

# **2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD 2021)**

**Virtual Conference  
13-15 October 2021**



**IEEE Catalog Number: CFP2155H-POD**  
**ISBN: 978-1-6654-2787-6**

**Copyright © 2021 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:    | CFP2155H-POD      |
| ISBN (Print-On-Demand): | 978-1-6654-2787-6 |
| ISBN (Online):          | 978-1-6654-2786-9 |

**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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# Table of Contents

|   |    |
|---|----|
| Automatic Segmentation of Texts based on Stylistic Features<br><i>Horia-Nicolai TEODORESCU, Speranta Cecilia BOLEA</i>  | 1  |
| Neural Networks for Automatic Environmental Sound Recognition<br><i>Svetlana SEGARCEANU, George SUCIU, Inge GAVAT</i>   | 7  |
| A Cognitive View on Intonational Meaning<br><i>Doina JITCĂ</i>  | 13 |
| A Review of Automated Intelligibility Assessment for Dysarthric Speakers<br><i>Andy HUANG, Kyle HALL, Catherine WATSON, Seyed Reza SHAHAMIRI</i>  | 19 |
| Speech under Combined Schizophrenia and Salivary Flow Alterations – Preliminary Data and Results<br><i>Horia-Nicolai TEODORESCU, Sorin ANDRIAN, Cristina Angela GHIORGHE, Stefan GHELȚU, Ionuț TĂRĂBOANȚĂ</i> | 25 |
| A Data-Reuse Approach for an Optimized LMS Algorithm<br><i>Alexandru-George RUSU, Laura-Maria DOGARIU, Silviu CIOCHINĂ, Constantin PALEOLOGU</i>  | 31 |
| Infant Vocal Tract Development Analysis and Diagnosis by Cry Signals with CNN Age Classification<br><i>Chunyan JI, Yi PAN</i>   | 37 |
| Dysarthric Speech Augmentation Using Prosodic Transformation and Masking for Subword End-to-end ASR<br><i>Mohammad SOLEYMANPOUR, Michael T. JOHNSON, Jeffery BERRY</i>  | 42 |

|   |     |
|---|-----|
| Effects of F0 Estimation Algorithms on Ultrasound-Based Silent Speech Interfaces<br><i>Pengyu DAI, Mohammed Salah AL-RADHI, Tamás Gábor CSAPÓ</i>   | 47  |
| Comparison in Suprasegmental Characteristics between Typical and Dysarthric Talkers<br>at Varying Severity Levels<br><i>Mohammad SOLEYMANPOUR, Michael T. JOHNSON, Jeffery BERRY</i>  | 52  |
| Results on the MFCC Extraction for Improving Audio Capabilities of TIAGo Service Robot<br><i>Toma TELEMBICI, Lacrimioara GRAMA, Lorena MUSCAR, Corneliu RUSU</i>  | 57  |
| Mispronunciation Detection and Diagnosis for Mandarin Accented English Speech<br><i>Subash KHANAL, Mohammad SOLEYMANPOUR, Narjes BOZORG, Michael T. JOHNSON</i>   | 62  |
| Versatility and Population Diversity of Evolutionary Algorithms in Automated Circuit<br>Sizing Applications<br><i>Cătălin VIȘAN, Octavian PASCU, Marius STĂNESCU, Horia CUCU,<br/>Cristian DIACONU, Andi BUZO, Georg PELZ</i> | 68  |
| Detecting Psychological Stress from Speech using Deep Neural Networks and Ensemble<br>Classifiers<br><i>Serban MIHALACHE, Dragos BURILEANU, Corneliu BURILEANU</i>  | 74  |
| Towards Detection of Synthetic Utterances in Romanian Language Speech Forensics<br><i>Gheorghe POP, Dragos BURILEANU</i>  | 80  |
| The MARA Corpus: Expressivity in End-to-end TTS Systems using Synthesised Speech Data<br><i>Adriana STAN, Beáta LŐRINCZ, Maria NUȚU, Mircea GIURGIU</i>   | 85  |
| Human-Machine Interaction Speech Corpus from the ROBIN project<br><i>Vasile PĂIȘ, Radu ION, Andrei-Marius AVRAM, Elena IRIMIA,<br/>Verginica BARBU MITITELU, Maria MITROFAN</i>   | 91  |
| Synthetic Speech Detection Using Neural Networks<br><i>Ricardo REIMAO, Vassilios TZERPOS</i>  | 97  |
| Project Văc: Can a Text-to-Speech Engine Generate Human Sentiments?<br><i>Shivam KULKARNI, Luis BARBADO, Jordan HOSIER, Yu ZHOU,<br/>Siddharth RAJAGOPALAN, Vijay K. GURBANI</i>  | 103 |
| Influence of Silence and Noise filtering on Speech Quality Monitoring<br><i>Rahul JAISWAL</i>   | 109 |
| Assessment of Pronunciation in Language Learning Applications<br><i>Camelia-Georgiana STATIVĂ, Adrian IFTENE, Camelia-Maria MILUȚ</i>   | 114 |
| Romanian Printed Language, Statistical Independence and the Type II Statistical Error<br><i>Alexandru DINU, Adriana VLAD</i>  | 120 |

|   |     |
|---|-----|
| A Fully Autonomous Person Re-Identification System<br><i>Roxana MIHAESCU, Mihai CHINDEA, Serban-Vasile CARATA, Marian GHENESCU, Constantin PALEOLOGU</i>  | 126 |
| Establishing a Baseline of Romanian Speech-to-Text Models<br><i>Dan UNGUREANU, Madalina BADEANU, Gabriela Catalina MARICA, Mihai DASCALU, Dan Ioan TUFIS</i>  | 132 |
| Gesture Recognition using PYTHON<br><i>Alexandru Octavian CALIN, Mihaela HNATIUC, Jan Catalin IOV</i>   | 139 |
| MARS: the First Romanian Pollen Dataset using a Rapid-E Particle Analyzer<br><i>Mihai BOLDEANU, Cristina MARIN, Dragos ENE, Luminița MARMUREANU, Horia CUCU, Corneliu BURILEANU</i>   | 145 |
| Word Embeddings for Romanian Language and Their Use for Synonyms Detection<br><i>Claudiu Marius POPESCU, Corneliu RUSU, Lacrimioara GRAMA</i>   | 151 |
| Autoregressive Articulatory WaveNet Flow for Speaker-Independent Acoustic-to-Articulatory Inversion<br><i>Narjes BOZORG, Michael T. JOHNSON, Mohammad SOLEYMANPOUR</i>  | 156 |
| Speaker Verification Experiments Using Identity Vectors, on a Romanian Speakers Corpus<br><i>Oana-Mariana NOVAC, Stefan-Adrian TOMA, Emil BUREACA</i>   | 162 |
| Dunstan Baby Language Classification with CNN<br><i>Costin Andrei BRATAN, Mirela GHEORGHE, Ioan ISPAS, Eduard FRANTI, Monica DASCALU, Silvia Maria STOICESCU, Ioana ROSCA, Florentina GHERGHICEANU, Doina DUMITRACHE, Leonard NASTASE</i> | 167 |
| An Analysis of the Data Efficiency in Tacotron2 Speech Synthesis System<br><i>Georgiana SĂRACU, Adriana STAN</i>  | 172 |
| Improvements of Speed's Romanian ASR System During ReTeRom Project<br><i>Alexandru-Lucian GEORGESCU, Horia CUCU, Corneliu BURILEANU</i>   | 177 |
| <b><i>Author Index</i></b>  | 183 |