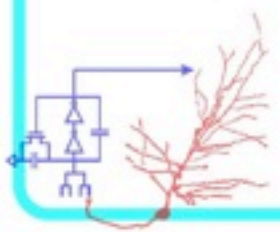
Robust Neural Encoding of Speech in Human Auditory Cortex

Nai Ding, Jonathan Z. Simon

Electrical Engineering / Biology

University of Maryland, College Park



Auditory Processing in Natural Scenes

How is the stable perception of sound generated from degraded acoustics?

Auditory Processing in Natural Scenes

How is the stable perception of sound generated from degraded acoustics?

Magnetoencephalography

(MEG)

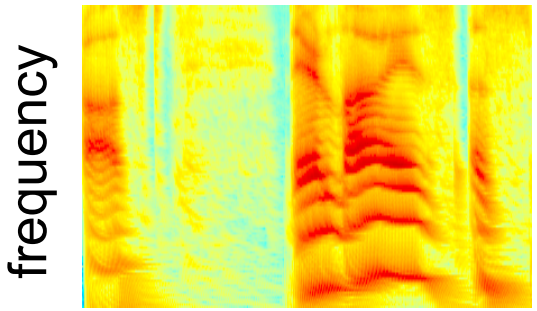
MEG measures spatially synchronized dendritic current.

Outline

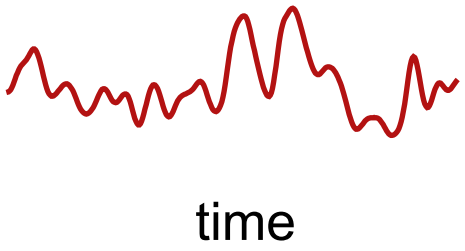
- Cortical Encoding of Speech in MEG
  Representation of Spectro-temporal Features
- Cortical Code despite Energetic Masking
  Speech in Stationary Noise
- Cortical Code despite Informational Masking
  Segregation of Simultaneous Speakers

