2019 22nd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA 2019)

Cebu, Philippines
25 – 27 October 2019

IEEE Catalog Number: CFP1908H-POD
ISBN: 978-1-7281-2450-6

Copyright © 2019 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:	CFP1908H-POD
ISBN (Print-On-Demand):	978-1-7281-2450-6
ISBN (Online):	978-1-7281-2449-0
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

2019 22nd Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)

S1

voisTUTOR Corpus: A Speech Corpus of Indian L2 English Learners for Pronunciation Assessment
Chiranjeevi Yarra (Indian Institute of Science, India), Aparna Srinivasan (Indian Institute of Science, India), Chandana Srinivasa (Arizona State University, India), Ritu Aggarwal (Indian Institute of Science, India), Prasanta Kumar Ghosh (Indian Institute of Science, India) 1

Automatic Pronunciation Generator for Indonesian Speech Recognition System Based on Sequence-to-Sequence Model
Devin Hoesen (PT. Prosa Solusi Cerdas, Indonesia), Fanda Yuliana Putri (PT. Prosa Solusi Cerdas, Indonesia), Dessi Lestari (Institut Teknologi Bandung, Indonesia) 7

Indian Languages Corpus for Speech Recognition
Joyanta Basu (CDAC, Kolkata, India), Soma Khan (CDAC, Kolkata, India), Rajib Roy (CDAC, Kolkata, India), Babita Saxena (CDAC, Noida, India), Dipankar Ganguly (CDAC, Noida, India), Sunita Arora (CDAC, Noida, India), Karunesh Kumar Arora (CDAC, Noida, India), Shweta Bansal (KIIT College of Engineering Gurgaon, India), Shyam Sunder Agrawal (KIIT College of Engineering Gurgaon, India) 13

X-vectors Based Urdu Speaker Identification for Short Utterances
Muhammad Umar Farooq (Center for Language Engineering, Al-Khawarizimi Institute of Computer Sciences, UET Lahore, Pakistan), Farah Adeeba (Center for Language Engineering, Al-Khawarizimi Institute of Computer Sciences, Pakistan), Sarmad Hussain (University of Engineering & Technology, Lahore, Pakistan) 19

A Great Reduction of WER by Syllable Toneme Prediction for Thai Grapheme to Phoneme Conversion
Sittipong Saychum (NECTEC, National Science and Technology Development Agency (NSTDA), Thailand), Anocha Rugchatjaroen (NECTEC, National Science and Technology Development Agency (NSTDA), Thailand), Chai Wutiwiwatchai (National Electronics and Computer Technology Center, Thailand) 24

Fast and Accurate Capitalization and Punctuation for Automatic Speech Recognition Using Transformer and Chunk Merging
Binh Thai Nguyen (Hanoi University of Science and Technology, Vietnam), Hung Nguyen (Vietnam Artificial Intelligence System, Vietnam), Nguyen Thi Thu Hien (Thai Nguyen University of Education, Vietnam), Phuong Ngoc Pham (Thai Nguyen University, Vietnam), The-Loc Nguyen (Vietnam Artificial Intelligence System, Vietnam), Truong Do (Vietnam Artificial Intelligence System, Vietnam), Mai Luong (Institute of Information Technology, Vietnam) 29

Recognition and Translation of Code-switching Speech Utterances
Sahoko Nakayama (Nara Institute of Science and Technology, Japan), Takatomo Kano (Nara Institute of Science and Technology, Japan), Andros Tjandra (Nara Institute of Science and Technology, Japan), Sakriani Sakti (Nara Institute of Science and Technology / RIKEN AIP, Japan), Satoshi Nakamura (Nara Institute of Science and Technology, Japan) 34

S2

LOTUS-BI: a Thai-English Code-mixing Speech Corpus
Sumonmas Thatphithakkul (National Electronics and Computer Technology Center, Thailand), Vataya Chunwijitra (National Electronics and Computer Technology Center & NECTEC, Thailand), Patcharika Chootrakool (National Electronics and Computer Technology Center, Thailand), Phuttapong Sertsi (National Electronics and Computer Technology Center, Thailand), Sawit Kasuriya (National Electronics and Computer Technology Center, Thailand) 40

An Acoustic Study of Affricates Produced by L2 English Learners in Harbin
Chenyang Zhao (University of Chinese Academy of Social Sciences, China), Aijun Li (Institute of Linguistics, Chinese Academy of Social Sciences, China), Zhiqiang Li (University of San Francisco, USA), Ying Tang (Jilin University, China) 45

