Interpretable model based phonetic selectivity using high density µECoG - recordings

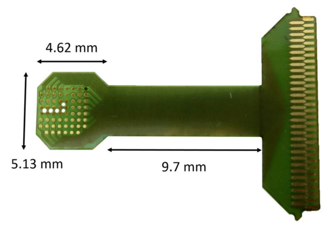
Suseendrakumar Duraivel[1], Chia-Han Chiang[1], Michael Trumpis[1], Charles Wang[1], Katrina Barth[1], Derek Southwell[2], Saurabh R. Sinha[3], Jonathan Viventi[1], Gregory B. Cogan[2]
Department of Biomedical Engineering[1], Neurosurgery[2], and Neurology[3], Duke University

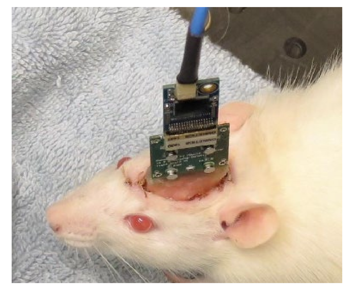
Introduction

- Understanding speech and language in the brain will help build reliable speech prosthetics
- **Phonemes** are fundamental units of speech and language (e.g., /m/ & /p/ in mad vs. pad) [1]
- Human posterior Superior Temporal Gyrus (pSTG) encodes spectrotemporal properties of phonetic information, but its fine-grained spatial resolution is unknown [2]
- **Micro-electrocorticography (µECoG)** can sample the pSTG at high spatial resolution
- Here, we combine µECoG with **interpretable decision-tree analysis** to:
  - Validate the technique in rat auditory cortex where the spatial tonotopy is well established
  - Establish **phoneme spatial maps** in human pSTG

Methods

Rat



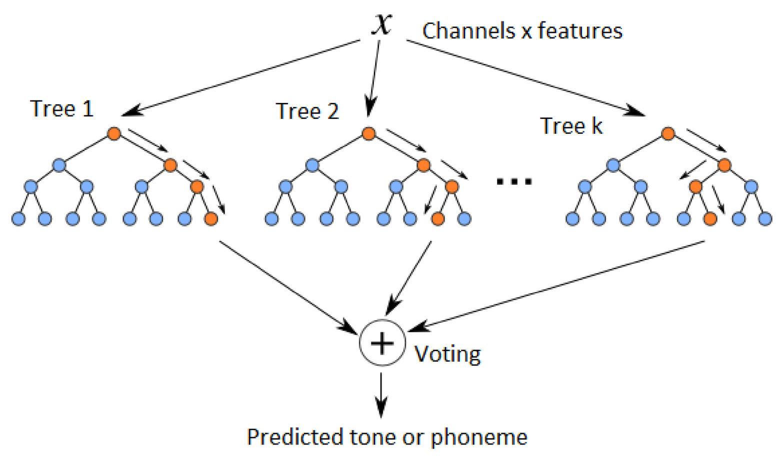


- **61-channel µECoG array**
  - 8 x 8 grid
  - 200 µm diameter electrodes
  - 400 µm spacing
- **Electrode Placement**
  - Implanted epidurally over primary auditory cortex
- **Stimuli**
  - Measured neural responses to 13 different tones (0.5-32 kHz, 0.5 octave spacings, 50 ms in duration)
- **Neural Features**
  - Broadband (2 – 100 Hz) evoked responses in 50 ms window, 200 ms after each tone onset

Human

- **256-channel µECoG array**
  - 12 x 24 grid
  - 200 µm diameter electrodes
  - 1.72 mm spacing
- **Electrode Placement**
  - Implanted subdurally over the pSTG of a patient suffering from epilepsy during resective surgery
- **Stimuli**
  - 58 sentences selected from TIMIT database (Garofolo *et al.*, 1993)
- **Neural Features**
  - High gamma (70 – 150 Hz) power integrated in 100 ms window centered at 150 ms post phoneme onset