MEG Response to Speech

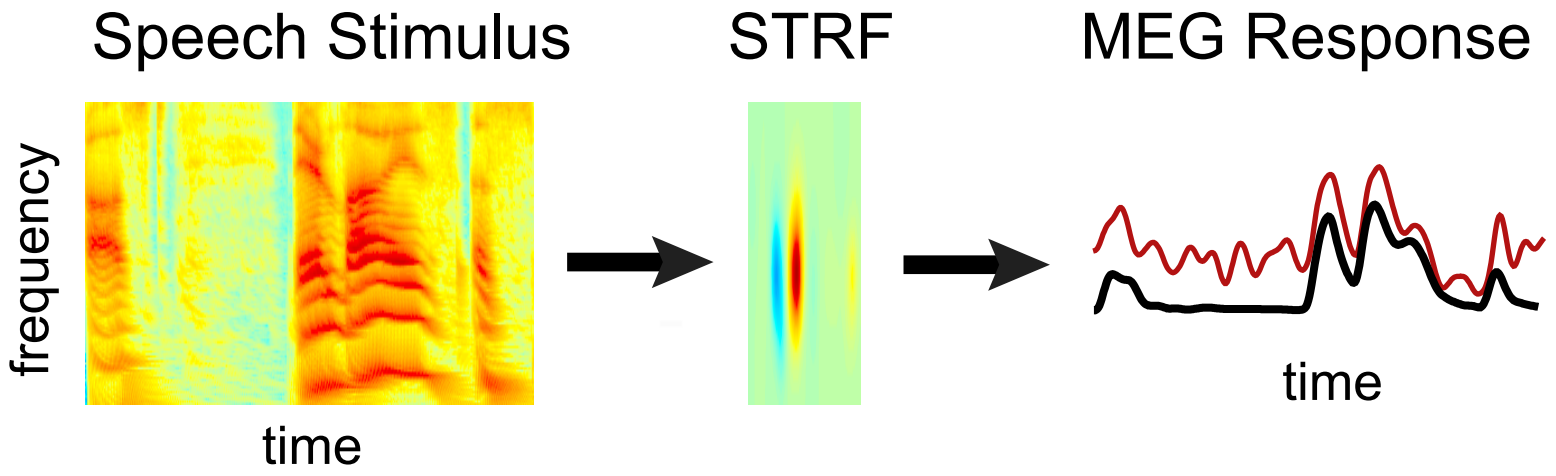
Speech Stimulus

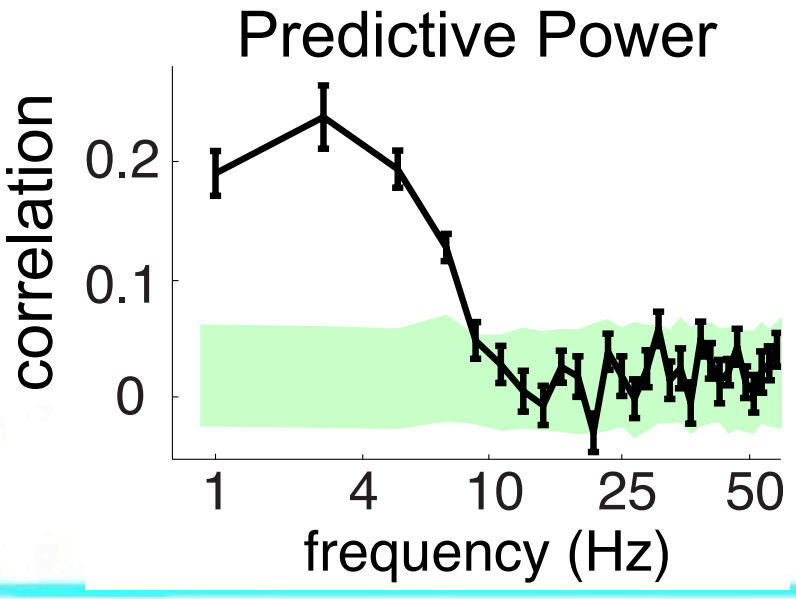


MEG Response



MEG Response to Speech





Large-scale synchronized cortical activity is phase locked to slow temporal modulations of speech.

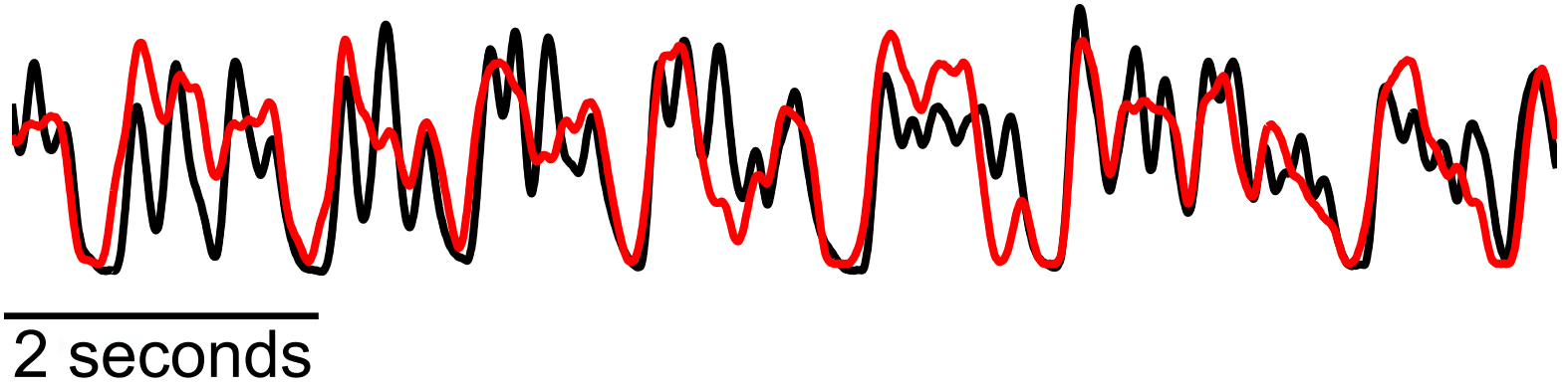
Ding & Simon (in press) *J. Neurophysiol.*

Neural Reconstruction

The temporal envelope of speech can be reconstructed from the MEG response.