An Acoustic-Articulatory Database of VCV Sequences and Words in Toda at Different Speaking Rates
Shankar Narayanan (IISc Bangalore, India), Aravind Illa (IISc Bangalore, India), Nayan Anand (ICFAI Hyderabad, India), Ganesh Sinisetty (IISc Bangalore, India), Karthick Narayanan (JNU Delhi, India), Prasanta Kumar Ghosh (Indian Institute of Science, India) 51

Acquisition of English Retroflex Vowel [ɝ] by EFL Learners from Chinese Dialectal Regions- A Case Study of Beijing and Changsha
Bin Li (University Of Chinese Academy Of Social Sciences, China), Yuan Jia (Institute of Linguistics, China) 57

The Effect of Focus on Trisyllabic Syllable Duration in Mandarin
Ziyu Xiong (CASS, China), Qiguang Lin (ZIBOT, Inc., USA), Maolin Wang (Jinan University, China), Zhouyu Chen (Jinan University, China) 62

Three-year-old Children's Production of Native Mandarin Chinese Lexical Tones
Ao Chen (Beijing Language and Culture University, China), Hintat Chung (Education University of Hong Kong, Philippines), Yuchen Li (Utrecht University, Philippines), Liqun Gao (Beijing Language and Culture University, China) 67

An Analysis of Voice Quality of Chinese Patients with Depression
Yuan Jia (Institute of Linguistics, CASS, China), Yuzhu Liang (Nankai Uiversity, China), Tingshao Zhu (Institute of Psychology, Chinese Academy of Sciences, China) 72

S3

A New Corpus of Elderly Japanese Speech for Acoustic Modeling, and a Preliminary Investigation of Dialect-Dependent Speech Recognition
Meiko Fukuda (Tokushima University, Japan), Ryota Nishimura (Tokushima University, Japan), Hiromitsu Nishizaki (University of Yamanashi, Japan), Yurie Iribe (Aichi Prefectural University, Japan), Norihide Kitaoka (Toyohashi University of Technology, Japan) 78

Recent Progress of Mandrain Spontaneous Speech Recognition on Mandrain Conversation Dialogue Corpus
Yu-Chih Deng (NCTU, Taiwan), Chen-Yu Chiang (NTPU, Taiwan), Sin-Horng Chen (National Chiao Tung University, Taiwan), Yih-Ru Wang (National Chiao Tung University, Taiwan) 84

Indic TIMIT and Indic English Lexicon: A Speech Database of Indian Speakers Using TIMIT Stimuli and a Lexicon from Their Mispronunciations
Chiranjeevi Yarra (Indian Institute of Science, India), Ritu Aggarwal (Indian Institute of Science, India), Avni Rajpal (DA-IICT, India), Prasanta Kumar Ghosh (Indian Institute of Science, India) 90

Motion Detection of Articulatory Movement with Paralinguistic Information Using Real-Time MRI Movie
Asai Takuya (WASEDA University, Japan), Hideaki Kikuchi (WASEDA University, Japan), Kikuo Maekawa (National Institute for Japanese Language, Japan) 96

RSL2019: A Realistic Speech Localization Corpus
Rohan Sheelvant (Indian Institute of Technology Dharwad, Dharwad, Singapore), Bidisha Sharma (National University of Singapore, Singapore), Maulik Madhavi (Dhirubhai Ambani Institute of Information Communication and Technology, India), Rohan Kumar Das (National University of Singapore, Singapore), Mahadeva Prasanna (IIT Dharwad, India), Haizhou Li (National University of Singapore, Singapore) 102

S4

XDF-REPA: A Densely Labeled Dataset Toward Refined Pronunciation Assessment for English Learning
Zhigang Ou (New Oriental Education and Technology Group, China), Yun Gao (New Oriental Education and Technology Group, China), Jianfeng Cheng (New Oriental Education and Technology Group, China), Yong Ruan (Institute of Computing Technology, Chinese Academy of Sciences, China), Xiangdong Wang (Institute of Computing Technology, Chinese Academy of Sciences, China), Yueliang Qian (Institute of Computing Technology, Chinese Academy of Sciences xdwang, China) 108

A Large Collection of Sentences Read Aloud by Vietnamese Learners of Japanese and Native Speakers' Reverse Shadowings
Shintaro Ando (The University of Tokyo, Japan), Zhenchao Lin (The University of Tokyo, Japan), Tasavat Trisitichoke (The University of Tokyo, Japan), Yusuke Inoue (The University of Tokyo, Japan), Fuki Yoshizawa (The University of Tokyo, Japan), Daisuke Saito (The University of Tokyo, Japan), Nobuaki Minematsu (The University of Tokyo, Japan) 114

