

2025 28th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA 2025)

**Yogyakarta, Indonesia
12-14 November 2025**

IEEE Catalog Number:
ISBN:

CFP2508H-POD
979-8-3315-8963-9



Copyright © 2025, IEEE

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:	CFP2508H-POD
ISBN (Print-On-Demand):	979-8-3315-8963-9
ISBN (Online):	979-8-3315-8962-2
ISSN:	2163-3479

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

TABLE OF CONTENTS

Tonal Coarticulation in Jotsoma Angami Compounds Vs Non-Compounds	1
<i>Zhonei i Gwirie, Priyankoo Sarmah, Sanasam Ranbir Singh</i>	
A Deep Learning Approach to Low-Resource Sanskrit Speech Recognition Using Ctc Loss	7
<i>Suhani Singh, Amita Dev, Poonam Bansal</i>	
Production Patterns and Prosodic Features of Chinese Tones by Learners from Five Central Asian Countries	13
<i>Yuan Jia, Mingshuai Yin</i>	
Improving Multi-Speaker Transcription for Live News Broadcasts With Canary 1B and Pyannote Diarization	19
<i>Muhammad Rifqi Adli Gumay, Rahmat Bryan Naufal, Alvin Xavier Rakha Wardhana, Kurniawati Azizah</i>	
Cross-Domain Effects on Syntactic Ambiguity Processing by Chinesemandarin Speaking Children	25
<i>Yu Jue, Zhang Xiaoxue</i>	
Enhancing Indonesian Deepfake Speech Localization with Pathological Features	31
<i>Edia Zaki Naufal Ilman, Dessi Puji Lestari, Candy Olivia Mawalim</i>	
Tuning Tone With Age: Adapting Dialogue Response Generation Based on Llms and Self-Supervised Speaker Age Estimation	37
<i>Riichi Yagi, Wangzixi Zhou, Hongwei Hu, Yuta Hirano, Sakriani Sakti</i>	
Development of Acehx for Sentiment Analysis Using a Bert-Based Model	43
<i>Doni Sumito Sukiswo, Hammam Riza, Muhammad Subianto, Taufik Fuadi Abidin, Afnan</i>	
PQSR: A Speech Corpus of Polar Questions and Spontaneous Responses in Standard Chinese with Complex Intentions Annotated	49
<i>Yingyi Luo, Yue Huang, Qingke Sun, Shuwen Chen</i>	
Case Studies on Error Checking for Tagalog and Bikol Language	55
<i>Zhean Robby Ganituen, Stephen Borja, Justin Ethan Ching, Nathaniel Oco</i>	
The Influence of Emotional Prosody on Preschoolers' Perception of Mandarin Tones Under Noise: Benefits from Visual-Articulatory Cues	61
<i>Wenyu Xiang, Yindan Weng, Shuimei Wang, Ping Tang</i>	
Generation and Automatic Evaluation of Soap Notes from Medical Dialogue Using Large Language Models	67
<i>Tsubasa Saito, Rikuto Yamanaka, Yukoh Wakabayashi, Norihide Kitaoka</i>	
Measuring Emotion Preservation in Expressive Speech-to-Speech Translation	73
<i>Bagus Tris Atmaja, Toru Shirai, Sakriani Sakti</i>	
Toward Natural Emotional Text-to-Speech System with Fine-Grained Non-Verbal Expression Control	79
<i>Wangzixi Zhou, Bagus Tris Atmaja, Sakriani Sakti</i>	
Automatic Classification of Disyllabic Tone Sandhi in Wuhan Dialect Based on Functional Principal Component Analysis	85
<i>Wanping Xu, Aijun Li</i>	