stimulus speech envelope

speech envelope reconstructed from MEG response

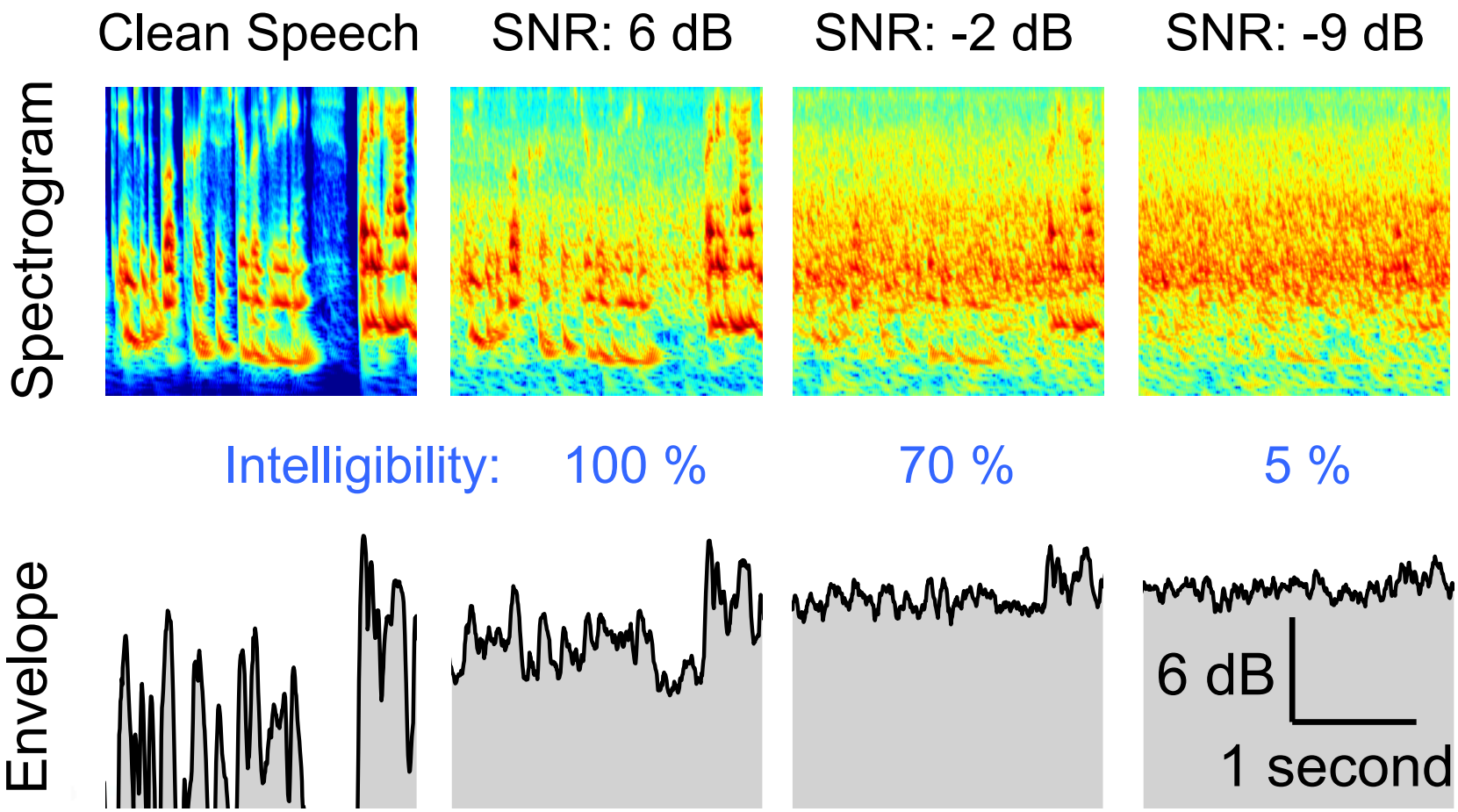


Subject: R1747

Outline

- Cortical Encoding of Speech in MEG
  Representation of Spectro-temporal Features
- **Neural Coding under Energetic Masking**
  Speech in Stationary Noise
- Neural Coding under Informational Masking
  Segregation of Simultaneous Speakers

