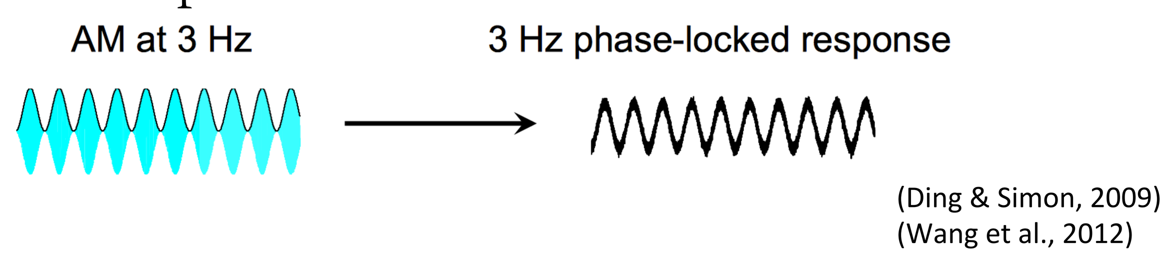
Neural Representation of Noisy Reverberant Speech in Human Auditory Cortex

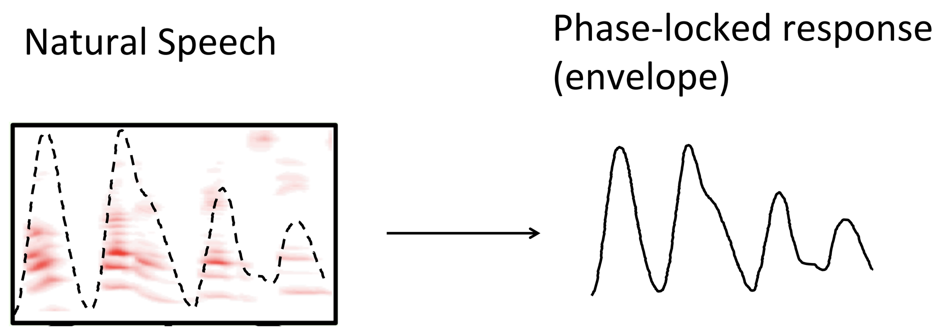
Krishna Puvvada[1], Jonathan Z Simon[1,2,3]
[1]Dept of Electrical & Computer Engineering, [2]Dept of Biology, [3]Institute for Systems Research
University of Maryland College Park

Computational Sensorimotor Systems Lab

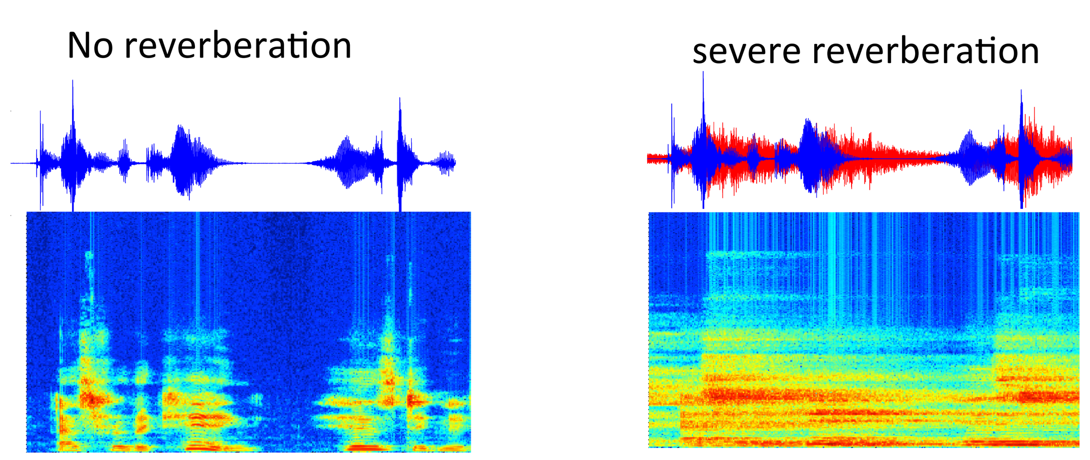
Background:

- Magnetoencephalography (MEG) activity is phase-locked to temporal modulations of stimulus.





- 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 it self? ('reverb' model).



- Reverberation causes temporal and spectral smearing.

