

Cortical Encoding of Auditory Objects in the Cocktail Party Problem

Jonathan Z. Simon
University of Maryland

Introduction

- Auditory Objects
- Magnetoencephalography (MEG)
- Decoding Neural Signals/Encoding Stimuli
- Cortical Representations of Auditory Objects I
- Cortical Representations of Auditory Objects II

Auditory Objects

- What is an Auditory Object?
 - Perceptual/Psychophysical Construct
 - Some Commonalities with Visual Objects
 - *I know it when I see it* vs. Formal Definition

Auditory Object Definition

- E. g., Griffiths & Warren
 - an object corresponds with something in the sensory world
 - information related to the object is separate from information related to the rest of the sensory world
 - abstracted so that object information can be generalized among particular sensory experiences in any one sensory domain

Auditory Objects at the Cocktail Party



Alex Katz,
The Cocktail Party

Auditory Objects at the Cocktail Party



Alex Katz,
The Cocktail Party

Auditory Objects at the Cocktail Party



Alex Katz,
The Cocktail Party

Auditory Objects at the Cocktail Party



Alex Katz,
The Cocktail Party

Auditory Objects at the Cocktail Party



Alex Katz,
The Cocktail Party

Auditory Objects at the Cocktail Party



Alex Katz,
The Cocktail Party

Auditory Objects at the Cocktail Party



Alex Katz,
The Cocktail Party

Auditory Objects at the Cocktail Party



Alex Katz,
The Cocktail Party

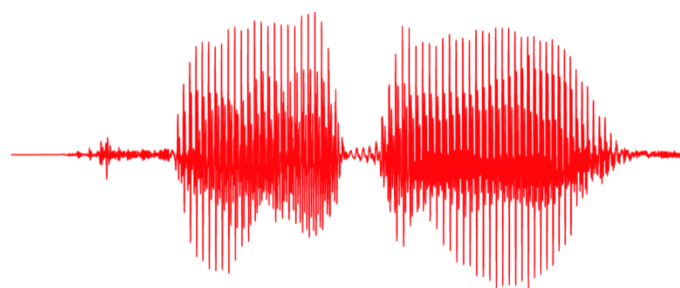
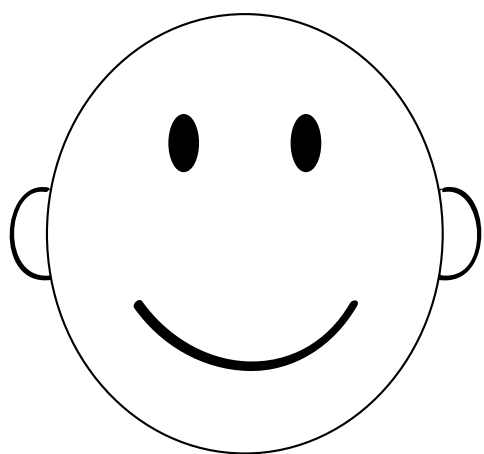
Auditory Objects at the Cocktail Party

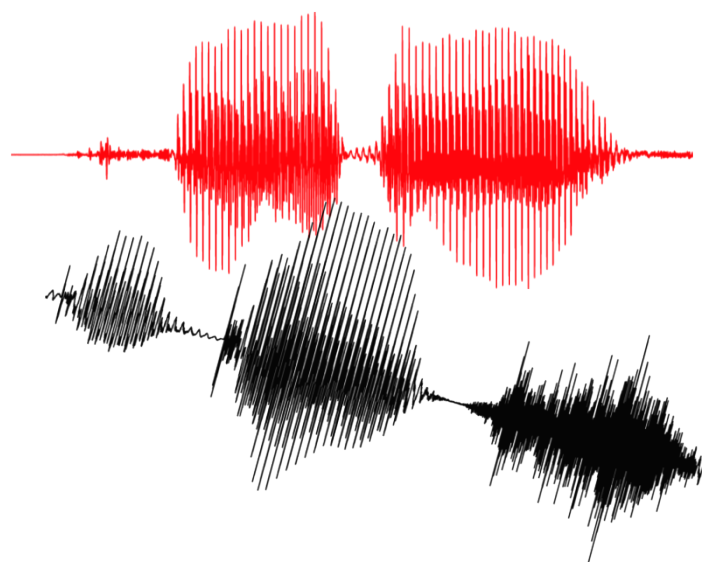
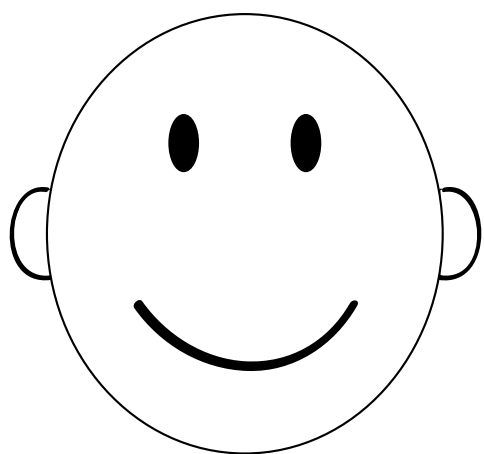


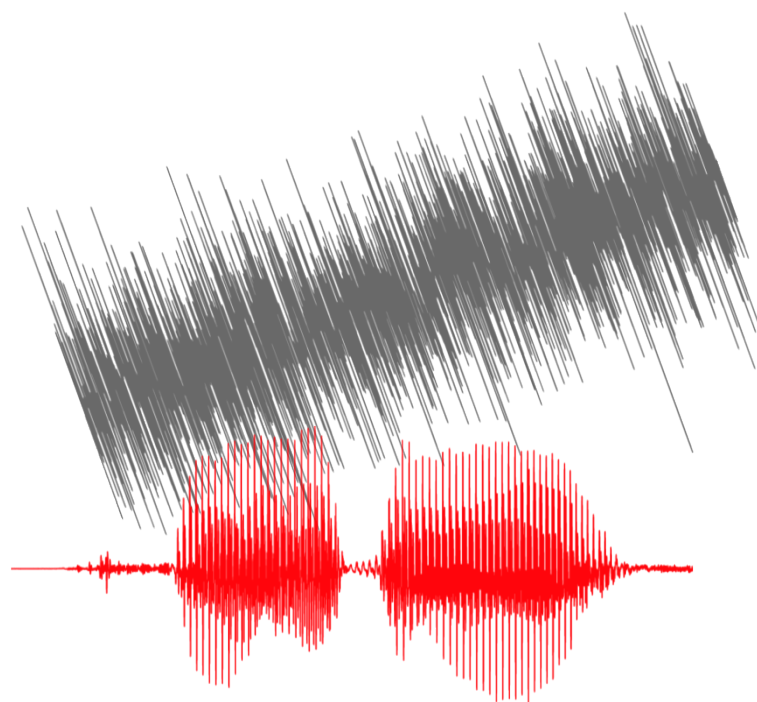
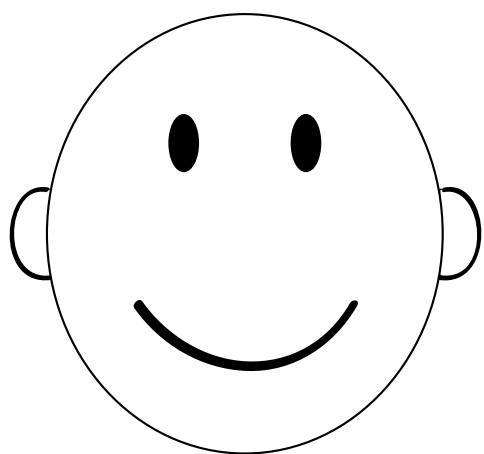
Alex Katz,
The Cocktail Party

