ROCLING 2025

The 37th Conference on Computational Linguistics and Speech Processing (ROCLING 2025)

Proceedings of the Conference

November 20-22, 2025

The ROCLING organizers gratefully acknowledge the support from the following sponsors.

Sponsors

©2025 Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-379-1

Message from the Conference Chairs

It is our great pleasure to welcome you to the 37th Conference on Computational Linguistics and Speech Processing (ROCLING 2025), held at National Taiwan University in Taipei, Taiwan, November 20 to 22, 2025.

This year's conference continues the tradition of serving as a premier forum for presenting cutting-edge research and showcasing innovative systems and techniques across the broad fields of natural language processing and speech processing.

This year marks a major milestone for ROCLING with the introduction of a dual-track submission system featuring both archival and non-archival papers. This new format provides authors with greater flexibility to present their research at different stages, combining rigorous review with opportunities for early idea exchange and community feedback.

ROCLING 2025 also launches its first Round Table Forum, a special event that brings together experienced researchers, industry experts, early-career scholars, and students for in-depth small-group discussions. The forum offers a valuable opportunity to exchange ideas across disciplines, receive constructive feedback, and build lasting professional connections.

We are also proud to note that this year's Program Committee is chaired by a team of outstanding young scholars from Taiwan. Their vision and dedication have shaped a high-quality, forward-looking program that reflects the energy, creativity, and diversity of our growing research community.

The conference program features two keynote speeches delivered by world-renowned scholars, who will share their insights into the future of language understanding and speech generation, two tutorials addressing AI-driven hearing assistive technologies and audio intelligence, and three special sessions exploring key directions in persuasive language in the age of AI, sentiment and medical text analysis, and speech recognition for Taiwanese Hakka languages.

We thank all authors, reviewers, organizers, and volunteers, as well as our sponsors, for their dedicated contributions. We hope that ROCLING 2025 will inspire new ideas, foster lasting collaborations, and strengthen our shared mission to advance human language and speech technologies.

Warm regards,

Prof. Yun-Nung Chen, National Taiwan University
Prof. Hung-Yi Lee, National Taiwan University
Prof. Pu-Jen Cheng, National Taiwan University

Conference Chairs, ROCLING 2025

Organizing Committee

Conference Chairs

Yun-Nung Chen, National Taiwan University
Hung-Yi Lee, National Taiwan University
Pu-Jen Cheng, National Taiwan University

Program Chairs

Kai-Wei Chang, Massachusetts Institute of Technology
Ke-Han Lu, National Taiwan University
Chih-Kai Yang, National Taiwan University
Zhi-Rui Tam, National Taiwan University
Wen-Yu Chang, National Taiwan University
Chung-Che Wang, National Taiwan University

Publication Chair

Kai-Wei Chang, Massachusetts Institute of Technology

Tutorial Chair

Lung-Hao Lee, National Yang Ming Chiao Tung University

Special Session Chairs

Tzu-Mi Lin, National Yang Ming Chiao Tung University
Siaw-Fong Chung, National Chengchi University
Yuan-Fu Liao, National Yang Ming Chiao Tung University

Registration Chair

Yu Tsao, Academia Sinica

Web Chairs

Yik Ming Chin, National Taiwan University
Jui-Wei Fu, National Taiwan University
Wei-Tang Hsu, National Taiwan University