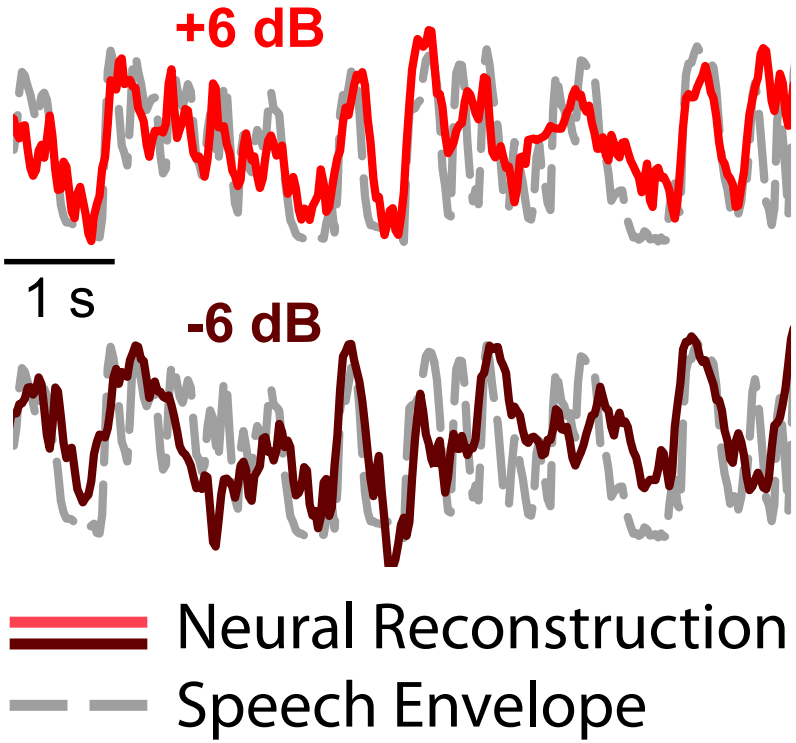
Speech Embedded in Noise

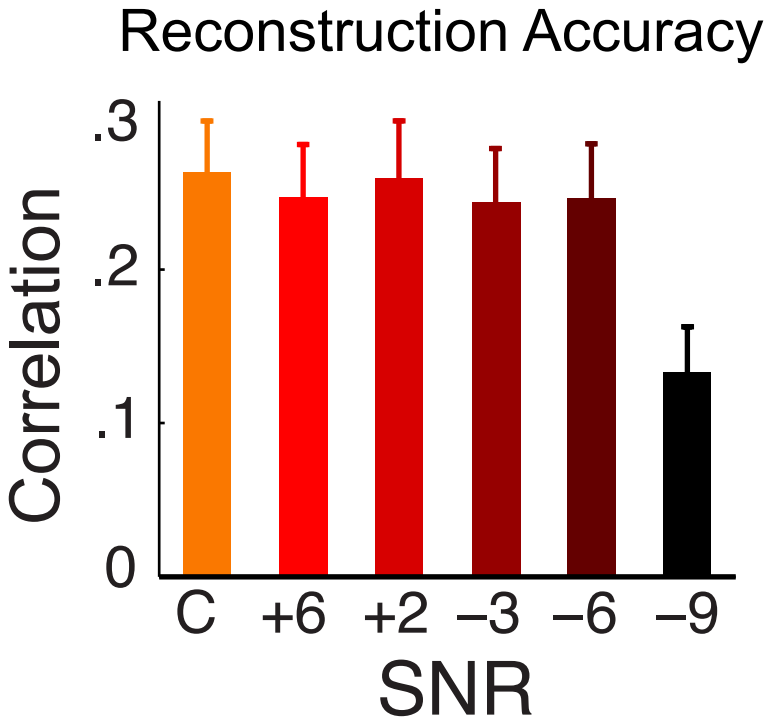


10 participants; 2 minutes of stimulus in each condition

Neural Reconstruction of Speech

The temporal envelope of the underlying speech is reconstructed neurally from cortical response.

