Investigating artificial neural networks optimized for ecological auditory tasks as a normative model for pitch perception

Bryan J. Medina[1,2,4], Mark R. Saddler[2,3,4], & Josh H. McDermott[2,3,4,5]

[1] Department of Computer Science, UCF, [2] Department of Brain and Cognitive Sciences, MIT, [3] McGovern Institute for Brain Research, MIT, [4] Center for Brains, Minds, and Machines, MIT, [5] Program in Speech and Hearing Biosciences and Technology, Harvard University

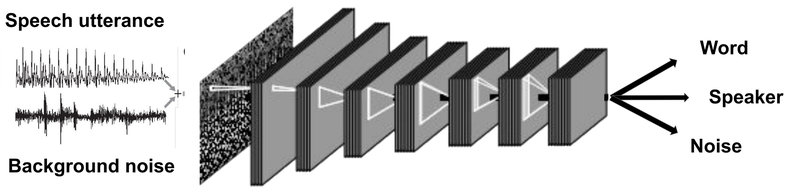
UNIVERSITY OF CENTRAL FLORIDA

Introduction

- Pitch perception is an important aspect of human hearing
- Pitch is classically envisioned as a the perceptual correlate of fundamental frequency (F0); however, humans can make accurate pitch judgments with inharmonic stimuli
- To investigate how human pitch perception may have been shaped by the demands of ecologically-important tasks, we trained deep artificial neural networks to perform different combinations of three tasks
- **HYPOTHESIS:** Artificial neural networks trained on ecologically-important tasks will learn representations of pitch that resemble those of humans

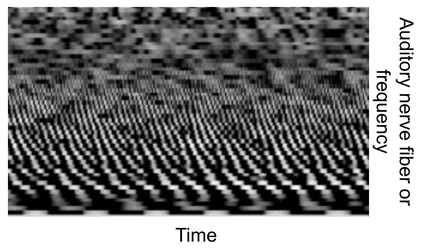
Natural auditory tasks

- Three auditory recognition tasks:
  - **Word:** "which word appeared in the middle of the utterance?"
  - **Speaker:** "which speaker made the utterance?"
  - **Noise:** "which environmental sounds appeared in the background noise?"

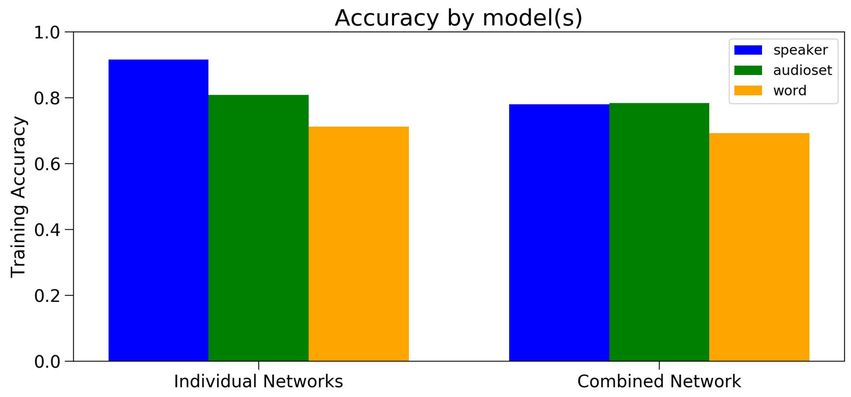


- Convolutional neural networks (CNNs) received simulated auditory nerve representations of speech-in-noise as input

Auditory nerve model:
Bruce et al., 2018 Hear. Res.



