

2025 International Conference on Speech Technology and Human-Computer Dialogue (SpeD 2025)

**Cluj-Napoca, Romania
19-22 October 2025**



**IEEE Catalog Number: CFP2555H-POD
ISBN: 979-8-3315-7487-1**

**Copyright © 2025 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:	CFP2555H-POD
ISBN (Print-On-Demand):	979-8-3315-7487-1
ISBN (Online):	979-8-3315-7486-4
ISSN:	2771-5779

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 Contents

1. Audio Deepfakes and Forensics

Audio Splicing Detection Using Self-Supervised Representations <i>Octavian Pascu, Horia Cucu</i>	1
Augmented Transfer Learning for Synthetic Speech Detection <i>Irina Mutică, Șerban Mihalache, Gheorghe Pop, Dragoș Burileanu</i>	7
Spoofed Speech Detection for Physical and Logical Access Applications using Deep Neural Networks <i>Cristian-Teodor Neamțu, Șerban Mihalache, Dragoș Burileanu</i>	12
Detecting Audio Deepfakes on the Edge: Lightweight SSL-Based Detection in a Browser Plugin <i>Octavian Pascu, Dan Oneață, Horia Cucu, Nicolas Müller</i>	17

2. Natural Language Processing

A Study on Metric-Human Alignment in Code Summarization by LLMs through Kolmogorov-Arnold Networks <i>Georgian Nicolae, Corneliu Burileanu</i>	23
Effects of Attention Head Pruning on Encoder-only Language Models for Multilingual Recipe Classification <i>Alin-Gabriel Angheluș, Vlad Andrei Negru, Camelia Lemnaru, Rodica Potolea</i>	29
Safeguarding Online Trust: NLP-ML Detection of Computer-Generated Reviews <i>Cornelia Ionela Bădoi</i>	34
Transcription and Classification of Compound and Special Numeral Entities Using Artificial Intelligence and Rule-Based Methods <i>Vasile Mîtruț, Ștefania Ștefănescu, Ștefan-Adrian Toma</i>	40

3. Emotion, Speaker Recognition, Diarization, Multimodal

Improved Visually Prompted Keyword Localisation in Real Low-resource Settings <i>Leanne Nortje, Dan Oneață, Gabriel Pîrlogeanu, Herman Kamper</i>	46
On the Contribution of Lexical Features to Speech Emotion Recognition <i>David Combei</i>	52
From Lab to Field: Practical Deployment and Evaluation of a Speaker Recognition System <i>Raluca-Ionela Costan, Mihai Coca, Ștefan-Adrian Toma</i>	57
Expanding and Refining RoMEMEs: A Multimodal Corpus of Romanian Memes for Advanced AI Analysis <i>Vasile Păiș, Daniela Gîfu</i>	63

4. Synthetic Speech

Latent Insights: Exploring Phoneme Diversity in Natural and Synthetic Speech through Latent Representations <i>Diptasree Debnath, Helard Becerra Martinez, Andrew Hines</i>	69
Style-Controlled VALL-E for Few-Shot Emotional German TTS <i>Rami Kammoun, Mohammed Salah Al-Radhi</i>	75
Towards Hungarian-English Code-Switching Speech Dataset Construction by using Multi Speaker-adaptive Text-to-Speech Synthesis <i>Piroska Zsófia Barta, Péter Mihajlik</i>	81
Adding Emotion Conditioning in Speech Synthesis via Multi-Term Classifier-Free Guidance <i>Radu-George Bolborici, Ana Antonia Neacșu</i>	86
Improving Speech Synthesis by Using a Cognitive View on F0 Contours <i>Doina Jitcă</i>	92

5. Automatic Speech Recognition

Impact of Text Origin and Real-Synthetic Data Ratio in TTS-Augmented Low-Resource ASR	97
---	----

Dalai Mengke, Péter Mihajlik

Open Source State-of-the-Art Solution for Romanian Speech Recognition 102
Gabriel Pîrlogeanu, Alexandru-Lucian Georgescu, Horia Cucu

Cross-lingual Transfer Learning Experiments for Arabic ASR 108
Amin Hassairi, Péter Mihajlik

6. Signal Processing

Pretrained Speech Models Learn Boundaries, Not Patterns: An Analysis of 114
Supervised vs. Unsupervised Capabilities
Abdul Rehman, Kavisha Jayathunge, Jian-Jun Zhang, Xiaosong Yang

ADNAC: Audio Denoiser using Neural Audio Codec 120
Lucian Daniel Jimon, Mircea Florin Vaida, Adriana Stan

A Robust Decomposition-Based RLS Algorithm for Echo Cancellation 125
Applications
*Radu-Andrei Otopoleanu, Camelia Elisei-Iliescu, Constantin Paleologu,
Jacob Benesty, Cristian-Lucian Stanciu, Cristian Anghel*

About the Spatial Arrangement of Acoustic Sensors for Monitoring Protected 130
Wildlife Area
Corneliu Rusu

Range of Lip Movement during the Production of Selected Consonants in 135
Kannada Speakers
Ajish K. Abraham, V. Sivaramkrishnan, N. Swapna, N. Manohar