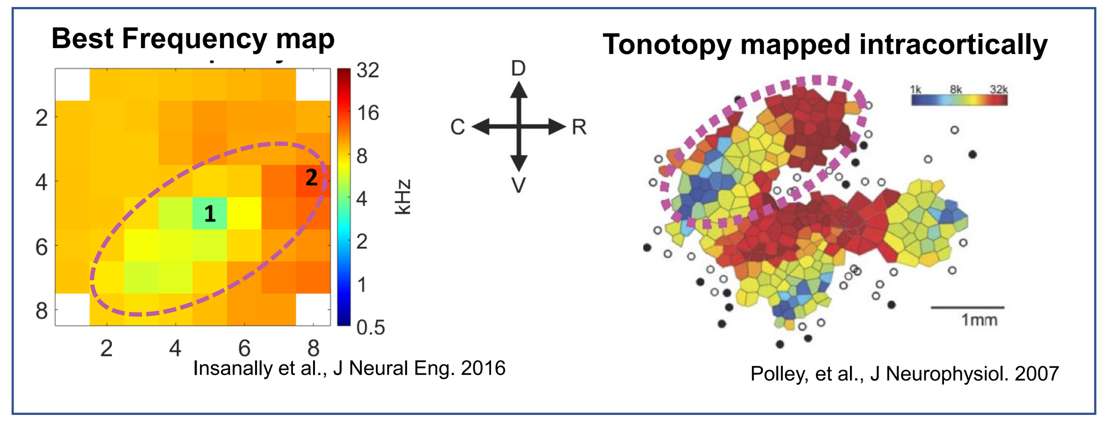
Decision Trees as Interpretable Models

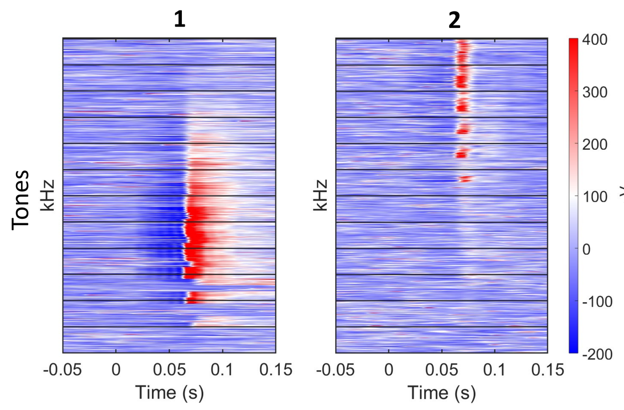


- Ensemble decision tree can quantify the importance of ECoG electrodes in decoding tones or phonemes
- The electrode importance can be interpreted as a **spatial selectivity metric** for tonotopy and phoneme maps