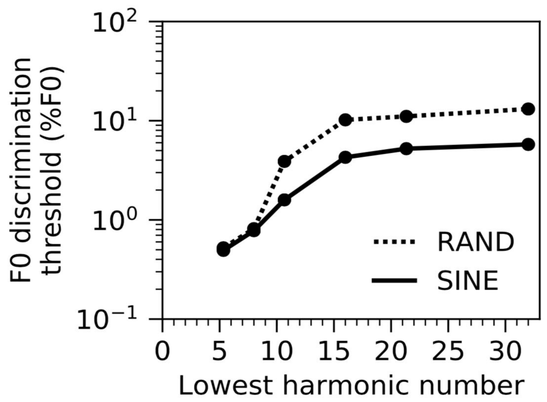
Auditory model training

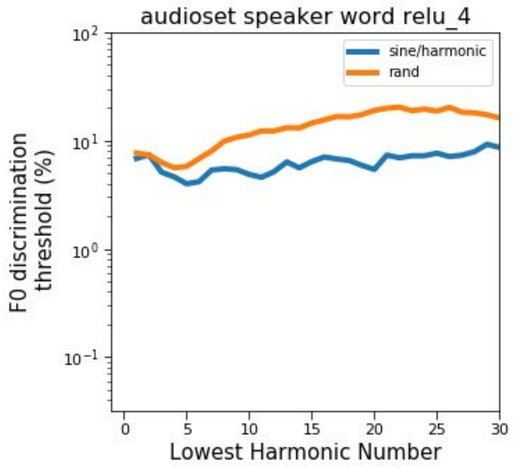


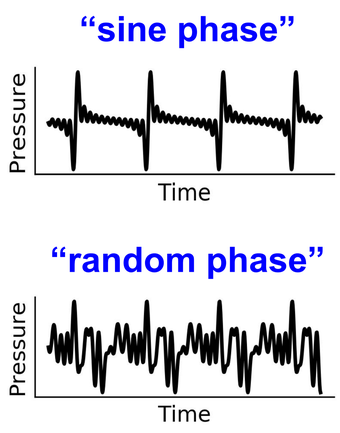
Pitch psychophysics

- Human pitch discrimination exhibits well-documented dependencies on harmonic phase and number

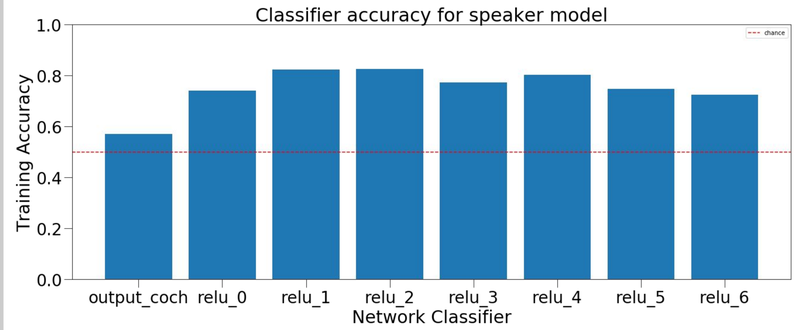
HUMAN DATA from Bernstein & Oxenham, 2005 JASA

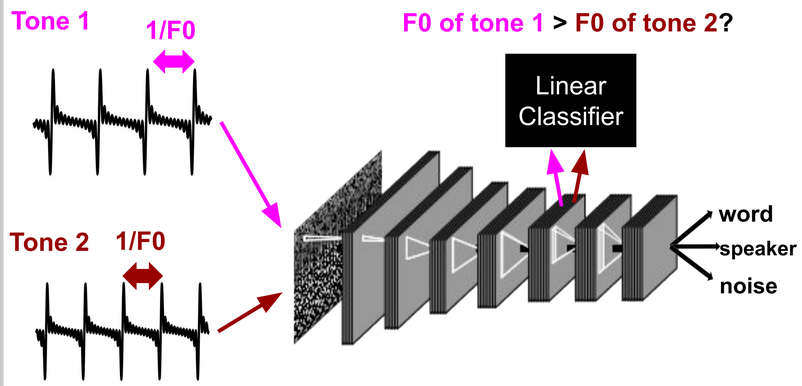






Pitch discrimination classifiers





Future directions

- Perform psychophysical experiments on trained classifiers to compare learned pitch representations to those of humans.
- Investigate how learned pitch representations depend on the auditory task optimized for
- Do the pitch mechanisms underlying speech and music tasks differ?
- Our in silico approach enables us to causally test how pitch perception may have been shaped by the demands of particular tasks

Acknowledgements

- Funded by the Center for Brains, Minds, and Machines MSRP-Bio Program
- Mandana Sassanfar (Director of MSRP-Bio)
- Ronald E. McNair Scholars Program at the University of Central Florida
- Jenelle Feather, Andrew Francl, and Ray Gonzalez for contributions to shared codebase and training dataset