Limvo: a Less Is More Approach for Visual Reasoning in Knowledge Based Visual Question Answering	91
<i>Ade Rohmat Maulana, Arie Ardiyanti Suryani, Ema Rachmawati</i>	
Exploring Rhythm Formant Analysis for Indic Language Classification	97
<i>Parismita Gogoi, Sishir Kalita, Priyankoo Sarmah, S.R Mahadeva Prasanna</i>	
Is Beijing Mandarin Stress-Timed? Examining Rhythmic Patterns in Spontaneous and Read Speech	103
<i>Zhiwei Wang, Li Aijun</i>	
The Effect of Question Intonation on Focus: A Comparative Study of Tianjin Mandarin and American English	109
<i>Binbin Sun, Shuang Yuan, Hui Feng, Aijun Li</i>	
Accent Conversion: Preserving Speaker Identity in Native English Synthesis	115
<i>Sabyasachi Chandra, Puja Bharati, Debolina Pramanik, Shyamal Kumar Das Mandal, Riya Sil</i>	
Stage-Wise Acoustic-Linguistic Fine-Tuning for Overlapped Speech Recognition: Does Ordering Matter?	121
<i>Saddam Annais Shaquille, Dessi Puji Lestari, Sakriani Sakti</i>	
Exploring the Impact of Data Quantity on Asr in Extremely Low-Resource Languages	127
<i>Yao-Fei Cheng, Li-Wei Chen, Hung-Shin Lee, Hsin-Min Wang</i>	
ThaiMRC: A Comprehensive Corpus for Advancing Machine Reading Comprehension in Thai	133
<i>Chaianun Damrongrat, Pornpimon Palingoon, Santipong Thaiprayoon, Sumonmas Thatphithakkul Vataya Chunwijitra</i>	
Chinese Learners' Processing of English Prosodic Boundaries: an Erp Study	139
<i>Ji Xiaoli, Cai Feier, Sun Pixiang, Yang Yanqin, Li Aijun</i>	
A Prosodically Annotated Bengali and Assamese Audiobook Corpus for Sentence Boundary Detection	145
<i>Priyanjana Chowdhury, Sanghamitra Nath, Utpal Sharma</i>	
Vietnamese Speech Database for No-Reference Telecommunication Quality Assessment	151
<i>Nhat H. Tran, Bao Thang Ta, Van Hai Do</i>	
Taiwanese Pos Tagging Without Training Data: an Llm Model Mergingbased Approach With Chinese Resources	157
<i>Chun-Hsuan Chen, Hsiao-Wen Chu, Yuan-Fu Liao</i>	
Bridging Disfluent to Fluent in Speech Translation: Effective Tagging and Fine-Tuning Strategies	163
<i>Yuka Ko, Katsuhito Sudoh, Satoshi Nakamura, Sakriani Sakti</i>	
Myanmar-English Code-Switching Speech Dataset: Measr	169
<i>Theingi Aye, Win Pa Pa, Hay Mar Soe Naing</i>	
Application of Data Augmentation to Reduce Session Variability in an I-Vector-Based Speaker Identification System	175
<i>Muhammad Hanan, Dessi Puji Lestari</i>	
Development of Chatbot Module in an Intelligent Tutoring System for English Language Learning Using Large Language Model	181
<i>Ziyad Dhia Rafi, Ayu Purwarianti, Samsu Sempena</i>	