Program Committee

Program Committee

Jason S. Chang, National Tsing Hua University
Kai-Wei Chang, Massachusetts Institute of Technology
Wen Yu Chang, National Taiwan University
Yung-Chun Chang, Taipei Medical University
Chien-Cheng Chen, National Taiwan University
Chien-Hung Chen, National Taiwan University
Chung-Chi Chen, National Institute of Advanced Industrial Science and Technology
Kuan-Yu Chen, National Taiwan University of Science and Technology
Pin-Er Chen, National Taiwan University
Po-Chun Chen, National Taiwan University
Wei-Chih Chen, National Taiwan University
Jou-An Chi, National Taiwan University
Huang-Cheng Chou, National Tsing Hua University
Siaw-Fong Chung, National Chengchi University
Min-Yuh Day, National Taipei University
Yao-Chung Fan, National Chung Hsing University
Bo-Han Feng, National Taiwan University
Chi-Yuan Hsiao, National Taiwan University
Chan-Jan Hsu, MediaTek Research
En-Pei Hu, National Taiwan University
Chao-Shih Huang, National Yang Ming Chiao Tung University
Chien-Yu Huang, Carnegie Mellon University
Hen-Hsen Huang, Academia Sinica
Hsiao-Ying Huang, National Taiwan University
Sung-Feng Huang, NVIDIA
Wei Ping Huang, National Taiwan University
Chih-Chung Kuo, National Yang Ming Chiao Tung University
Lung-Hao Lee, National Yang Ming Chiao Tung University
Yu-Ting Lee, National Chengchi University
Yuan-Fu Liao, National Yang Ming Chiao Tung University
Claire Lin, National Taiwan University
Tzu-Quan Lin, National Taiwan University
Yi-Cheng Lin, National Taiwan University
Yu-Xiang Lin, National Taiwan University
Chi-Liang Liu, Kotoba Technologies
Ke-Han Lu, National Taiwan University
Wenze Ren, National Taiwan University
Ming-Hsiang Su, Soochow University
Zhi Rui Tam, National Taiwan University
Yu-Chien Tang, National Yang Ming Chiao Tung University
Ming-Feng Tsai, National Chengchi University
Hou-Chiang Tseng, National Taiwan University of Science and Technology
Wei-Cheng Tseng, University of Texas at Austin
Yuan Tseng, The University of Texas at Austin
Chung-Che Wang, National Tsing Hua University
Hsin-Min Wang, Academia Sinica

Jenq-Haur Wang, National Taipei University of Technology
Jun-You Wang, National Taiwan Normal University
Jheng-Long Wu, Soochow University
Chih-Kai Yang, National Taiwan University
Pei-Ying Yang, National Taiwan University
Wen-Chao Yeh, National Tsing Hua University
An-Zi Yen, National Yang Ming Chiao Tung University
Liang-Chih Yu, Yuan Ze University

Keynote Talk

Towards Social Agents

Asli Celikyilmaz
Meta FAIR

November 21st, 2025 – Time: **09:10 - 10:10** – Room: **2F, Space M Session Room / 1F, R117**

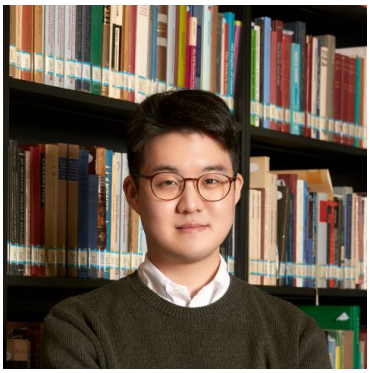
Abstract: As language models evolve into social agents, new challenges arise around reasoning, collaboration, and understanding others' minds. I'll share three directions that move us beyond next-word prediction and towards social agents: collaborative reasoning, where agents learn to communicate, coordinate, and build on each other's ideas; mental modeling (theory of mind), the ability to represent what others know, believe, or intend; and social preference alignment, where models learn human values not just from isolated feedback but through extended, context-rich interaction. Together, these efforts aim to build agents that not only converse fluently but also reason jointly, interpret intentions, and evolve toward more adaptive, long-horizon social intelligence.

Bio: Asli Celikyilmaz is a Senior Research Manager at Fundmendals AI Research (FAIR). Formerly, she was Senior Principal Researcher at Microsoft Research (MSR) in Redmond, Washington. She is also an Affiliate Associate Member at the University of Washington. She has received Ph.D. Degree in Information Science from University of Toronto, Canada, and later continued her Postdoc study at Computer Science Department of the University of California, Berkeley. Her research interests are mainly in deep learning and natural language, specifically on language generation with long-term coherence, language understanding, language grounding with vision, and building intelligent agents for human-computer interaction She is serving on the editorial boards of Transactions of the ACL (TACL) as area editor and Open Journal of Signal Processing (OJSP) as Associate Editor. She has received several "best of" awards including NAFIPS 2007, Semantic Computing 2009, CVPR 2019, EMNLP 2023.

Keynote Talk

Giving Voice and Face to AI

Joon Son Chung
KAIST



November 22nd, 2025 – Time: **09:10 - 10:10** – Room: **2F, Space M Session Room / 1F, R117**

Abstract: As AI systems advance, building natural and intuitive multimodal interfaces is becoming increasingly critical. This talk examines technologies that equip AI with both a voice and a face, improving their capacity for seamless, expressive communication with humans. We will discuss how incorporating visual and linguistic signals into speech synthesis enables alignment between acoustic output, facial and textual attributes, yielding more natural and expressive speech generation. Our recent work synthesises speech directly from visual inputs, enabling communication where audio signals are limited or absent. In parallel, we present our talking head synthesis system, where audio inputs generate lifelike facial animations, effectively giving a face to the AI's voice and enriching the multimodal interaction.