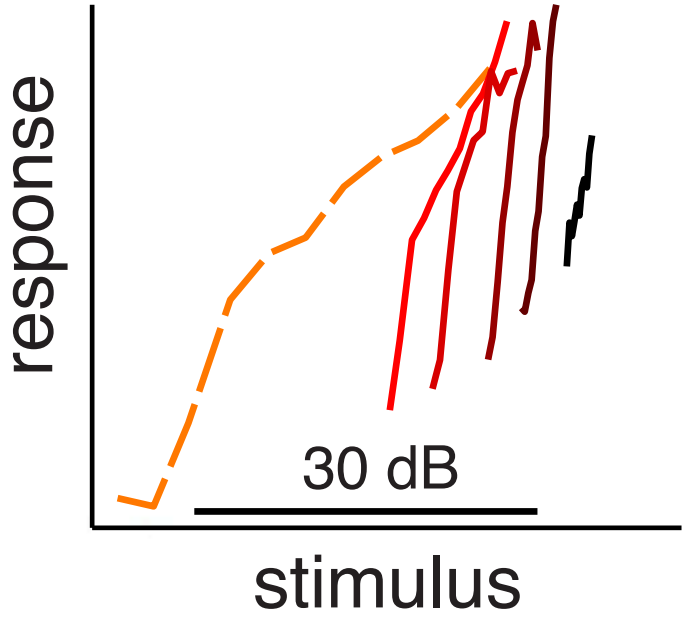




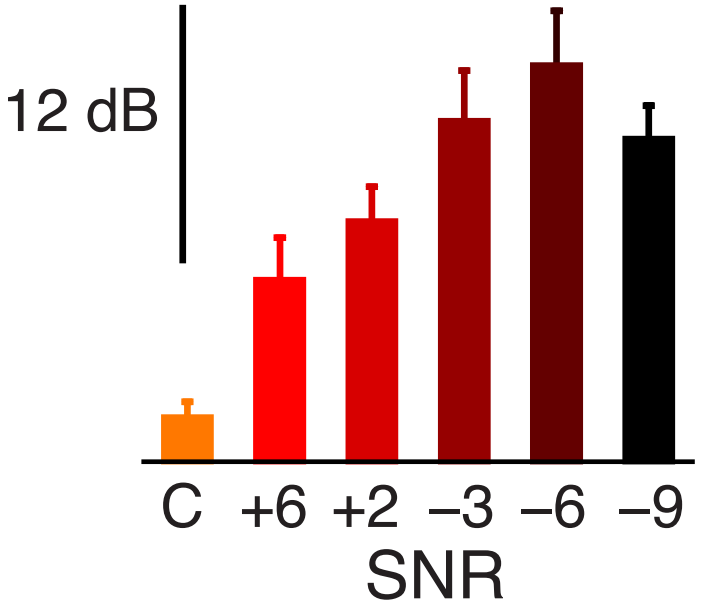
Contrast Gain Control

Neural compensation for noise-induced loss of stimulus contrast

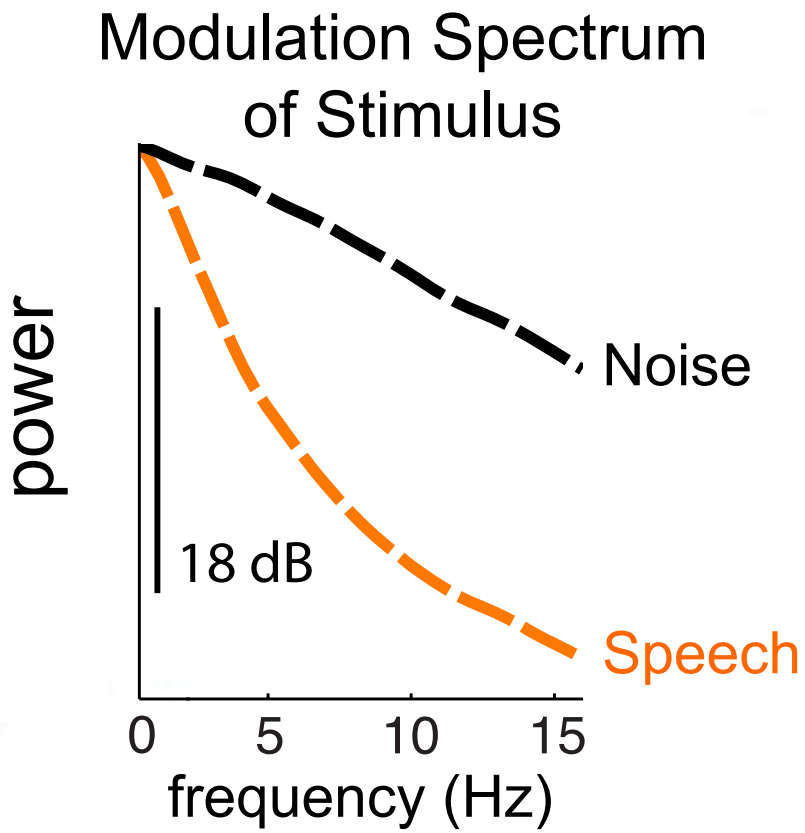
Amplitude-Intensity Function



Amplitude Growth Rate



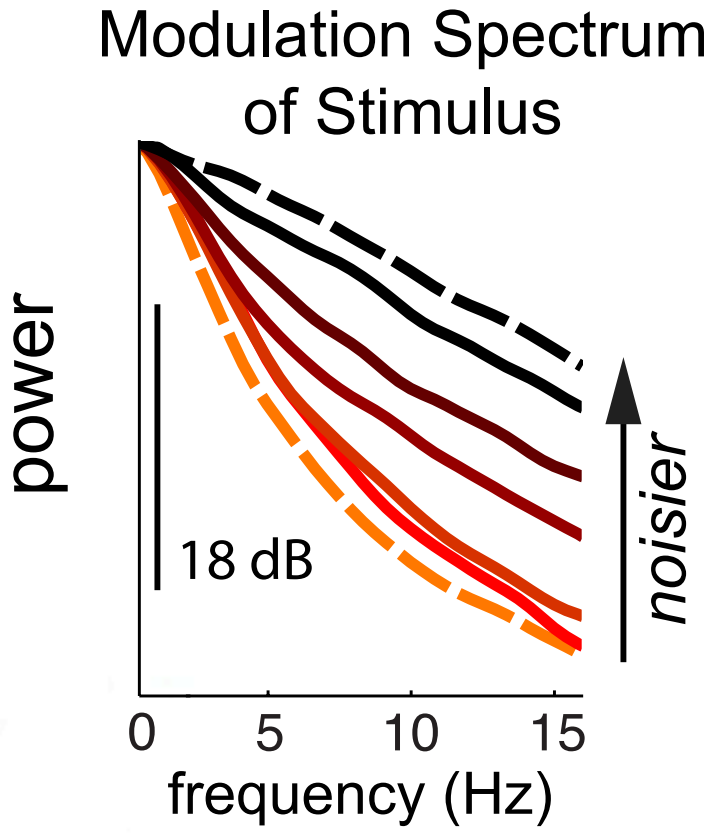
Adaptive Encoding of Modulations

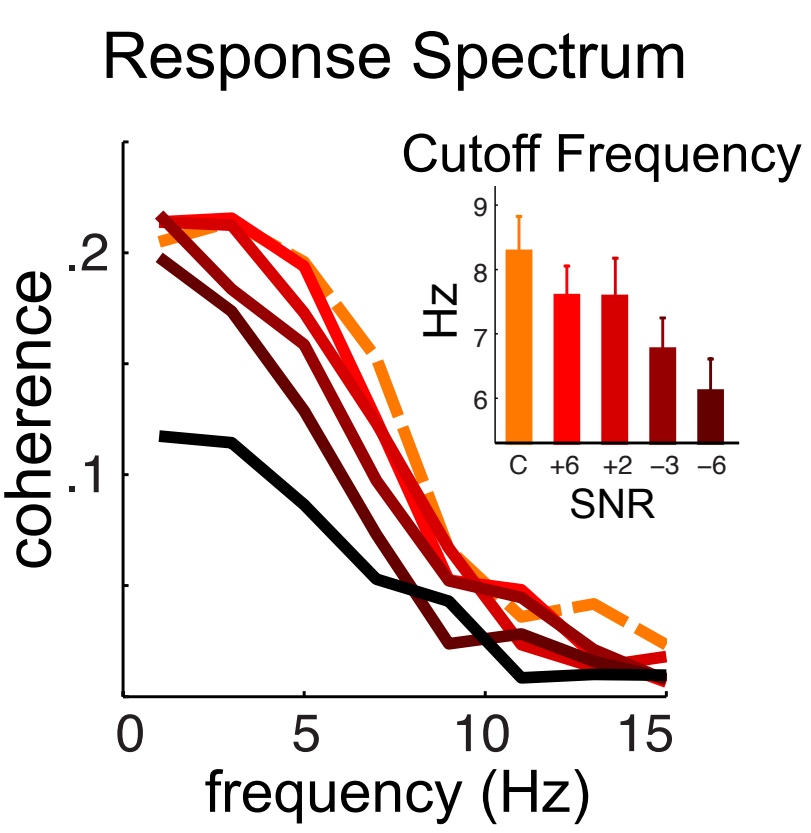


Noise contains more energy at higher modulation rate, and therefore interfere with speech more at high modulation rates.

Adaptive Encoding of Modulations