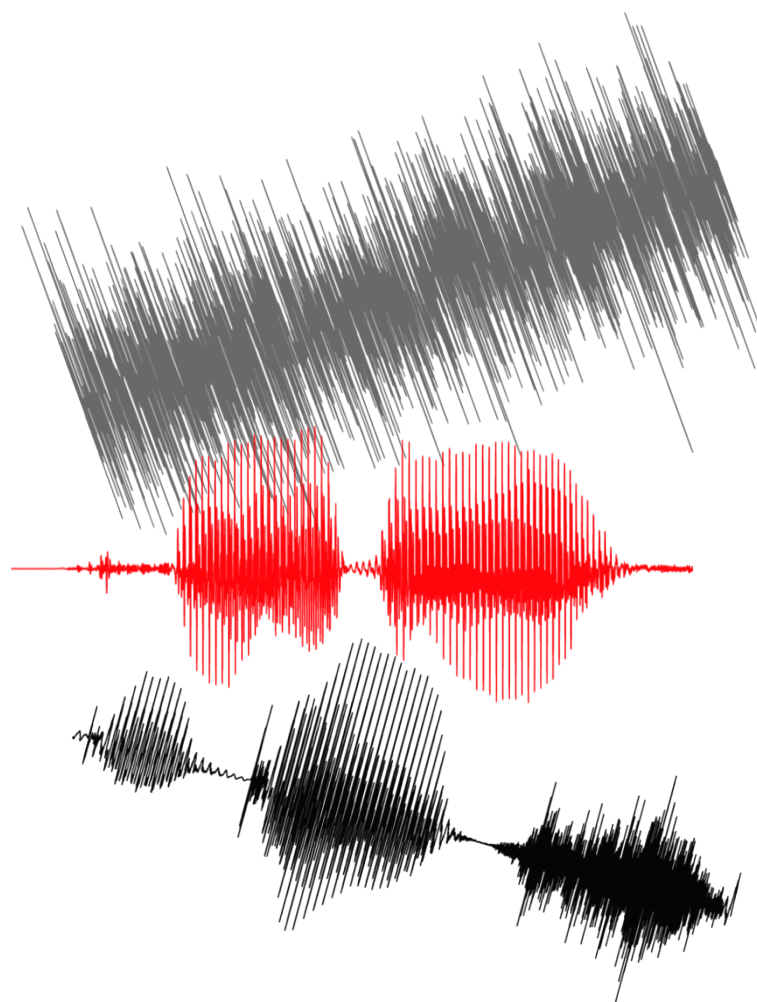
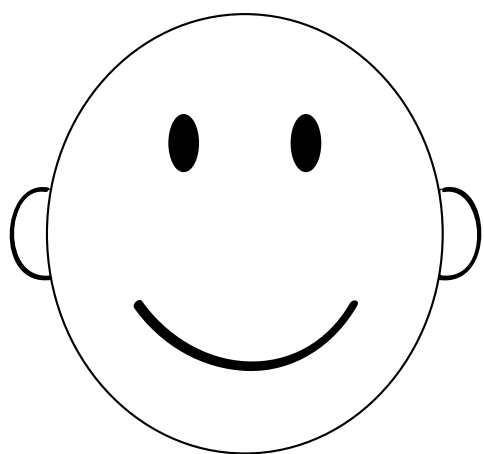
Auditory Objects at the Cocktail Party

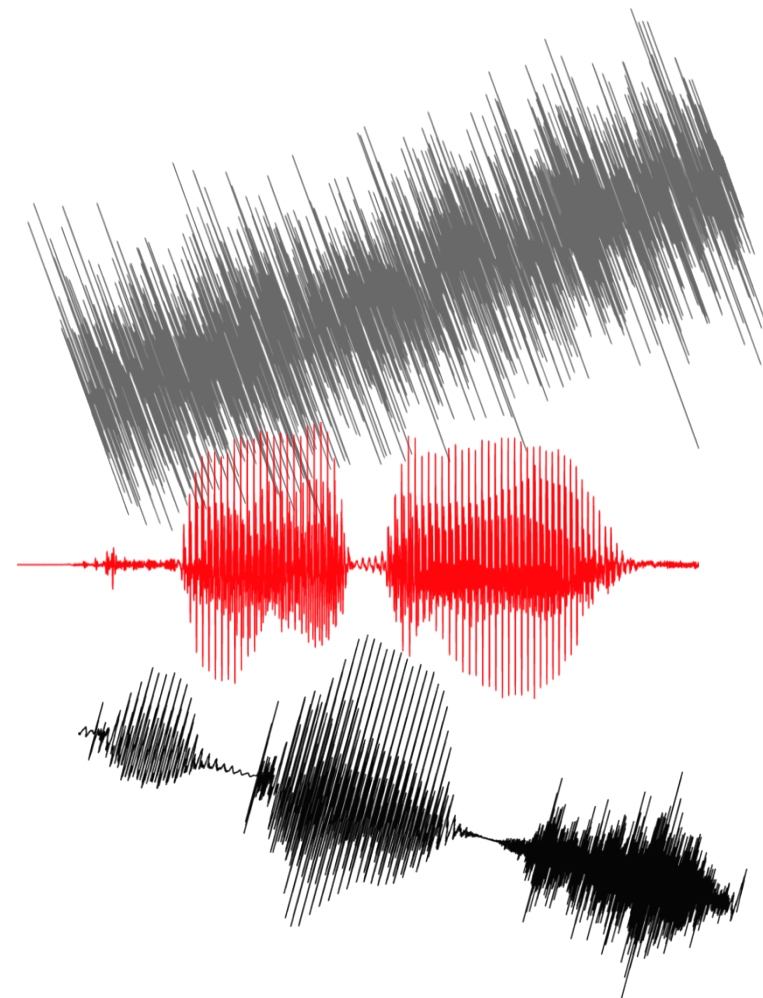












Magnetoencephalography (MEG)