Distributing and Sharing Resources for Automatic Speech Recognition Applications
Sila Chunwijitra (National Electronics and Computer Technology Center (NECTEC), Thailand), Surasak Boonkla (National Electronics and Computer Technology Center, Thailand), Vataya Chunwijitra (National Electronics and Computer Technology Center & NECTEC, Thailand), Nattapong Kurpukdee (National Electronics and Computer Technology Center & NECTEC, Thailand), Phuttapong Sertsi (National Electronics and Computer Technology Center, Thailand), Sawit Kasuriya (National Electronics and Computer Technology Center, Thailand) 120

Challenges Posed by Voice Interface to Child-Agent Collaborative Storytelling
Ethel Ong (De La Salle University, Philippines), Junlyn Bryan Alburo (De La Salle University, Philippines), Christine Rachel De Jesus (De La Salle University, Philippines), Luisa Katherine Gilig (De La Salle University, Philippines), Dionne Tiffany Ong (De La Salle University, Philippines) 126

A Linguistic Representation Scheme for Depression Prediction - With a Case Study
Yuan Jia (Institute of Linguistics, CASS, China), Yuzhu Liang (Nankai Uiversity, China), Tingshao Zhu (Institute of Psychology, Chinese Academy of Sciences, China) 132

Phoneme-Level Speaking Rate Variation on Waveform Generation Using GAN-TTS
Mayuko Okamoto (Nara Institute of Science and Technology, Japan), Sakriani Sakti (Nara Institute of Science and Technology / RIKEN AIP, Japan), Satoshi Nakamura (Nara Institute of Science and Technology, Japan) 137

The Architecture of Speech-to-Speech Translator for Mobile Conversation
Agung Santosa (Agency for the Assessment and Application of Technology (BPPT), Indonesia), Asril Jarin (Badan Pengkajian dan Penerapan Teknologi (BPPT), Indonesia), Gunarso Gunarso (Agency for the Assessment and Application of Technology (BPPT), Indonesia), Gita Citra Puspita (Agency for the Assessment and Application of Technology (BPPT), Indonesia), Muhammad Teduh Uliniansyah (Agency for the Assessment and Application of Technology (BPPT), Indonesia), Andi Djalal Latief (Agency for the Assessment and Application of Technology (BPPT), Indonesia), Hammam Riza (Agency for the Assessment and Application of Technology (BPPT), Indonesia) 143

Sequence-to-Sequence Models for Grapheme-to-Phoneme Conversion on Large Myanmar Pronunciation Dictionary
Aye Mya Hlaing (University of Computer Studies, Yangon, Myanmar), Win Pa Pa (University of Computer Studies, Yangon, Myanmar) 149

S5

Comparison Between Read and Spontaneous Speech Assessment of L2 Korean
Seung Hee Yang (Seoul National Uniersity, Korea (South)), Minhwa Chung (Seoul National University, Korea (South)) 154

Annotation and Preliminary Analysis of Utterance Decontextualization in a Multiactivity
Haruka Amatani (National Institute of Japanese Language and Linguitstics, Japan), Yayoi Tanaka (National Institute of Japanese Language and Linguistics, Japan) 159

The Study of Prosody-Pragmatics Interface with Focus Functioning as Pragmatic Markers: The Case of Question and Statement
Siyi Cao (Nanjing University of Aeronautics and Astronautics, China), Yizhong Xu (Nanjing University of Aeronautics and Astronautics, Japan), Xiaoli Ji (Jiangsu University of Science and Technology, China) 164

Characteristics of Everyday Conversation Derived from the Analysis of Dialog Act Annotation
Yuriko Iseki (National Institute for Japanese Language and Linguistics, Japan), Keisuke Kadota (Waseda University, Japan), Yasuharu Den (Chiba University, Japan) 169

Statistical Studies on Japanese Sonority by Using Loudness Calibration Scores
Takayuki Kagomiya (National Institute for Japanese Language and Linguistics, Japan) 175

An Investigation of Prosodic Features Related to Next Speaker Selection in Spontaneous Japanese Conversation
Yuichi Ishimoto (National Institute for Japanese Language and Linguistics, Japan), Takehiro Teraoka (Takushoku University, Japan), Mika Enomoto (Tokyo University of Technology, Japan) 180