Neural sensitivity profile shifts away from the modulation rates heavily corrupted by noise.





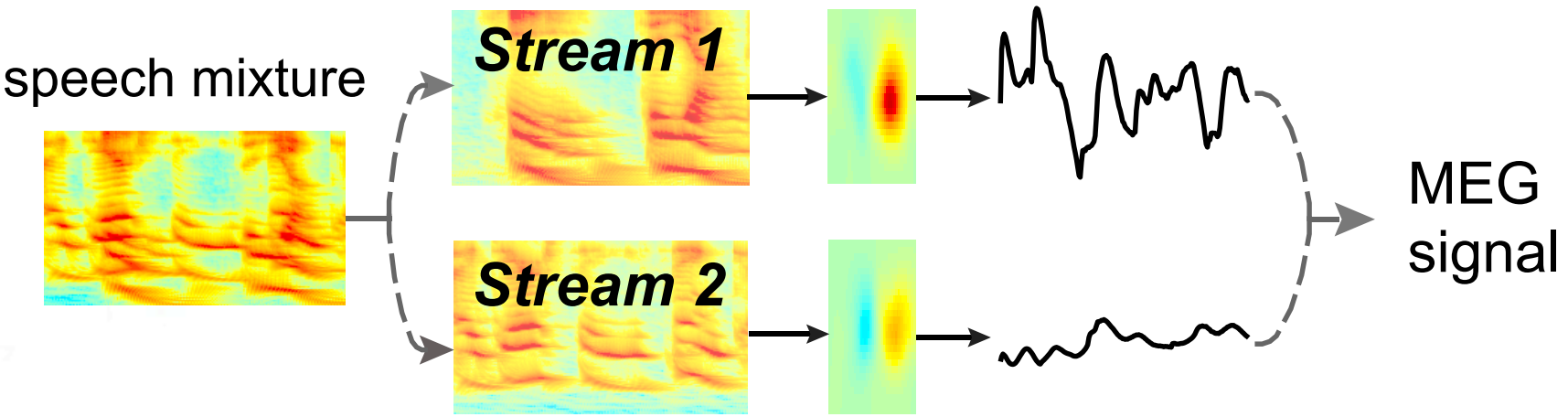
Outline

- Cortical Encoding of Speech in MEG
  Representation of Spectro-temporal Features
- Neural Coding under Energetic Masking
  Speech in Stationary Noise
- **Neural Coding under Informational Masking**
  Segregation of Simultaneous Speakers

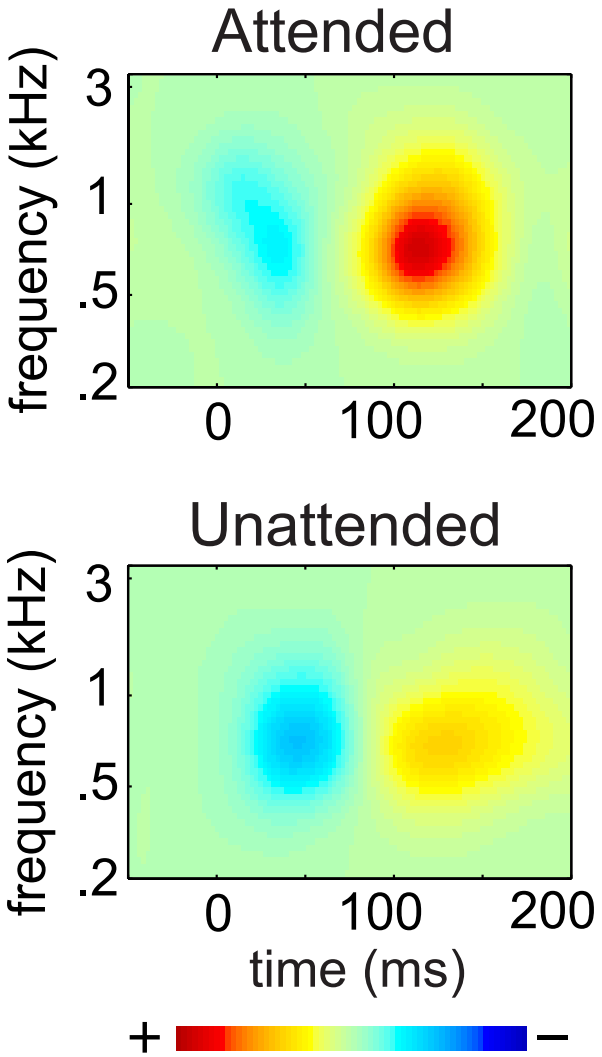
Diotic Speech Segregation

Two speakers, one male and one female, were mixed and presented diotically. The subjects were instructed to focus on one or the other speaker.