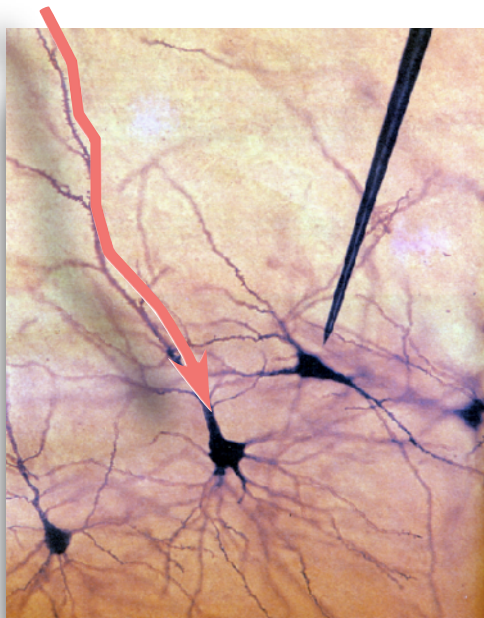
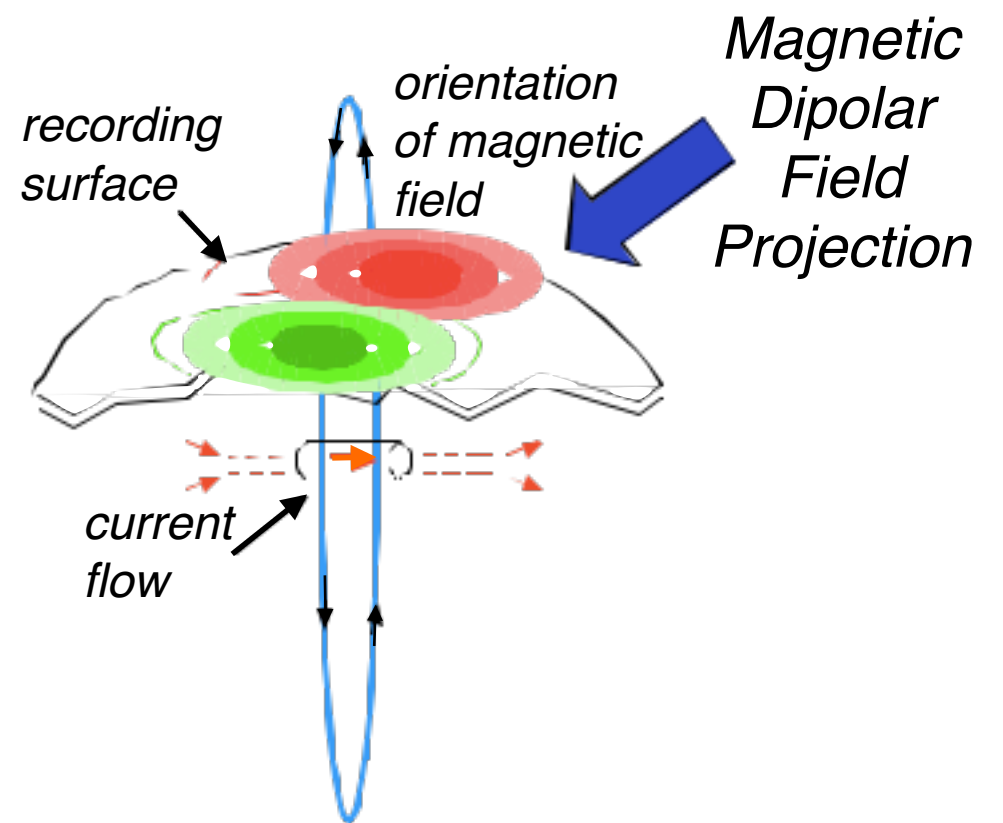
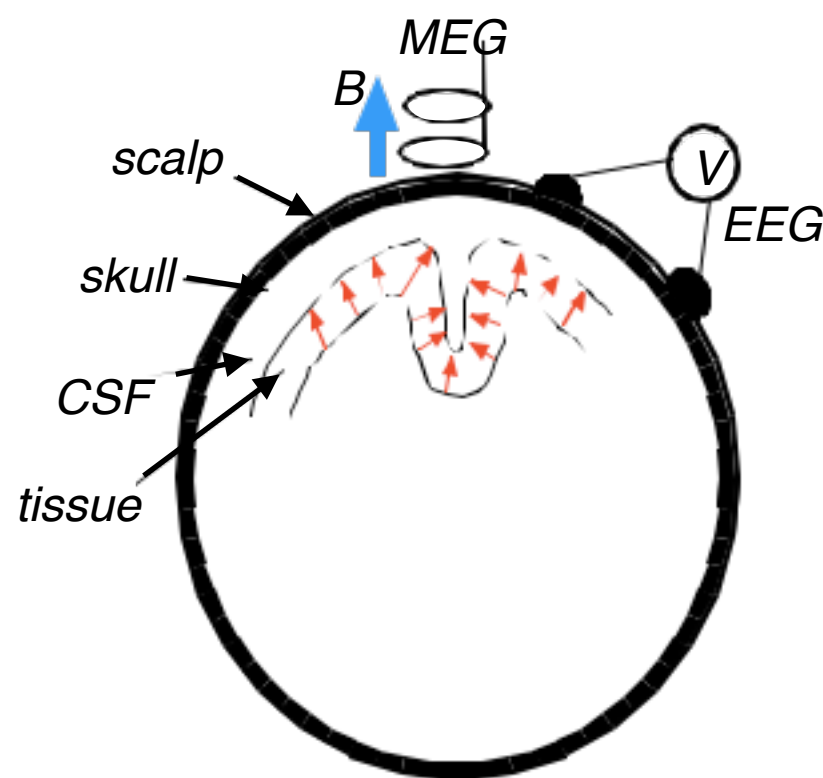
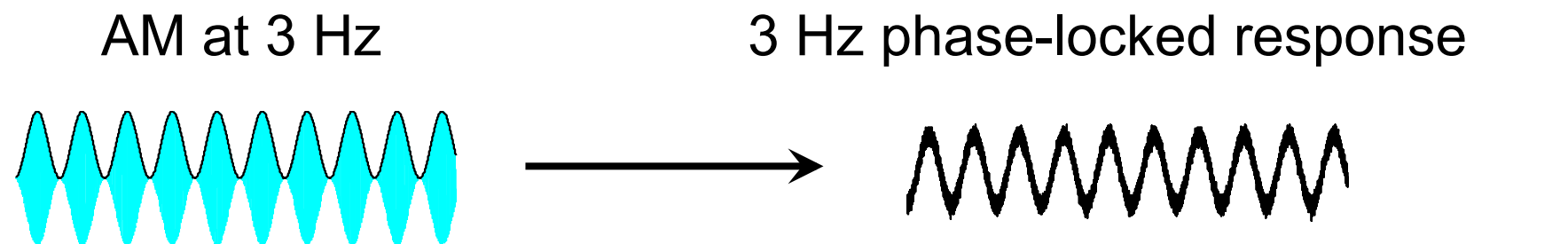