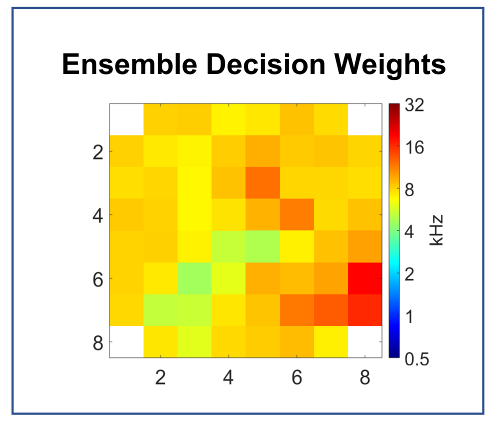
Rat µECoG Tonotopy

Control



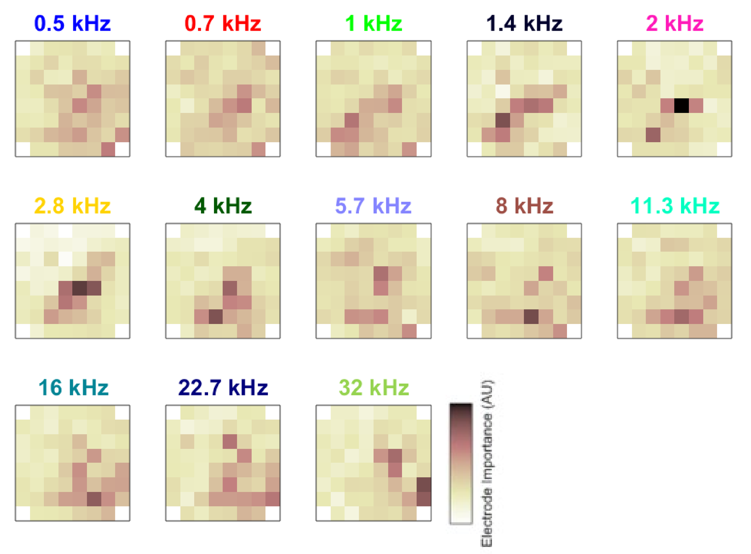


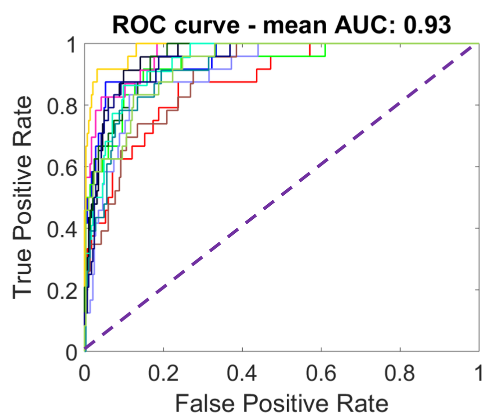
Experiment



- Subdural acute recording
- Responses to 50ms tones:
  - 0.5 – 32 kHz
  - 70 dB
- Clear tonotopic organization
- Best frequency map using:
  - Mahalanobis distance
  - Ensemble decision weights

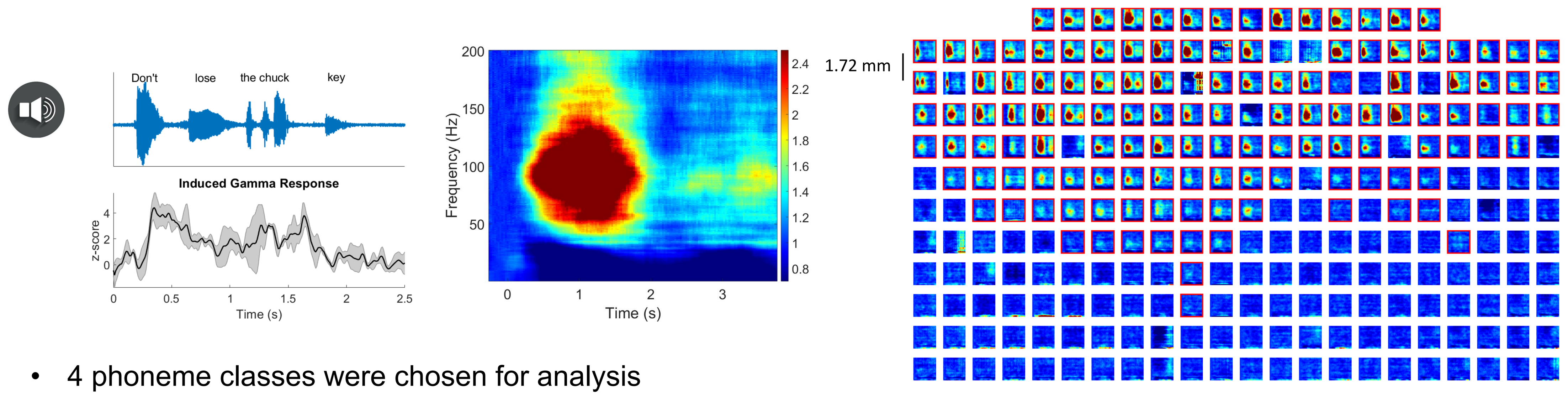
Decision weights quantify channel importance for each tone