Bio: Joon Son Chung is an associate professor at the School of Electrical Engineering, KAIST, where he is directing the Multimodal AI Lab. Previously, he was a research team lead at Naver Corporation, where he managed the development of speech recognition models for various applications including Clova Note and LINE CLOVA AI Speaker. He received his BA and PhD from the University of Oxford, working with Prof. Andrew Zisserman. His work has been published in top-tier venues such as TPAMI and IJCV, and he has received several paper awards, including at Interspeech and ACCV. His research interests include speaker recognition, multimodal learning, visual speech synthesis and audio-visual speech recognition. He is a co-author of the well-known audio-visual dataset for human speech, VoxCeleb. According to Google Scholar, his work has accumulated over 17,000 citations.

Table of Contents

Training a Chinese Listenability Model Using Word2Vec to Predict the Difficulty of Spoken Texts
Yen-Hsiang Chien, Hou-Chiang Tseng, Kuan-Yu Chen and Yao-Ting Sung 1

Cubicpower Agentic Mixture of Experts(AMoE) Framework for Fine-Tuning NLP Tasks Without GPUs
Chao-Yih Hsia 11

Design and Evaluation of a Courtroom Examination AI Simulation System with Behavioral Fidelity
Hsien-Jyh Liao 20

Multimodal Approaches for Stress Recognition: A Comparative Study Using the StressID Dataset
Chia-Yun Lee, Matúš Pleva, Daniel Hladek and Ming-Hsiang Su 29

Beyond Binary: Enhancing Misinformation Detection with Nuance-Controlled Event Context
Elijah Frederick Albertson, Retnani Latifah and Yi-Shin Chen 35

A Preliminary Study of RAG for Taiwanese Historical Archives
Claire Lin, Bo-Han Feng, Xuanjun Chen, Te-Lun Yang, Hung-Yi Lee and Jyh-Shing Roger Jang 45

Bridging Underspecified Queries and Multimodal Retrieval: A Two-Stage Query Rewriting Approach
Szu-Ting Liu, Wen-Yu Cho, Hsin-Wei Wang and Berlin Chen 63

The Study of a Traffic Accident Information Collection Agent System Based on Fine-tuned Open-Source Large Language Models
Jo-Chi Kung and Chia-Hui Chang 71

Automatic Generation of Corpus-Based Exercises Using Generative AI
Adrian Jan Zasina 80

Diversity is the Key: Enhancing LLM-based Post-processing for Automated Audio Captioning
Seyed Ali Farokh, Mohammad Mehdi Homayounpour and Ahmad Nickabadi 87

Memory-Efficient Training for Text-Dependent SV with Independent Pre-trained Models
Seyed Ali Farokh and Hossein Zeinali 95

Information-theoretic conditioning in terminological alternations in specialized domains: The cases of Taiwan Mandarin legal language and English biomedical language
Po-Hsuan Huang and Hsuan-Lei Shao 103

Voice Spoofing Detection via Speech Rule Generation Using wav2vec 2.0-Based Attention
Qian-Bei Hong, Yu-Chen Gao, Yu-Ying Xiao, Yeou-Jiunn Chen and Kun-Yi Huang 108

Computational Approaches to Quantitative Analysis of Pause Duration in Taiwan Mandarin
I-Ping Wan, Yu-Ju Lai and Pu Yu 116

A Novel Chinese-Idiom Automatic Error Correction Method Based on the Hidden Markov Model
Rongbin Zhang, Anlu Gui, Peng Cao, Lingfeng Wu, Feng Huang and Jiahui Li 124

Toward Traditional Chinese ModernBERT: A Preliminary Study
Yi-En Chen, Qiao-Ying He and Kuan-Yu Chen 133

Effective Speaker Diarization Leveraging Multi-task Logarithmic Loss Objectives
Jhih-Rong Guo, Tien-Hong Lo, Yu-Sheng Tsao, Pei-Ying Lee, Yung-Chang Hsu and Berlin Chen 140

Leveraging Weak Segment Labels for Robust Automated Speaking Assessment in Read-Aloud Tasks
Yue-Yang He and Berlin Chen 146

Exploring the Feasibility of Large Language Model- and Rubric-Based Automatic Assessment of Elementary Students' Book Summaries
Qi-Zhen Huang, Hou-Chiang Tseng and Yao-Ting Sung 153

From Scarcity to Scalability: Lexicon and Grammar Enhanced Amis to Mandarin Translation with GPT Models
Joseph Lin, Kai-Ying Lin and Hung-Yu Kao 167

