict Changes in Sea Level	250
<i>Yi-Lin Hsieh and Ming-Hsiang Su</i>	
The Relevance Identification Between Housing Rental Texts And Legal Provisions	255
<i>Min-Chin Ho, Ya-Mien Cheng and Jheng-Long Wu</i>	
Solving Linguistic Olympiad Problems with Tree-of-Thought Prompting	262
<i>Zheng-Lin Lin, Chiao-Han Yen, Jia-Cheng Xu, Deborah Watty and Shu-Kai Hsieh</i>	
Can generative models be used to detect hate speech related to body shaming?	270
<i>Yuan-Shiang Tsai and Yu-Yun Chang</i>	
Lexical Complexity Prediction using Word Embeddings	279
<i>Cheng-Zen Yang, Jin-Jian Li and Shu-Chang Lin</i>	
Analysis of Chinese Irony on PTT Corpus-Using "Tested Positive" and "Hope" as the Key Words	288
<i>Pin-Wen Wang and Siaw-Fong Chung</i>	
Evaluating Interfaced LLM Bias	292

Kai-Ching Yeh, Jou-An Chi, Da-Chen Lian and Shu-Kai Hsieh

A Novel Named Entity Recognition Model Applied to Specialized Sequence Labeling	300
<i>Ruei-Cyuan Su, Tzu-En Su, Ming-Hsiang Su, Matus Pleva and Daniel Hladek</i>	
SCU-MESCLab at ROCLING-2023 Shared Task; GNamed Entity Recognition Using Multiple Classifier Model	311
<i>Tzu-En Su, Ruei-Cyuan Su, Tsung-Hsien Yang and Ming-Hsiang Su</i>	
YNU-HPCC at ROCLING 2023 MultiNER-Health Task: A transformer-based approach for Chinese health-care NER	316
<i>Chonglin Pang, You Zhang and Xiaobing Zhou</i>	
YNU-ISE-ZXW at ROCLING 2023 MultiNER-Health Task: A Transformer-based Model with LoRA for Chinese Healthcare Named Entity Recognition	324
<i>Xingwei Zhang, Jin Wang and Xuejie Zhang</i>	
Overview of the ROCLING 2023 Shared Task for Chinese Multi-genre Named Entity Recognition in the Healthcare Domain	332
<i>Lung-Hao Lee, Tzu-Mi Lin and Chao-Yi Chen</i>	
CrowNER at ROCLING 2023 MultiNER-Health Task: Enhancing NER Task with GPT Paraphrase Augmentation on Sparsely Labeled Data	338
<i>Yin-Chieh Wang, Wen-Hong Wu, Feng-Yu Kuo, Han-Chun Wu, Te-Yu Chi, Te-Lun Yang, Sheh Chen and Jyh-Shing Roger Jang</i>	
LingX at ROCLING 2023 MultiNER-Health Task: Intelligent Capture of Chinese Medical Named Entities by LLMs	349
<i>Xuelin Wang and Qihao Yang</i>	
ISLab at ROCLING 2023 MultiNER-Health Task: A Three-Stage NER Model Combining Textual Content and Tagged Semantics	358
<i>Jun-Jie Wu, Tao-Hsing Chang and Fu-Yuan Hsu</i>	
Accelerating Hakka Speech Recognition Research and Development Using the Whisper Model	366
<i>Ching-Yuan Chen, Yun-Hsiang Hsu and Chenchi Chang</i>	
Enhancing Automatic Speech Recognition Performance Through Multi-Speaker Text-to-Speech	370
<i>Po-Kai Chen, Bing-Jhih Huang, Chi-Tao Chen, Hsin-Min Wang and Jia-Ching Wang</i>	
The DMS-ASR System for the Formosa Speech Recognition Challenge 2023	376
<i>Hsiu Jui Chang and Wei Yuan Chen</i>	
NSYSU-MITLab Speech Recognition System for Formosa Speech Recognition Challenge 2023	379
<i>Hong-Jie Hu and Chia-Ping Chen</i>	
The North System for Formosa Speech Recognition Challenge 2023	385
<i>Li-Wei Chen, Hung-Shin Lee and Kai-Chen Cheng</i>	
WhisperHakka: A Hybrid Architecture Speech Recognition System for Low-Resource Taiwanese Hakka . . .	389
<i>Ming-Hsiu Chiang, Chien-Hung Lai and Hsuan-Sheng Chiu</i>	
The NTNU ASR System for Formosa Speech Recognition Challenge 2023	396
<i>Hao-Chien Lu, Chung-Chun Wang, Jhen-Ke Lin and Tien-Hong Lo</i>	
The Taiwan AI Labs Hakka ASR System for Formosa Speech Recognition Challenge 2023	402
<i>Yuan-Hsiang Lu, Chung-Yi Li and Zih-Wei Lin</i>	
A preliminary study on Hakka speech recognition by using the Branchformer	408
<i>Jia-Jyu Su, Dong-Min Li and Chen-Yu Chiang</i>	

The NTNU Super Monster Team (SPMT) system for the Formosa Speech Recognition Challenge 2023 -
Hakka ASR 413
Tzu-Ting Yang, Hsin Wei Wang and Meng-Ting Tsai, Berlin Chen

Whisper Model Adaptation for FSR-2023 Hakka Speech Recognition Challenge 422
Yi-Chin Huang and Ji-Qian Tsai