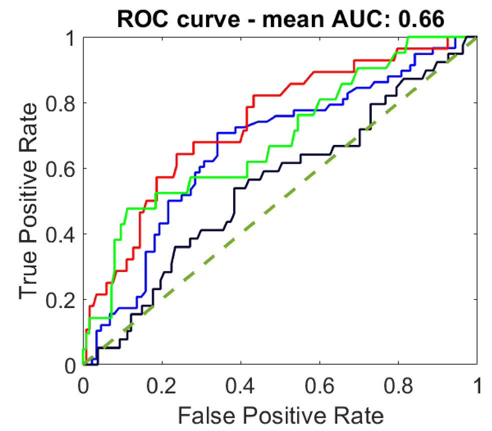


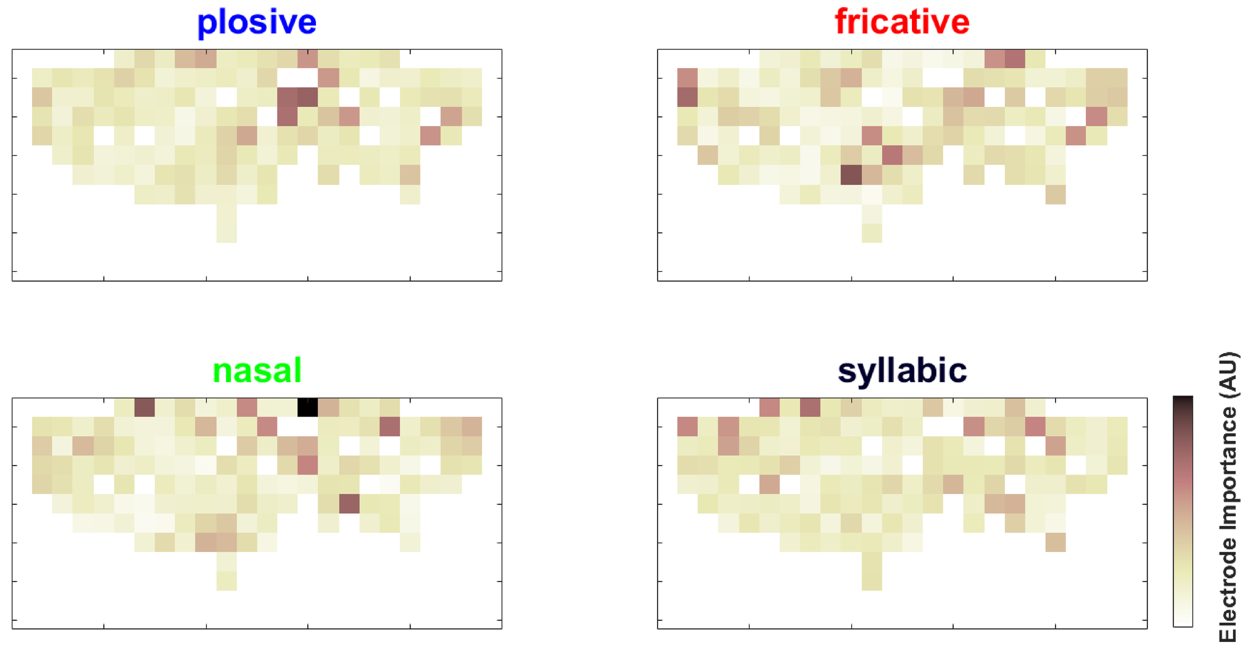
Human µECoG Phoneme Selectivity

- Multi-taper spectrograms reveal **high gamma (70 – 150 Hz)** responses at the onset of each sentence.
  - 132 significant channels (1 – sided permutation test p < 0.05, FDR corrected)



- 4 phoneme classes were chosen for analysis
  - plosive: b, d, k, p, t
  - fricative: s, z
  - nasal: m, n
  - syllabic: ɚ, ɑ, oʊ, ə, ɪ̈
- Phoneme attribute maps from decision trees
  - Decision weights characterize the **spatial importance** of electrodes





Discussion

- Ensemble decision trees can characterize the µECoG electrode importance in decoding tones and phonemes
  - In rat, tone maps established using decision weights agrees with the intracortically mapped tonotopy
  - In humans, decision tree weights identify spatially important electrodes responsive to phoneme attributes
- In the future, the spatial importance metric from decision tree models will enable effective cortical targeting of BCI devices

References

[1] Liberman, Alvin M., et al. "The discrimination of speech sounds within and across phoneme boundaries." *Journal of experimental psychology* 54.5 (1957): 358.
[2] Hullett, Patrick W., et al. "Human superior temporal gyrus organization of spectrotemporal modulation tuning derived from speech stimuli." *Journal of Neuroscience* 36.6 (2016): 2014-2026.
[3] Mesgarani, Nima, et al. "Phonetic feature encoding in human superior temporal gyrus." *Science* 343.6174 (2014): 1006-1010.
[4] Insanally, Michele, et al. "A low-cost, multiplexed µECoG system for high-density recordings in freely moving rodents." *Journal of neural engineering* 13.2 (2016): 026030.
[5] Polley, Daniel B., et al. "Multiparametric auditory receptive field organization across five cortical fields in the albino rat." *Journal of neurophysiology* 97.5 (2007): 3621-3638.

ACKNOWLEDGEMENTS: We would like to thank Anna Thirakul for help with consenting participants, Seth Foster for his help with the task design, and Shervin Rahimpour for help with surgical preparation.