

2023 International Conference on Speech Technology and Human-Computer Dialogue (SpeD 2023)

**Bucharest, Romania
25 – 27 October 2023**



**IEEE Catalog Number: CFP2355H-POD
ISBN: 979-8-3503-2798-4**

**Copyright © 2023 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:	CFP2355H-POD
ISBN (Print-On-Demand):	979-8-3503-2798-4
ISBN (Online):	979-8-3503-2797-7
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. Speech Analysis, Representation and Models. Audio Signal Processing and Applications

Liveness Detection – Automatic Classification of Spontaneous and Pre-recorded Speech for Biometric Applications <i>Cristian-Teodor Neamtu, Serban Mihalache, Dragos Burileanu</i>	1
How F0 Contours Reflect Speech Nuclear Events <i>Doina Jitcă</i>	6
Automatic Audio Upmixing Based on Source Separation and Ambient Extraction Algorithms <i>Marian Negru, Bogdan Moroşanu, Ana Neacşu, Dragoş Drăghicescu, Cristian Negrescu</i>	12
Regularized RLS Adaptive Algorithm with Conjugate Gradient Method <i>Cristian-Lucian Stanciu, Cristian Anghel, Camelia Elisei-Iliescu</i>	18
Personalized Multi-Track Leveling Algorithm <i>Bogdan Moroşanu, Marian Negru, Ana Neacşu, Cristian Negrescu, Constantin Paleologu</i>	24
Classification and Identification of Certain Types of Car Accidents Based on Sound Information <i>Tudor-Vasile Şerban-Moga, Lacrimioara Grama, Corneliu Rusu</i>	30

2. Special Session: Research Advances in Child Speech Technologies

Synthetic Speaking Children – Why We Need Them and How to Make Them <i>Muhammad Ali Farooq, Dan Bigioi, Rishabh Jain, Wang Yao, Mariam Yiwere, Peter Corcoran</i>	36
A Comparative Analysis between Conformer-Transducer, Whisper, and Wav2vec2 for Improving the Child Speech Recognition <i>Andrei Barcovschi, Rishabh Jain, Peter Corcoran</i>	42
Automated Child Voice Generation: Methodology and Implementation <i>Shaimaa Alwaisi, Mohammed Salah Al-Radhi, Géza Németh</i>	48
Improved Child Text-to-Speech Synthesis through Fastpitch-based Transfer Learning <i>Rishabh Jain, Peter Corcoran</i>	54
Comparison of Data Augmentation Techniques on Filipino ASR for Children’s Speech <i>Cathlyn Abion, Niel Carlo Lumapag, John Cairu Ramirez, Christchelle Resulto, Crisron Rudolf Lucas</i>	60

Data Center Audio/Video Intelligence on Device (DAVID) - An Edge-AI Platform for Smart-Toys	66
<i>Gabriel Costache, Francisco Salgado, Rishabh Jain, Cosmin Rotariu, George Sterpu, Peter Corcoran</i>	

3. Multimodal Processing. Speech Analysis for Medical Applications. Speech Emotion Recognition

Fairness Constraints in Speech-to-Text Applications – Clinical Effects of Salivation	72
<i>Mike H.M. Teodorescu, Ionut Taraboanta, Sorin Andrian, Angela C. Ghiorghe, Horia-Nicolai Teodorescu, Mihaela Grigorie</i>	
Dysarthric Speech Recognition using Depthwise Separable Convolutions: Preliminary Study	78
<i>Seyed Reza Shahamiri, Krishnendu Mandal, Sudeshna Sarkar</i>	
Analysis of Formant Correlations in Emotional Speech – Do Formants Correlate?	83
<i>Horia-Nicolai Teodorescu, Stefan Andrei Gheltu</i>	
Dysarthric Speech Recognition: A Comparative Study	89
<i>Dhvani Shah, Vanshika Lal, Zihan Zhong, Qianli Wang, Seyed Reza Shahamiri</i>	
Investigation of a Specific Effect of Alcohol on Formants	95
<i>Jan Malucha</i>	
An Analysis of Large Speech Models-based Representations for Speech Emotion Recognition	100
<i>Adrian Bogdan Stănea, Vlad Strilețchi, Cosmin Strilețchi, Adriana Stan</i>	
Speech Emotion Recognition for Emergency Services	105
<i>Bogdan Marghescu, Ștefan-Adrian Toma, Luciana Morogan, Ion Bica</i>	

4. Selected Topics in Signal Processing. Deep Learning Algorithms, Architectures and Applications

Analysis of Spectral Entropy and Maximum Power of EEG as Authentication Mechanisms	111
<i>Alexandru Vulpe, Marius Zamfirache, Alexandru Caranica</i>	
Applying Multi-objective Acquisition Function Ensemble for a Candidate Proposal Algorithm	116
<i>Cristian Manolache, Cristina Maria Andronache, Alexandru Caranica, Horia Cucu, Andi Buzo, Cristian Vasile Diaconu, Georg Pelz</i>	
Experimental Analysis of Network Traffic Databases for Anomaly Detection	122
<i>Maria-Mădălina Andronache, Alexandru Vulpe</i>	
A Study on Initial Population Sampling for Multi-Objective Optimization based on Differential Evolution and Bayesian Inference	128
<i>Georgian Nicolae, Catalin Visan, Dan Curavale, Mihai Boldeanu, Horia Cucu, Andi Buzo, Georg Pelz</i>	

5. Speech and Language Resources. Spoken Language Recognition

Information Technology and Geolinguistics	133
<i>Silviu-Ioan Bejinariu, Vasile Apopei, Manuela Nevaci, Florin-Teodor Olariu, Nicolae Saramandu</i>	
Evaluating a Fine-Tuned Whisper Model on Underrepresented Romanian Speech	141
<i>Vasile Păiș, Verginica Barbu Mititelu, Radu Ion, Elena Irimia</i>	
Uncovering Bias in ASR Systems: Evaluating Wav2vec2 and Whisper for Dutch Speakers	146
<i>Marcio Fuckner, Sophie Horsman, Pascal Wiggers, Iskaj Janssen</i>	

6. Text-to-Speech Synthesis

Advancing Limited Data Text-to-Speech Synthesis: Non-Autoregressive Transformer for High-Quality Parallel Synthesis	152
<i>Mohammed Salah Al-Radhi, Omnia Ibrahim, Ali Raheem Mandeel, Tamás Gábor Csapó, Géza Németh</i>	
Enhancing End-to-End Speech Synthesis by Modeling Interrogative Sentences with Speaker Adaptation	158
<i>Ali Raheem Mandeel, Mohammed Salah Al-Radhi, Tamás Gábor Csapó</i>	
Harmonic-plus-Noise Network with Linear Prediction and Perceptual Weighting Filters for Filipino Speech Synthesis	164
<i>Camille Marie Tatoy, Johanna Lindsey Pasco, Josaphat Ira Emmanuel Benedicto, Crisron Rudolf Lucas</i>	
Modeling Irregular Voice in End-to-End Speech Synthesis via Speaker Adaptation	170
<i>Ali Raheem Mandeel, Mohammed Salah Al-Radhi, Tamás Gábor Csapó</i>	
Nonparallel Expressive TTS for Unseen Target Speaker using Style-Controlled Adaptive Layer and Optimized Pitch Embedding	176
<i>Mohammed Salah Al-Radhi, Tamás Gábor Csapó, Géza Németh</i>	
Author Index	182