A Corpus-Based Investigation of Acoustic Features Influencing Intelligibility of Super-Elderly Japanese Speech	187
<i>Meiko Fukuda, Ryota Nishimura, Norihide Kitaoka</i>	
Farmsaathi: a Retrieval-Reranking Rag Framework for Multilingual Conversational Ai in Indian Agriculture	193
<i>Devansh Kapoor, Saanya Setia, Shivam Arora, Komal Bharti</i>	
Rhythm – Syntax Interaction Across Modalities: Evidence from Chinese Learners of English	199
<i>Tian Peng, Jue Yu</i>	
A Study on Chinese Tone and Intonation Errors Among Bangladeshi Learners an Investigation Based on Monosyllabic Characters and Sentences	205
<i>Jia Jing, Wen Cao</i>	
Speech Input Interface for Electronic Medical Record Supporting Automatic Soap Generation Using Large Language Models	211
<i>Rikuto Yamanaka, Tsubasa Saito, Yukoh Wakabayashi, Norihide Kitaoka</i>	
Effects of Speech Rate and Syllable Position on the Temporal and Spectral Characteristics of Cantonese Vowels	217
<i>Chu Yan Ho, Wai-Sum Lee</i>	
Joining Diarization and Multi-Speaker Automatic Speech Recognition with Overlap Handling for Long Conversations	223
<i>Myat Aye Aye Aung, Win Pa Pa, Sakriani Sakti</i>	
Acoustic Differences Between Coronal Nasal /n/ and Lateral /l/ in Standard Chinese	229
<i>Yijing He, Wai-Sum Lee</i>	
Quality Improvement of Low-Resourced Bahasa Indonesia Expressive Speech Synthesis Using Cross-Lingual Transfer Learning and Tacotron2	235
<i>Elok Anggrayni, Aprianto Dwi Prasetyo, Dhany Arifianto</i>	
HifiDiff: Two Stream Diffusion Models for High Fidelity Speech Generation of Unseen Languages	241
<i>Lhuqita Fazry, Kurniawati Azizah, Dipta Tanaya, Ayu Purwarianti, Dessi Lestari, Sakriani Sakti</i>	
A Gop-Based Automatic Pronunciation Scoring System for Taiwanese Hakka Using Transformer Regression Models	247
<i>Yi-Chin Huang, Yu-Heng Chen, Chih-Chung Kuo, Chao-Shih Huang, Yuan-Fu Liao</i>	
Rapid Model Adaptation of Code-Switching ASR in Low-Resource, High-Noise Industrial Domains	253
<i>Tian-Yi Chen, Chih-Chung Kuo, Yu-Siang Lan, Yuan-Fu Liao, Bo-Wei Chen, Yi Liu, Ming-Hsuan Wu</i>	
Comparative Evaluation of N-Gram and Transformer Based Language Models for ECoG-Based Speech Neuroprosthesis	259
<i>Bagas Aryo Seto, Nur Ahmadi, Dessi Puji Lestari</i>	
To Hesitate is to be Proficient: Acoustics and Speech Perception of Filled Pauses in Spontaneous L2 Hindi Speech by L1 Assamese Speakers	264
<i>Sridevi Ravi, Joyshree Chakraborty, Priyankoo Sarmah</i>	
Exploring Dialects with Speech Embeddings: Insights from Two Speech Databases in Assamese and Finnish	270
<i>Tuukka Törö, Antti Suni, Leena Dihingia, Juraj Šimko, Priyankoo Sarmah</i>	

Tokyo-Type Accent Production Among the North East Indian Students	276
<i>Nozomi Tokuma, Gulab Jha, Ruubino Peseyie</i>	
Japanese Articulatory Speech Dataset Acquired with 3D Electromagnetic Articulography	282
<i>Ikeda Eri, Yoshinaga Yukiyasu, Katsurada Kouichi, Wakamiya Kohei</i>	
Variations in Nasalance Values of Oral Vowels: Nasometric Evidence from Shanghainese	288
<i>Feifan Wang, Yaohua Jin</i>	
Optimization of Large-Scale Speaker Identification System Based on I-Vector with LDA, PCA, and LSH	294
<i>Muhammad Zaydan Athallah, Dessi Puji Lestari</i>	
Voiceless Laterals in Hmar	300
<i>Ruubino Peseyie, Priyankoo Sarmah</i>	
LAFAEK-Corpus-1M+: A Large-Scale Tetun Corpus to Build a Low-Resourced LLM for Speech and Text Processing	306
<i>Yuichi Nishida, Yuto Kuroda, Satoshi Tamura</i>	
Multilingual Multi-Task Learning with Gradient Manipulation Method for Local Languages in Indonesia	311
<i>Wilbert Fangderson, Ayu Purwarianti</i>	
Neural Network-Based Speech Emotion Recognition for the Indonesian Language	318
<i>Muhammad Iqbal Asrif, Muhammad Alif Ismady, Kurniawati Azizah</i>	