Photo by Fritz Goro



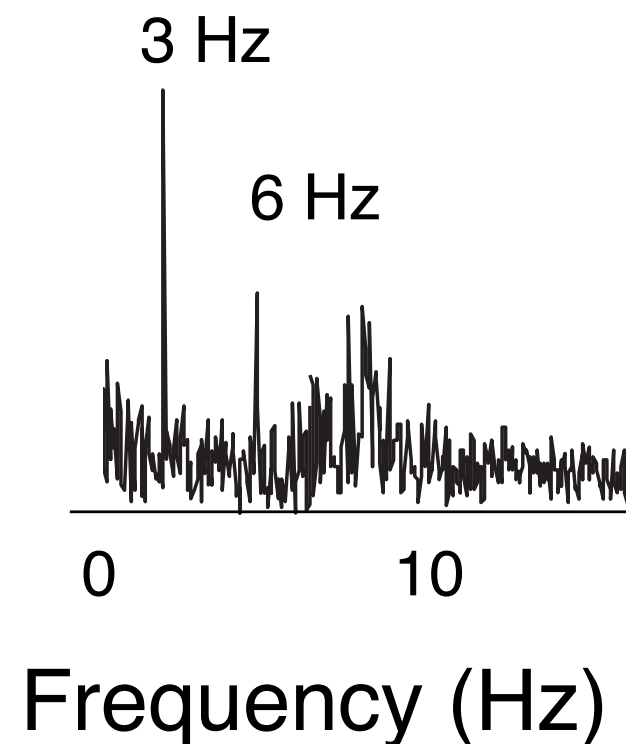
- Direct electrophysiological measurement
 - not hemodynamic
 - real-time
- No unique solution for distributed source
- Measures spatially synchronized cortical activity
- Fine temporal resolution (~ 1 ms)
- Moderate spatial resolution (~ 1 cm)