Experimental Design

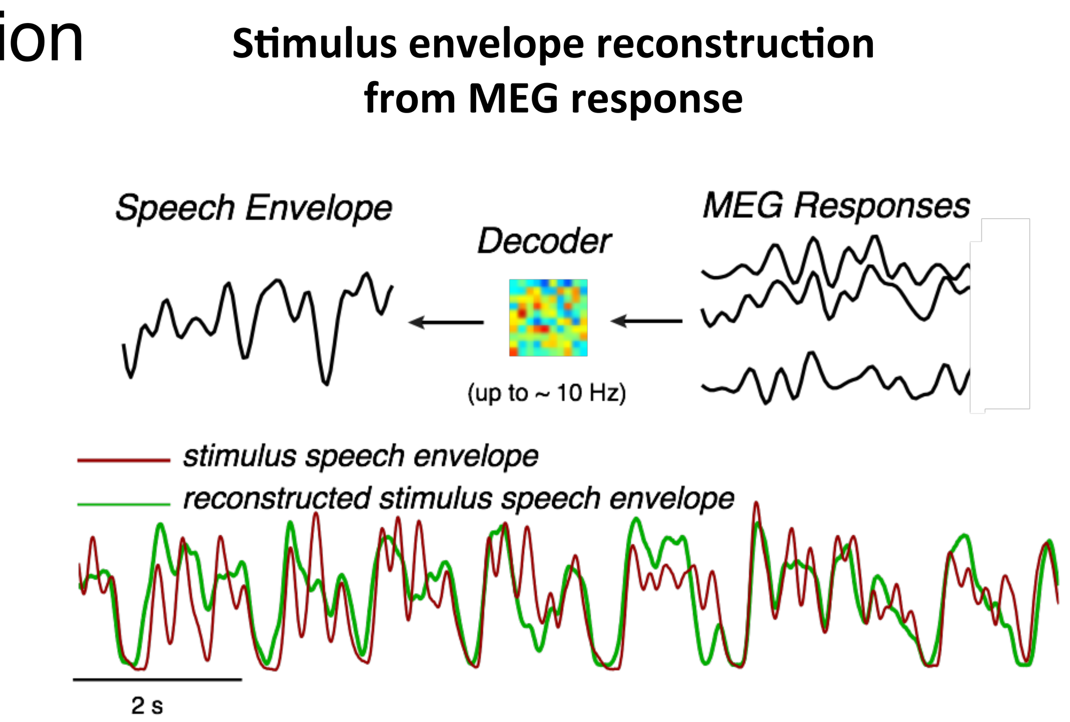
- 4 reverberant conditions.
  - No reverb, mild, medium and severe reverberation.
- 3 different noise conditions.
  - No noise, +3 dB SNR, +6 dB SNR.
- 12 conditions in total.
- 60 second long story segments, 3 repetitions.

- *N* = 12 Subjects.
- To maintain attention, in each repetition the subject counts the number of times a given keyword is heard in the story.
- MEG recording with 157 channels.
- 1kHz sampling , Time-shifted PCA based de-noising.
- Spatial filtering used to reduce 157 channels to 10, more reliable, virtual channels.

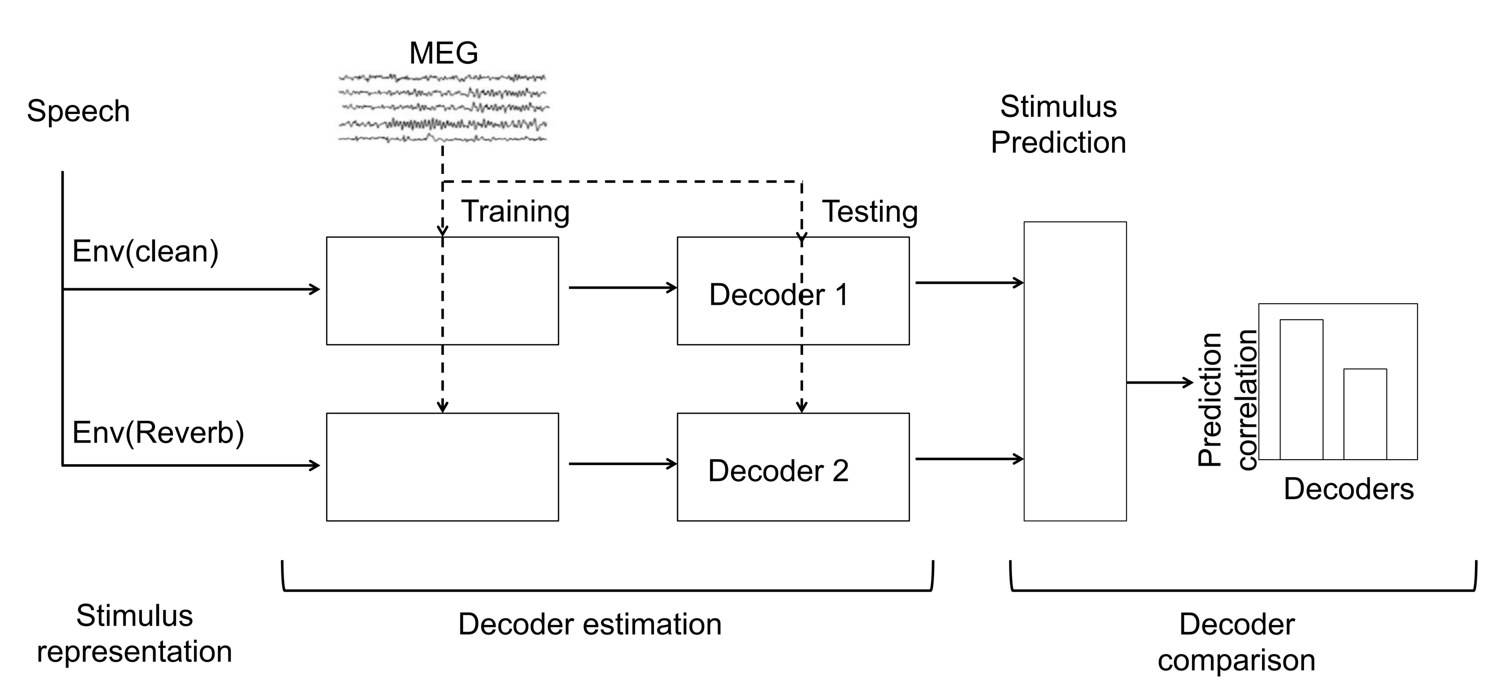
Method:

