Neural Representation of Noisy Reverberant Speech in Human Auditory Cortex

Krishna Puvvada¹, Jonathan Z Simon¹,²,³
¹Dept of Electrical & Computer Engineering, ²Dept of Biology, ³Institute for Systems Research
University of Maryland College Park

Background:
- Magnetoencephalography (MEG) activity is phase-locked to temporal modulations of stimulus.
  - AM at 3 Hz
  - 3 Hz phase-locked response

- Natural speech corrupted with additive noise is represented as ‘uncorrupted’ speech in cortex and is quite robust to level of degradation. (Ding, N. & J.Z. Simon, 2013)
- Reverberation is another major source of speech degradation.
- Is the neural representation of speech corrupted by reverberation (convolutive noise) an ‘uncorrupted’ version of speech (‘clean’ model) or reverberant version itself? (‘reverb’ model).

- Reverberation causes temporal and spectral smearing.

Experimental Design
- 4 reverberant conditions.
  - No reverb, mild, medium and severe reverberation.
- 3 different noise conditions.
  - No noise, +3 dB SNR, +6 dB SNR.
- 12 conditions in total.
- 60 second long story segments, 3 repetitions.

Method:
Stimulus reconstruction
- Temporal envelope of stimulus is reconstructed from cortical responses using optimum linear filters.
- Reconstruction based on integrating neural responses over a temporal window.
- Optimum decoders are designed to reconstruct the speech envelope under each model (clean/reverb).
- Correlation between reconstructed and presumed model envelope is used as metric as to how faithfully the speech envelope is represented under presumed model.

Analysis & Results

Stimulus reconstruction using Delta band (1-4 Hz) responses
- Both models performed significantly above chance.
- Significant test shows that neural responses are more correlated with reverberant envelope than clean envelope.

Stimulus reconstruction using Theta band (4-8 Hz) responses
- Both models performed significantly above chance.
- Neither model performed significantly better than the other.

Discussion
- Delta band fluctuations in speech reflect prosody level information; theta fluctuations reflect syllabic information.
- While Delta band neural responses reflect the reverberant version of speech envelope, theta band responses do not distinguish between the models.
- As the stimulus contrast is stronger in theta band than delta, the shift away from the reverberance dominated model provides weak evidence for reverberance removal in theta band.

References:

Acknowledgements: Funding from NIH R01 DC 014085