MEG Phase Locking to Slow Temporal Modulations

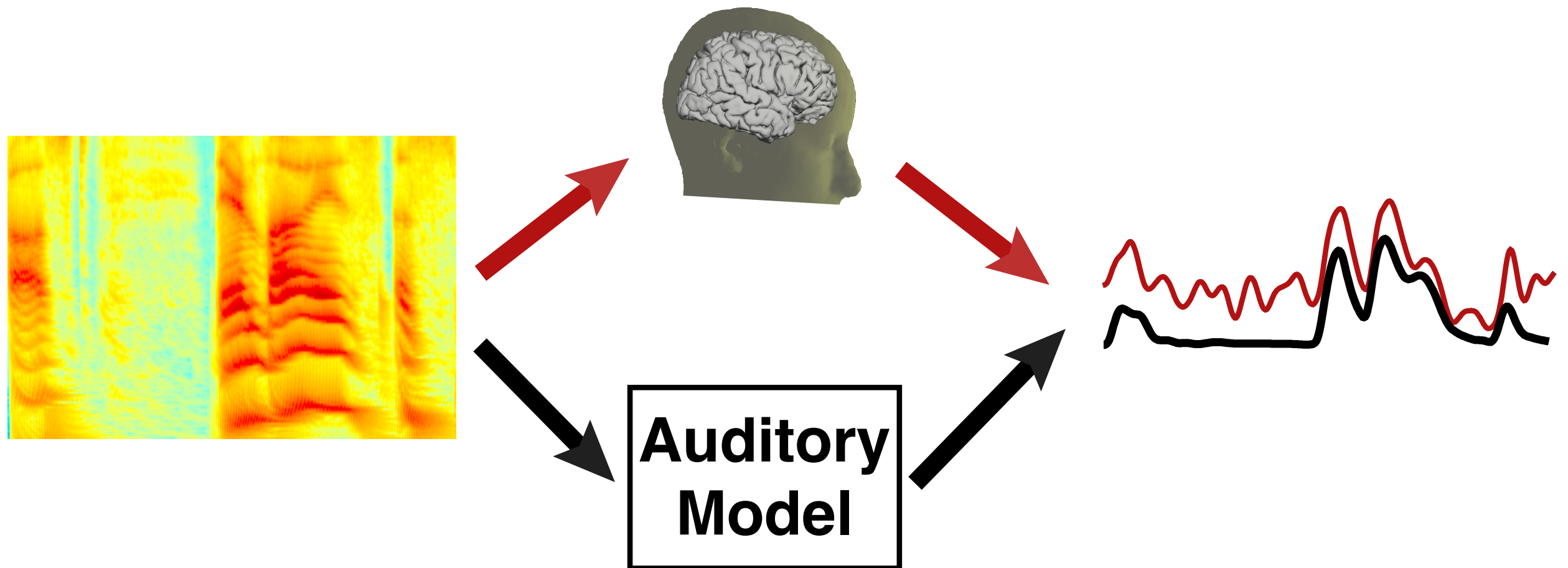


MEG activity is precisely phase-locked to temporal modulations of sound

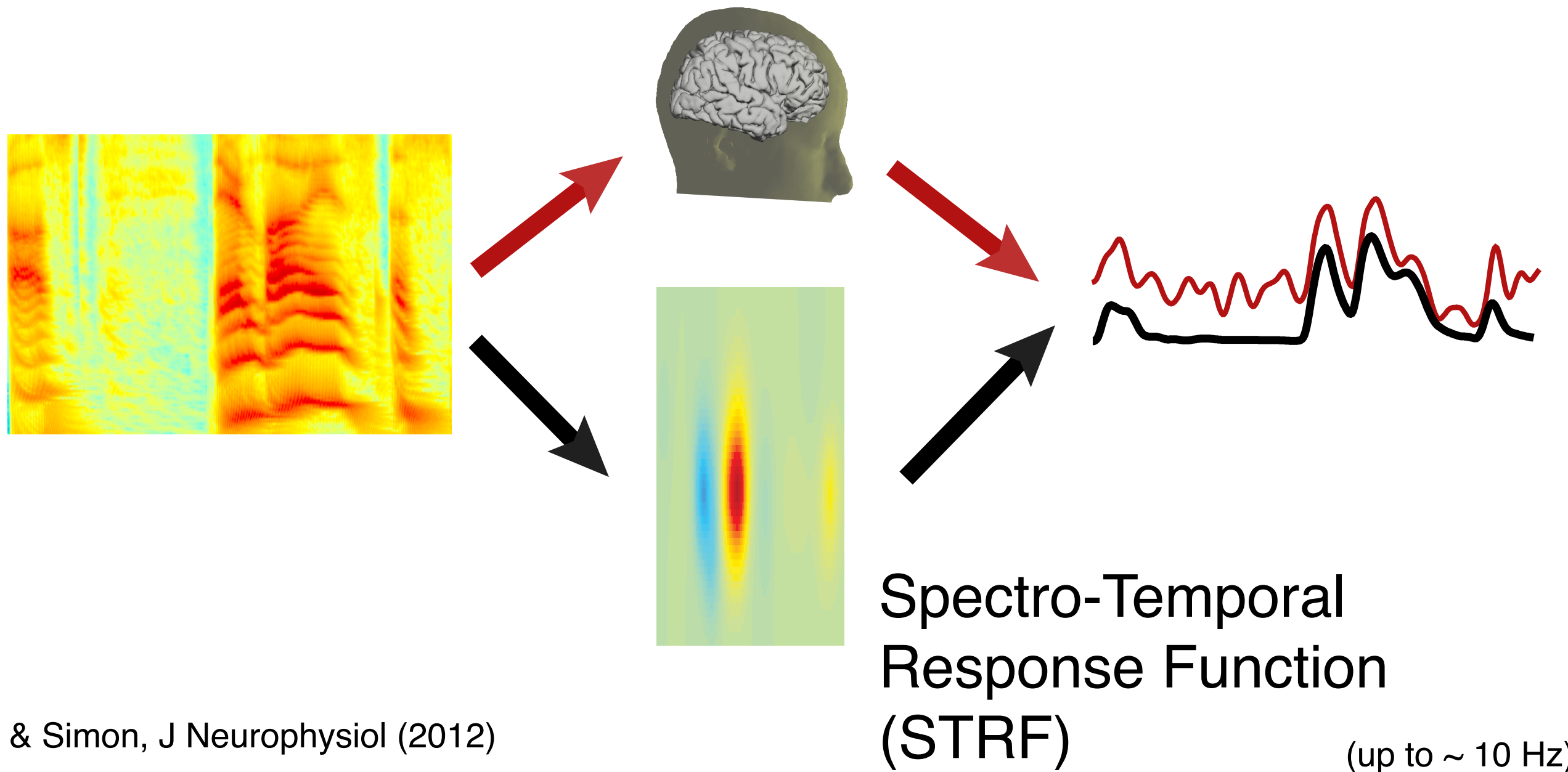
response spectrum (*subject R0747*)



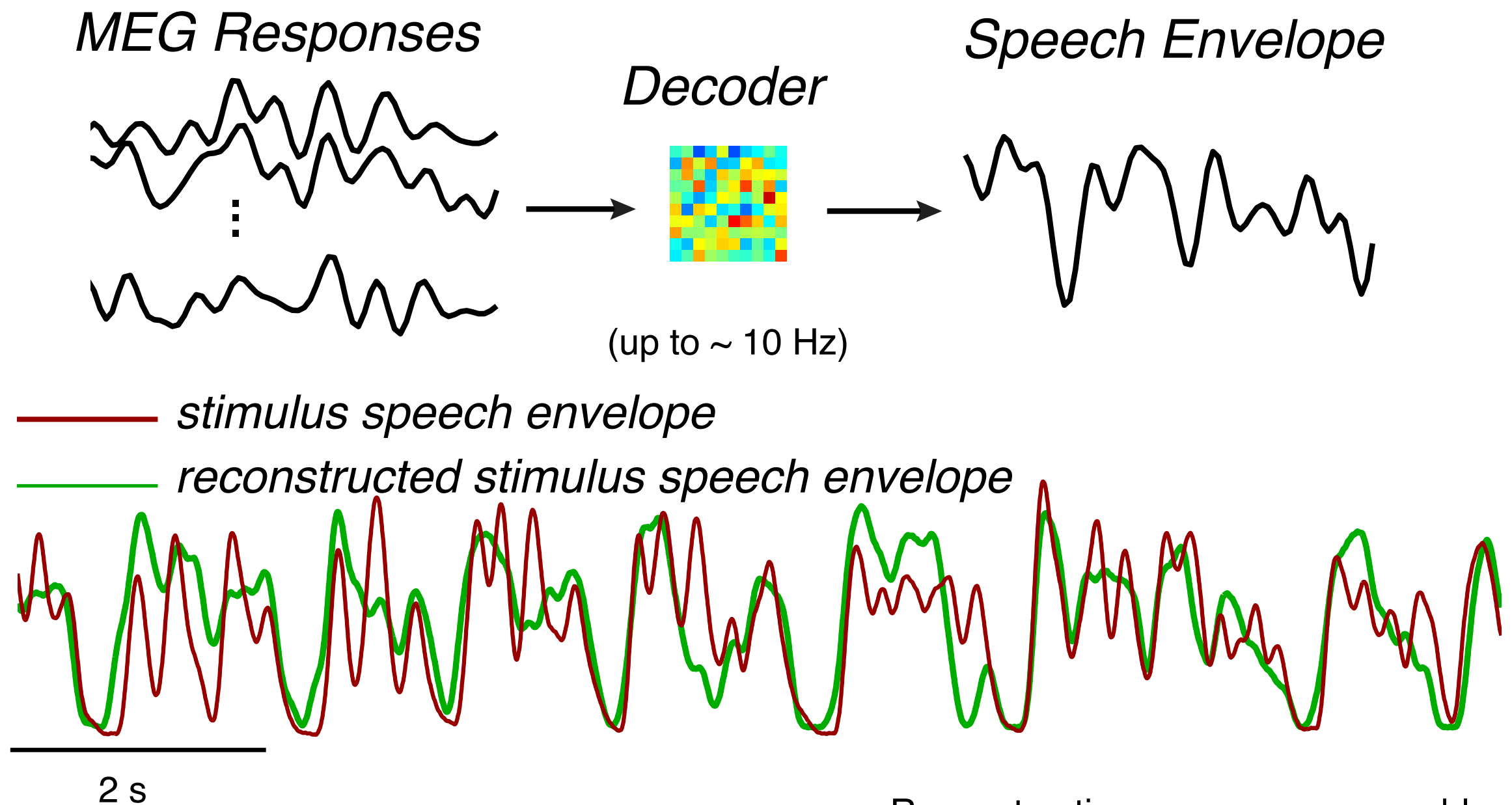
Modeling MEG Response to Speech Modulations