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.



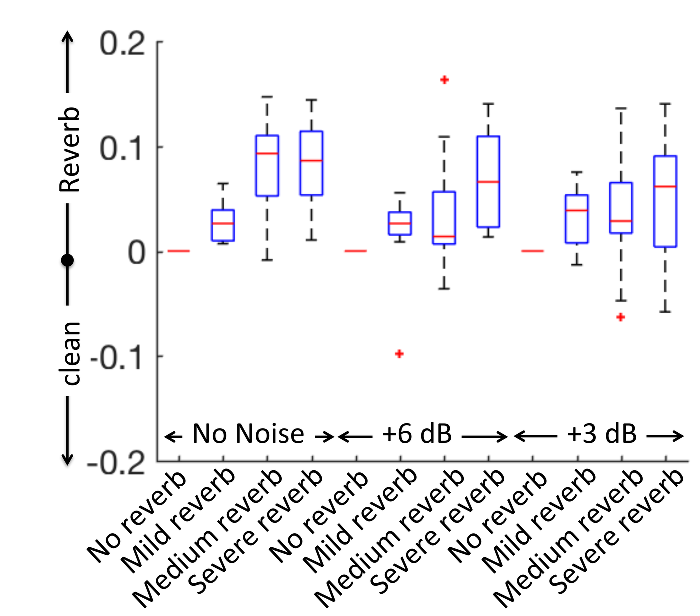
- Optimum decoders are designed to reconstruct the speech envelope under each model (clean/reverb).
- Correlation between reconstructed and presumed model envelope is used as metric as to how faithfully the speech envelope is represented under presumed model.



Analysis & Results

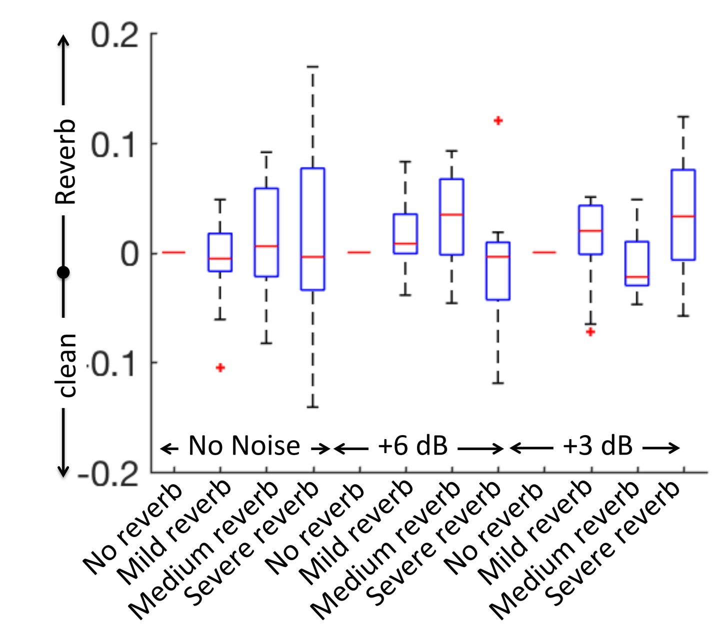
Stimulus reconstruction using Delta band (1-4 Hz) responses

- 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: 1) Ding,N. & J.Z.Simon (2009). Neural representations of complex temporal modulations in the human auditory cortex. *Journal of Neurophysiology.*
2) Wang et al., (2012). Sensitivity to Temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. *Journal of Neurophysiology.*
3) Ding N. and J. Z. Simon, (2013). Adaptive temporal encoding leads to a background insensitive cortical representation of speech. *Journal of Neuroscience.*
Acknowledgements: Funding from NIH R01 DC 014085