Matthew K. Perez

3844 Bob and Betty Beyster Building, Ann Arbor, MI 48103
Email: mkperez@umich.edu | Phone: (808) 286-4393
Website: https://matthewkperez.com/

RESEARCH INTEREST

My research interests center around leveraging the rich signal of human speech for characterizing health, behavior, and mood using machine learning models. My research has touched on various topics, including automatic speech recognition, speech synthesis, acoustic feature analysis, and computational paralinguistics. My thesis work largely involves developing novel approaches to enable individuals with speech impairments to communicate effectively, which has driven my passion for advancing the field of speech technology to benefit a broader range of populations.

EDUCATION

University of Michigan, Ann Arbor 2017 – 2023 (*expected)*
Ph.D. in Computer Science and Engineering
Advisor: Dr. Emily Mower Provost

University of Michigan, Ann Arbor 2017 – 2019
M.S. in Computer Science and Engineering
Advisor: Dr. Emily Mower Provost

University of Notre Dame 2013 – 2017
B.S. in Computer Science, *Cum Laude*

AWARDS

NSF Graduate Research Fellowship (GRFP) 2020 – 2023
GEM Ph.D. Engineering Fellowship 2019 – 2020
Dean's List, University of Notre Dame 2015 – 2017
Balfour-Hesburgh Scholar, University of Notre Dame 2013 – 2017

RESEARCH EXPERIENCE

Computational Human Artificial Intelligence (CHAI) Lab 2017 – Present
University of Michigan Ann Arbor, MI
Working on speech and signal processing methods for analyzing various neurological conditions.

Research Assistant: Mobile Computing Lab 2015-2016
University of Notre Dame Notre Dame, IN
Investigated speech analysis methods for classifying mild traumatic brain injuries (concussions).

PUBLICATIONS

Yang Yu, **Matthew Perez**, Ankur Bapna, Fadi Haik, Siamak Tazari, Yu Zhang. "PronScribe: Highly accurate multimodal phonemic transcription from speech and text". Interspeech 2023 (In submission)

James Tavernor, **Matthew Perez**, Emily Mower Provost. "Episodic Memory For Domain-Adaptable, Robust Speech Emotion Recognition". Interspeech 2023 (In submission)

Matthew Perez, Mimansa Jaiswal, Minx Niu, Cristina Gorrostieta, Matthew Roddy, Kye Taylor, Reza Lotfian, John Kane, Emily Mower Provost. "Mind the gap: On the value of silence representations to lexical-based speech emotion recognition". Interspeech 2022. (p*oster presentation*)

Amrit Romana, Minxue Niu, **Matthew Perez**, Angela Roberts, Emily Mower Provost. "Enabling Off-the-Shelf Disfluency Detection and Categorization for Pathological Speech". Interspeech 2022. (oral *presentation*)

Matthew Perez, Amrit Romana, Noelle Carlozzi, Praveen Dayalu, Jennifer Ann Miner, Angela Roberts, and Emily Mower Provost. "Articulatory Coordination for Speech Motor Tracking in Huntington Disease" Interspeech 2021. (*oral presentation*)

Amrit Romana, John Bandon, **Matthew Perez**, Stephanie Gutierrez, Richard Richter, Angela Roberts, Emily Mower Provost. “Automatically Detecting Errors and Disfluencies in Read Speech to Predict Cognitive Impairment in People with Parkinson's Disease”. Interspeech 2021 (oral presentation)

Zakaria Aldeneh, **Matthew Perez**, and Emily Mower Provost. “Learning Paralinguistic Features from Audiobooks through Style Voice Conversion” NAACL 2021. (v*irtual presentation*)

Matthew Perez, Zakaria Aldeneh, and Emily Mower Provost. “Aphasic Speech Recognition using a Mixture of Speech Intelligibility Experts” Interspeech 2020. (*virtual presentation*)

Matthew Perez, Wenyu Jin, Duc Le, Noelle Carlozzi, Praveen Dayalu, Angela Roberts, and Emily Mower Provost. “Classification of Huntington’s Disease Using Acoustic and Lexical Features.” Interspeech 2018. (*oral presentation*)

Louis Daudet, Nikhil Yadav, **Matthew Perez**, Christian Poellabauer, Sandra Schneider, Alan Huebner. “Portable mTBI Assessment Using Temporal and Frequency Analysis of Speech.” IEEE Journal of Biomedical and Health Informatics 2017.

PROFESSIONAL EXPERIENCE

Research Intern 2022
Google Mountain View, CA
Worked on frontend text-to-speech focusing on fine-tuning end-to-end networks for pronunciation learning.

Research Intern 2021
Cogito Corporation Virtual
Worked on the Speech Signals team where I researched learning silence tokens within a language modeling framework (GloVe/BERT) for speech emotion recognition.

Research Intern 2019
MIT Lincoln Laboratory Lexington, MA
Worked with Thomas Quatieri in the Bioengineering and Technologies Systems Group where my research focused on analyzing speech-based articulation features for neurological diseases such as depression.

iOS Developer Intern 2016
Garmin Olathe, KS
Contributed to the development of the Garmin iOS application, which syncs wearable device data to/from the iPhone. Specifically, implemented *Today Extensions* for the Garmin Connect Mobile app, which displays wearable health information like steps, activities, etc. I presented work to research teams and company executives.

Software Developer: ND Tours 2016
University of Notre Dame Notre Dame, IN
Developed mobile app using augmented reality that overlays the camera view with the history and information about specific landmarks. Coded in objective-c and uses firebase for backend location and data storage.

TALKS

Learning silence in language models for speech emotion recognition 2022
Google-Speech Reading Group, Mountain View

Speech Intelligibility in Aphasic Speech Recognition Modeling 2020
The National GEM Consortium Technical Workshop, Virtual

Multimodal Classification of Huntington Disease, *Student Poster Award* 2018
Graduate Engineering Research Symposium, Ann Arbor

SERVICES

Reviewer, Transactions on Audio, Speech and Language Processing 2023

Reviewer, Transactions of Affective Computing 2023

Reviewer, Affective Computing & Intelligent Interaction (ACII) 2021

Reviewer, Computer Speech & Language 2020

Poster Chair, Michigan AI Symposium 2019
University of Michigan, Ann Arbor

MEMBERSHIP

Member, International Speech Communication Association 2018 – Present
Member, IEEE 2016 – Present
National GEM Fellow 2019 – Present

SKILLS

Languages: Python, Bash, C++, Objective-C
Software Frameworks: Kaldi, Speechbrain, Pytorch-kaldi, OpenSmile, Librosa, NumPy, Pandas
Machine Learning: Pytorch, Tensorflow, Keras, scikit-learn