Modeling MEG Response using STRF model



Neural Reconstruction of Speech Envelope



Reconstruction accuracy comparable
to single unit recordings & ECoG

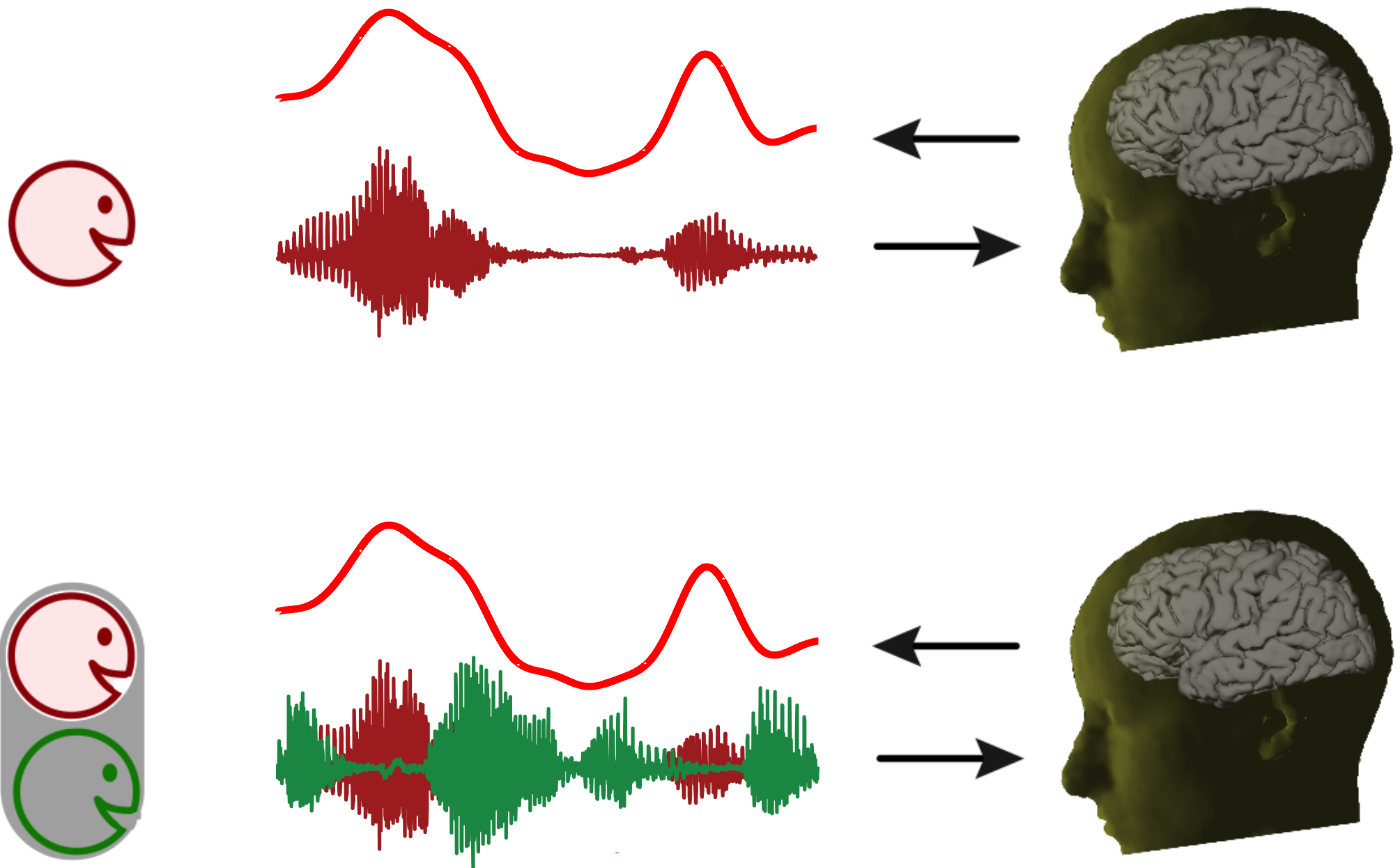
Speech Stream as an Auditory Object

- a speech stream corresponds with something in the sensory world
- information related to a speech stream is separate from information related to the rest of the sensory world, e.g. other speech streams or noise
- a speech stream is abstracted: it is generalized among different sensory experiences, e.g. different sound mixtures

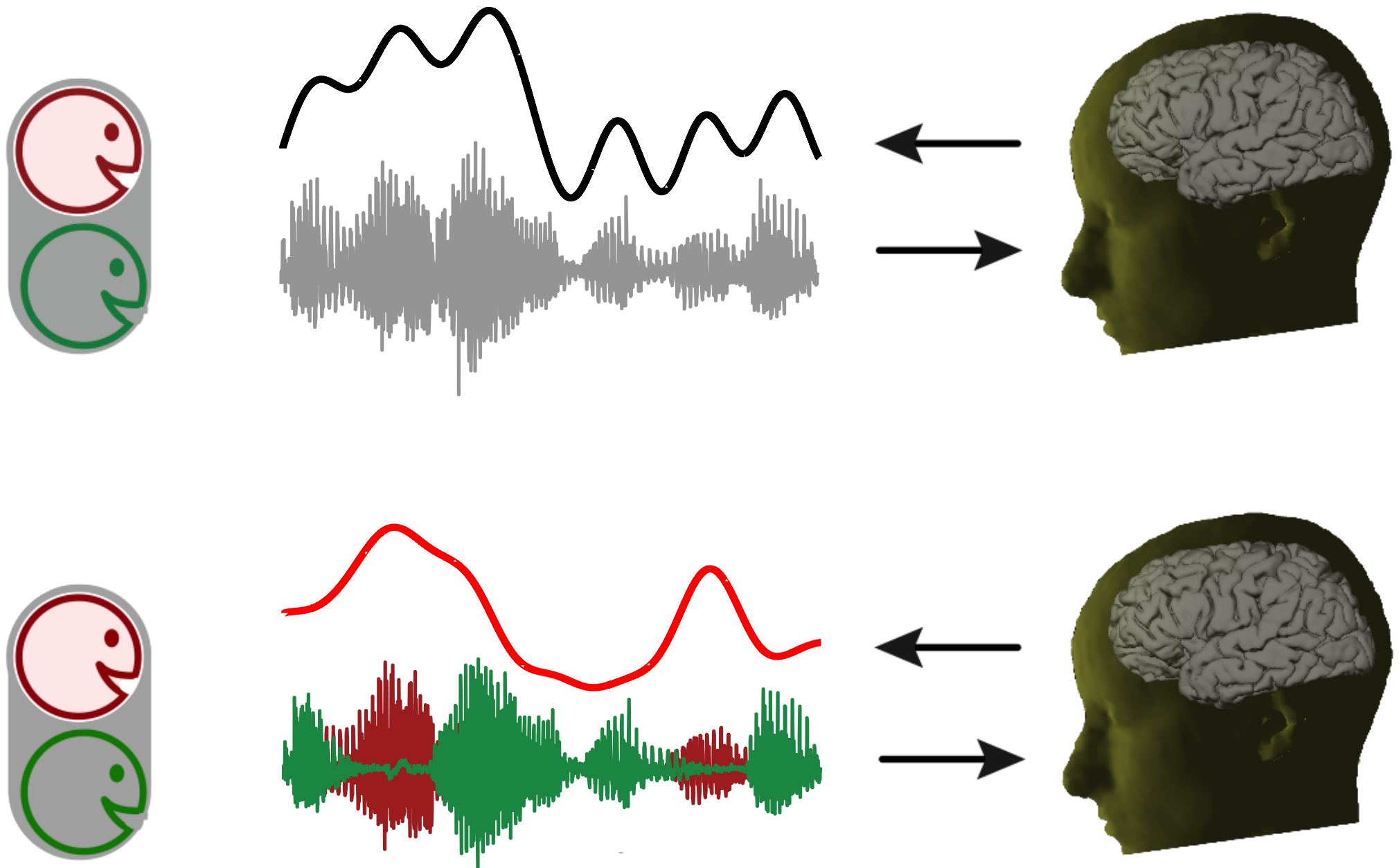
Neural Representation of an Auditory Object