CLiFT-ASR: A Cross-Lingual Fine-Tuning Framework for Low-Resource Taiwanese Hokkien Speech Recognition
Hung-Yang Sung, Chien-Chun Wang, Kuan-Tang Huang, Tien-Hong Lo, Yu-Sheng Tsao, Yung-Chang Hsu and Berlin Chen 176

MINAS: Mandarin Intelligent Narrative Assessment of Syntax for Children
Ruei-Ru Wang, Ya-Sin Li, Yi-Shuo Yin, Tao-Yu Chen, Hint-Tat Cheung and Ching-Tai Chen 184

LOBSTER: Linguistics Olympiad Benchmark for Structured Evaluation on Reasoning
Da-Chen Lian, Ri-Sheng Huang, Pin-Er Chen, Chunki Lim, You-Kuan Lin, Guan-Yu Tseng, Zhen-Yu Lin, Pin-Cheng Chen and Shu-Kai Hsieh 193

Cross-user Collaborative and Sequential Modeling for Recommendation
Qiao-Ying He, Yi-En Chen and Kuan-Yu Chen 230

Structured vs. Unstructured Inputs in LLMs: Evaluating the Semantic and Pragmatic Predictive Power in Abnormal Event Forecasting
Jou-An Chi and Shu-Kai Hsieh 237

Embodiment in Multimodal Semantics: Comparing Sensory, Emotional, and Visual Features in Chinese Color Metaphors
Yufeng Wu and Meichun Liu 249

Language Modeling Using Entanglement Enhanced Tensor Trains
Ellis Reyes and Yi-Shin Chen 258

Multimodal Fake News Detection Combining Social Network Features with Images and Text
Lawrence Yung Hak Low, Yen-Tsang Wu, Yan-Hong Liu and Jenq-Haur Wang 266

Speech-Driven Editing System for Chinese ASR Errors
Sji-Jie Ding, Chia-Hui Chang and Zi-Xuan Jian 277

A Fake News Detection Model Utilizing Graph Neural Networks to Capture Writing Styles
Yen-Tsang Wu, Lawrence Y. H Low and Jenq-Haur Wang 286

Revisiting Pre-trained Language Models for Conversation Disentanglement
Tung-Thien Lam and Cheng-Zen Yang 296

Multilingual Promise Verification in ESG Reports with Large Language Model Performance Evaluation
Wei-Chen Huang, Hsin-Ting Lu, Wen-Ze Chen and Min-Yuh Day 303

Exploring Sentence Stress Detection using Whisper-based Speech Models
Ting-An Hung, Yu-Hsuan Hsieh, Tien-Hong Lo, Yung-Chang Hsu and Berlin Chen 314

Integrating Sequential Information and Graph Structures for Anti-Money Laundering Anomaly Detection
Yin-Ju Wu, Gavin Tseng and Berlin Chen 320

A Multi-faceted Statistical Analysis for Logit-based Pronunciation Assessment
Chieh-Ren Liao and Berlin Chen 326

Learning User Common Interests for Unseen Group Recommendation
Yu-Ting Cheng, Pin-Hsin Hsiao, Chiou-Shann Fuh and Pu-Jen Cheng 334

Introduction: Persuasive Language in the Age of AI
Siaw-Fong Chung 342

Stance and Cohesion: The Use of However and While in AI-Human Argumentative Discourse
Yu-Che Yen and Siaw-Fong Chung 348

Quantum Perspectives on Persuasive Language in AI-Generated News: A QNLP-Based Analysis
Jung-Hua Liu 358

Interpretation of the level of ANGER in discussion forum
Suet Ching Soon 369

ROCLING-2025 Shared Task: Chinese Dimensional Sentiment Analysis for Medical Self-Reflection Texts
Lung-Hao Lee, Tzu-Mi Lin, Hsiu-Min Shih, Kuo-Kai Shyu, Anna S. Hsu and Peih-Ying Lu . 375

CYUT-NLP at ROCLING-2025 Shared Task: Valence–Arousal Prediction in Physicians' Texts Using BERT, RAG, and Multi-Teacher Pseudo-Labeling
Yi-Min Jian, An Yu Hsiao and Shih-Hung Wu 381

NTULAW at ROCLING-2025 Shared Task: Domain-Adaptive Modeling of Implicit Emotions in Medical Reflections
Sieh-Chuen Huang and Hsuan-Lei Shao 390