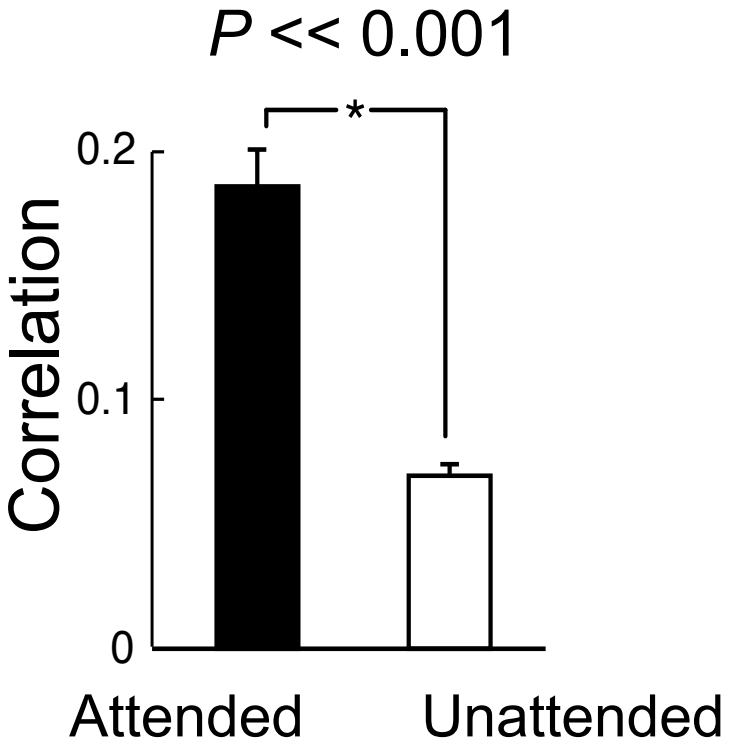
The MEG response is modeled using two STRFs, one for each speaker.



Neural Unmixing of Concurrent Speakers



Neurally decoded envelope is more correlated with the attended speaker in >90% of single trials.



Summary

- 1. Neural processing adapts to noise.
- 2. Simultaneous speakers can be neurally segregated and processed differently.
- 3. Cortical encoding is precise yet dynamic: modulated by both stimulus acoustics (bottom-up) and attention (top-down), and leading to a robust encoding of speech in natural scenes.

Acknowledgement

- We thank Stephen David, David Poeppel, Mary Howard, Shihab Shamma, and Monita Chatterjee for discussions!

- SfN poster: 172.11/KK6 (Sunday, 10 -11)

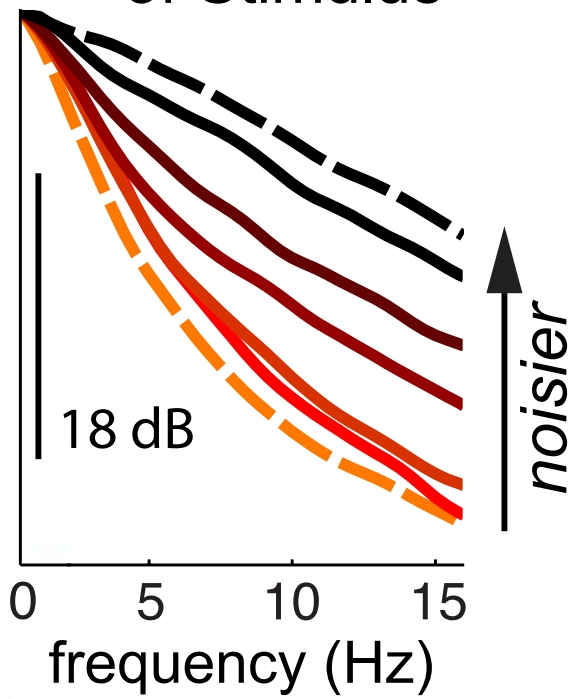
- Contact:
- gahding@umd.edu Nai Ding
- jzsimon@umd.edu Jonathan Z. Simon

Thank you!