TCU at ROCLING-2025 Shared Task: Leveraging LLM Embeddings and Ensemble Regression for Chinese Dimensional Sentiment Analysis
Hsin-Chieh Li and Wen-Cheng Lin 399

Hey Vergil at ROCLING-2025 Shared Task: Emotion-Space-Based System for Doctors' Self-Reflection Sentiment Analysis
Ting-Yi Lin, Cong-Ying Lin and Jui-Feng Yeh 407

KOLab at ROCLING-2025 Shared Task: Research on Emotional Dimensions in Chinese Medical Self-Reflection Texts
Chia-Yu Chan, Chia-Wen Wang and Jui-Feng Yeh 413

SCUNLP at ROCLING-2025 Shared Task: Systematic Guideline Refinement for Continuous Value Prediction with Outlier-Driven LLM Feedback
Hong Rui Pan and Jheng Long Wu 418

Taiwanese Hakka Across Taiwan Corpus and Formosa Speech Recognition Challenge 2025 – Dapu & Zhao'an Accents
Yuan-Fu Liao, Chih-Chung Kuo, Chao-Shih Huang, Yu-Siang Lan, Han-Chun Lai and Wen-Han Hsu 427

Speech Recognition for Low-resource Languages: A Comparative Study on Hakka Han Characters and Romanization
Yu-Hsiang Cheng and Yi-Syuan Wu 435

Applying Whisper Fine-tuning and Branchformer to Hakka Speech Recognition
Yu-Sheng Huang, Wei-Cheng Hong, Xin-Yu Chen and Szu-Yin Lin 441

Improving Low-Resource Speech Recognition with Whisper-MoE and Synthetic Data Augmentation: A Case Study on Hakka
Yuan-Chi Hsu, Liang-Chun Fang and Hong-Jie Dai 446

Whisper Finetuning For Hakka Recognition in Low Resource
Min Han Teng, Ci Dao Chen, You Ting Lin and Bing Jhih Huang 450

Hakka Speech Recognition with Whisper and Pinyin Post-processing for FSR-2025
Chia-Hsin Lee, Yung-Jun Chang, Jin-Yan Wu and Kuan-Yu Chen 454

A Study on a Low-Resource Speech Recognition System for Taiwan Hakka Based on Whisper and LoRA
Zheng-Ting Liu, Heng-You Wang, Yi-Xiang Liao, Zhong-Yuan Qiu and Zhao-Yi Huang 459

A Compact Whisper+LoRA Baseline for Taiwanese Hakka ASR in FSR-2025
Hung-Ting Hsieh 467

Optimizing Whisper Parameters and Training Data Processing for Formosa Speech Recognition Challenge 2025 - Hakka ASR II
Jhen-Hao Lee, Sheng-Wei Kuo, An-Che Cheng, Bing-Hua Chen and Yi-An Liu 471

The EZ-AI System for Formosa Speech Recognition Challenge 2025
Yu-Sheng Tsao, Hung-Yang Sung, An-Ci Peng, Jhih-Rong Guo and Tien-Hong Lo 476

A Multi-Module Error Detection and Correction System for Hakka ASR
Min-Chun Hu, Yu-Lin Xiao and Wen-Hsiang Lu 481

A Whisper-Based System with Multi-Faceted Data Augmentation for Low-Resource Language
Pin-Cheng Chen, Yu-Chi Chen, Chia-Chun Liang, Cheng-Yu Lin, Ping-Juei Tsai and Wei-Yun Ma 489

A Channel-Aware Anomaly-Guided Data Augmentation Framework for the FSR-2025 Hakka Speech Recognition Challenge
Siang-Ting Lin, Arthur Hao, Chiun-Yu Hua, Kuan-Tang Huang and Berlin Chen 499

The AS-SLAM system for Formosa Speech Recognition Challenge 2025
Chih-Hsi Chen, Pei-Jun Liao, Chia-Hua Wu, Pang-Cheng Wu and Hsin-Min Wang 504

Challenges and Limitations of the Multilingual Pre-trained Model Whisper on Low-Resource Languages: A Case Study of Hakka Speech Recognition
Pei-Chi Lan, Hsin-Tien Chiang, Ting-Chun Lin and Ming-Hsiang Su 512

The NPTU ASR System for FSR2025 Hakka Character/Pinyin Recognition: Whisper with mBART Post-Editing and RNNLM Rescoring
Yi-Chin Huang, Yu-Heng Chen, Jian-Hua Wang, Hsiu-Chi Wu, Chih-Chung Kuo, Chao-Shih Huang and Yuan-Fu Liao 518