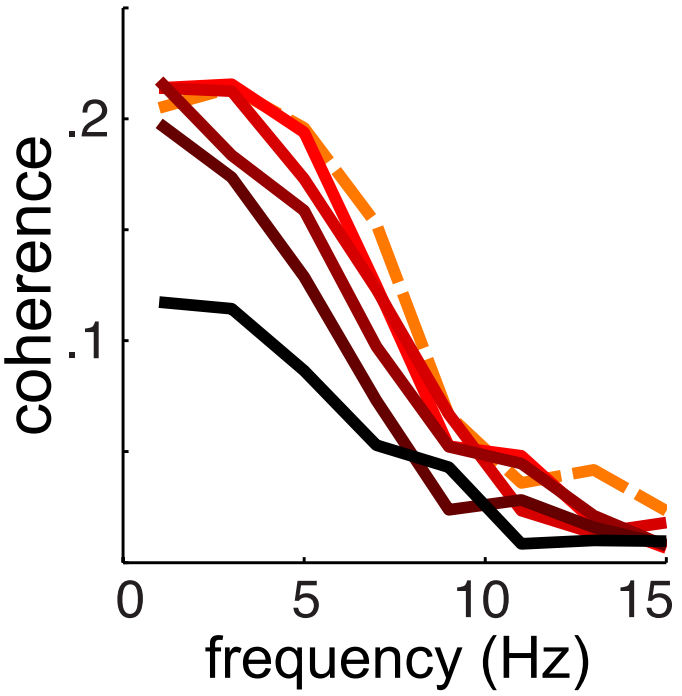
Adaptive Encoding of Modulations

Neural sensitivity profile shifts away from the modulation rates heavily corrupted by noise.

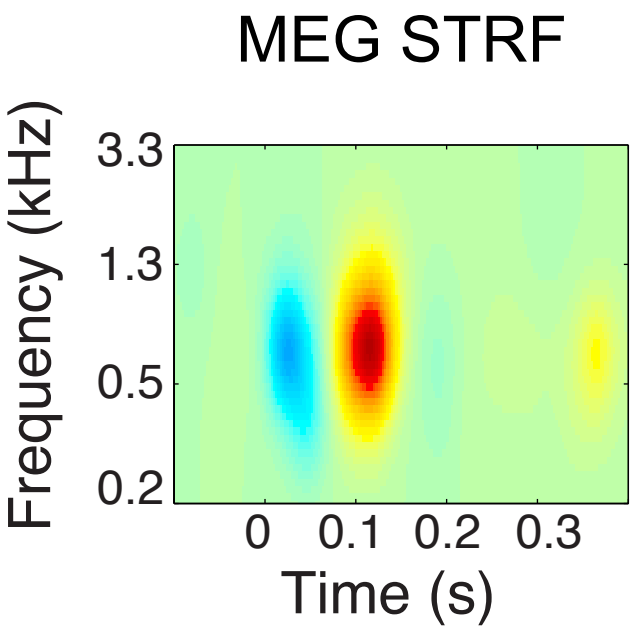
Modulation Spectrum of Stimulus

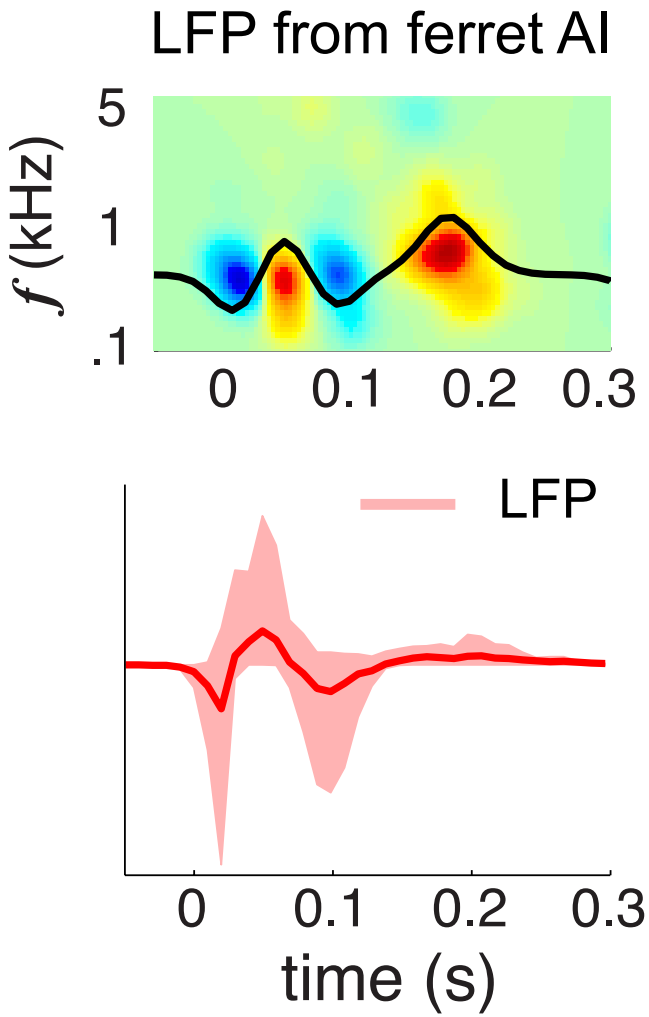


Response Spectrum



STRF from MEG and LFP





(in collaboration with Stephen David and Shihab Shamma)