- neural representation is of auditory object, something in sensory world
- when auditory object is with other sounds, the neural representation is of the auditory object, not the entire acoustic scene
- neural representation remains invariant under broad changes in acoustic representation of auditory object

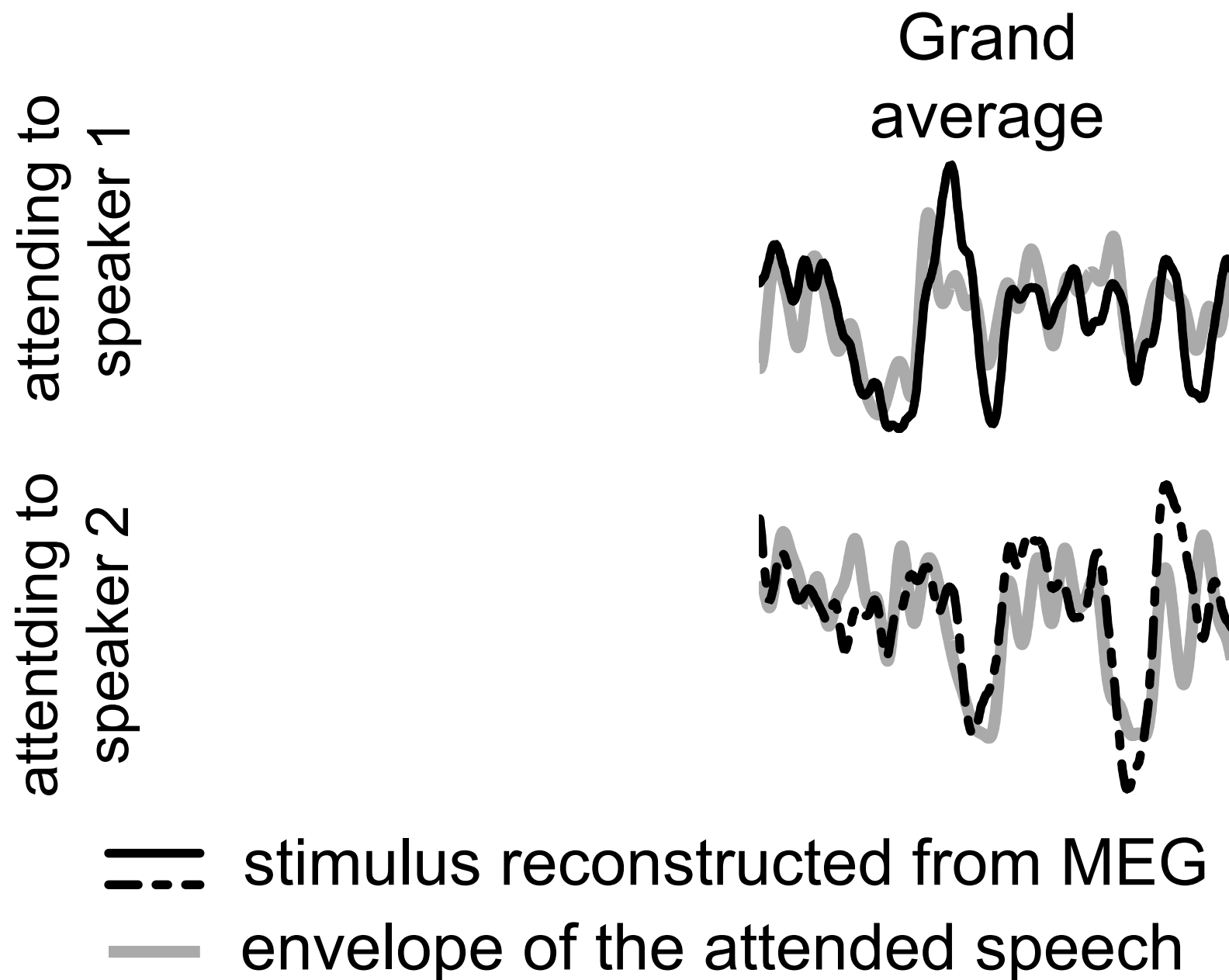
Selective Neural Encoding



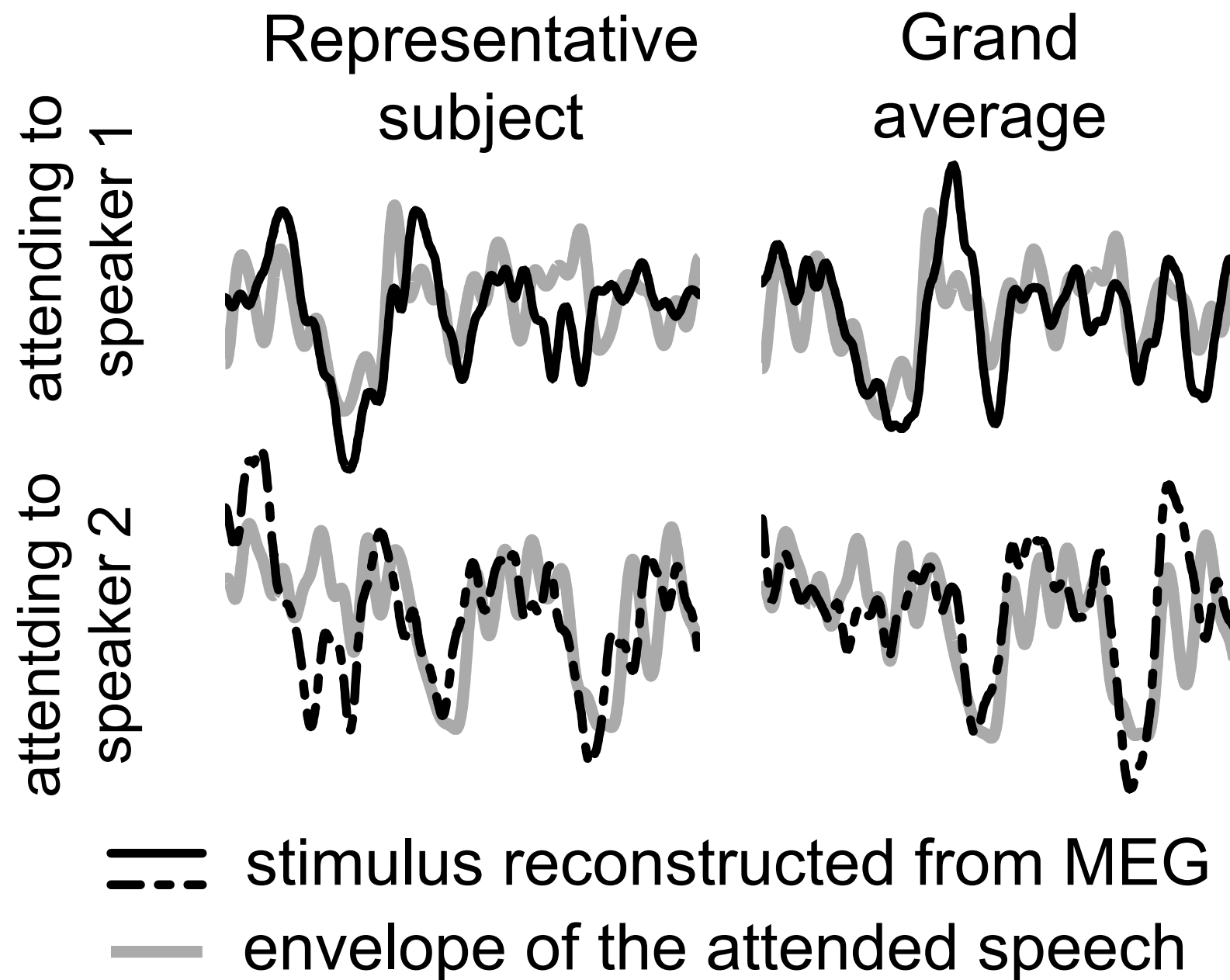
Unselective vs. Selective Neural Encoding



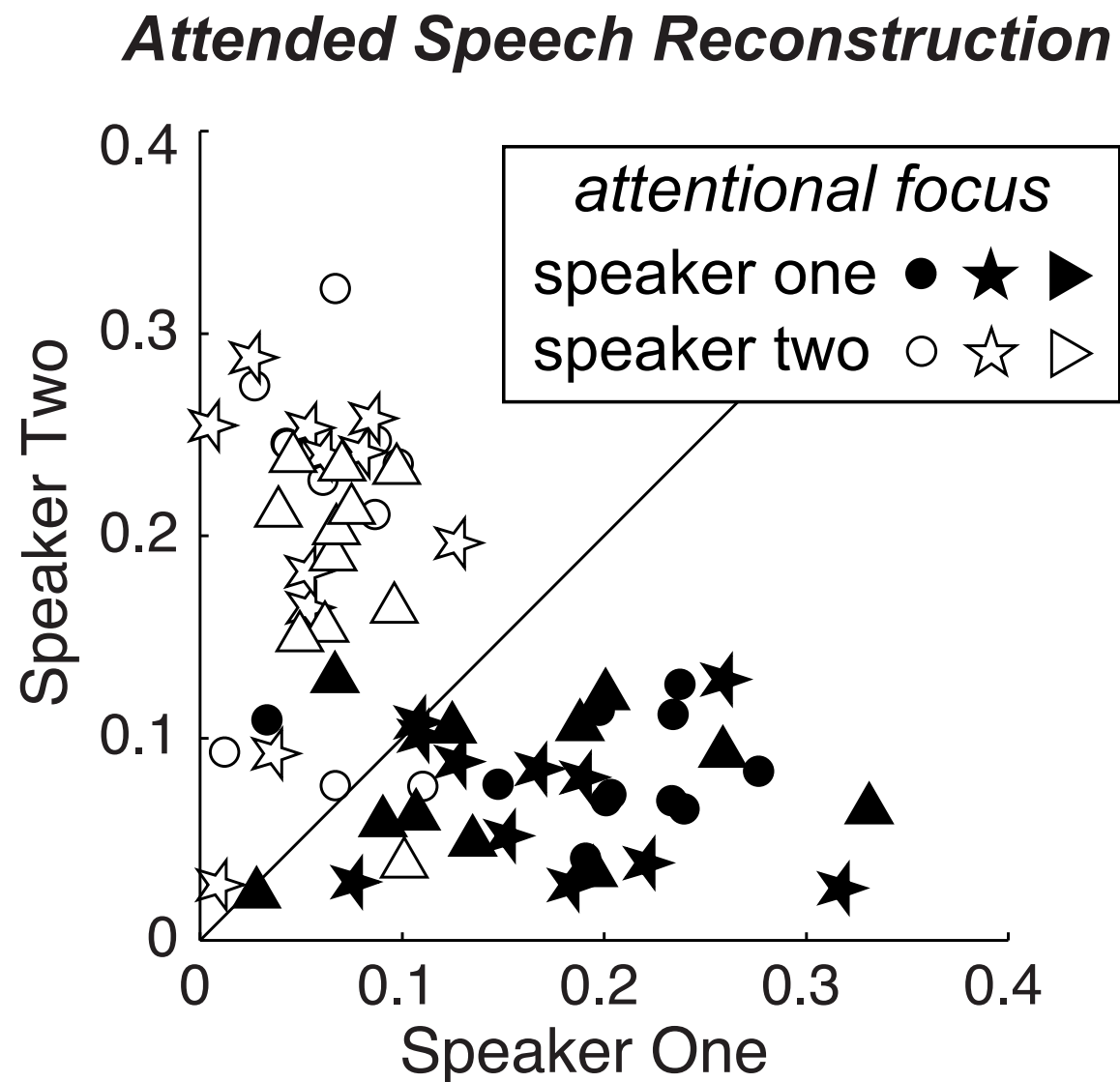
Auditory Object-Specific Representation



Auditory Object-Specific Representation

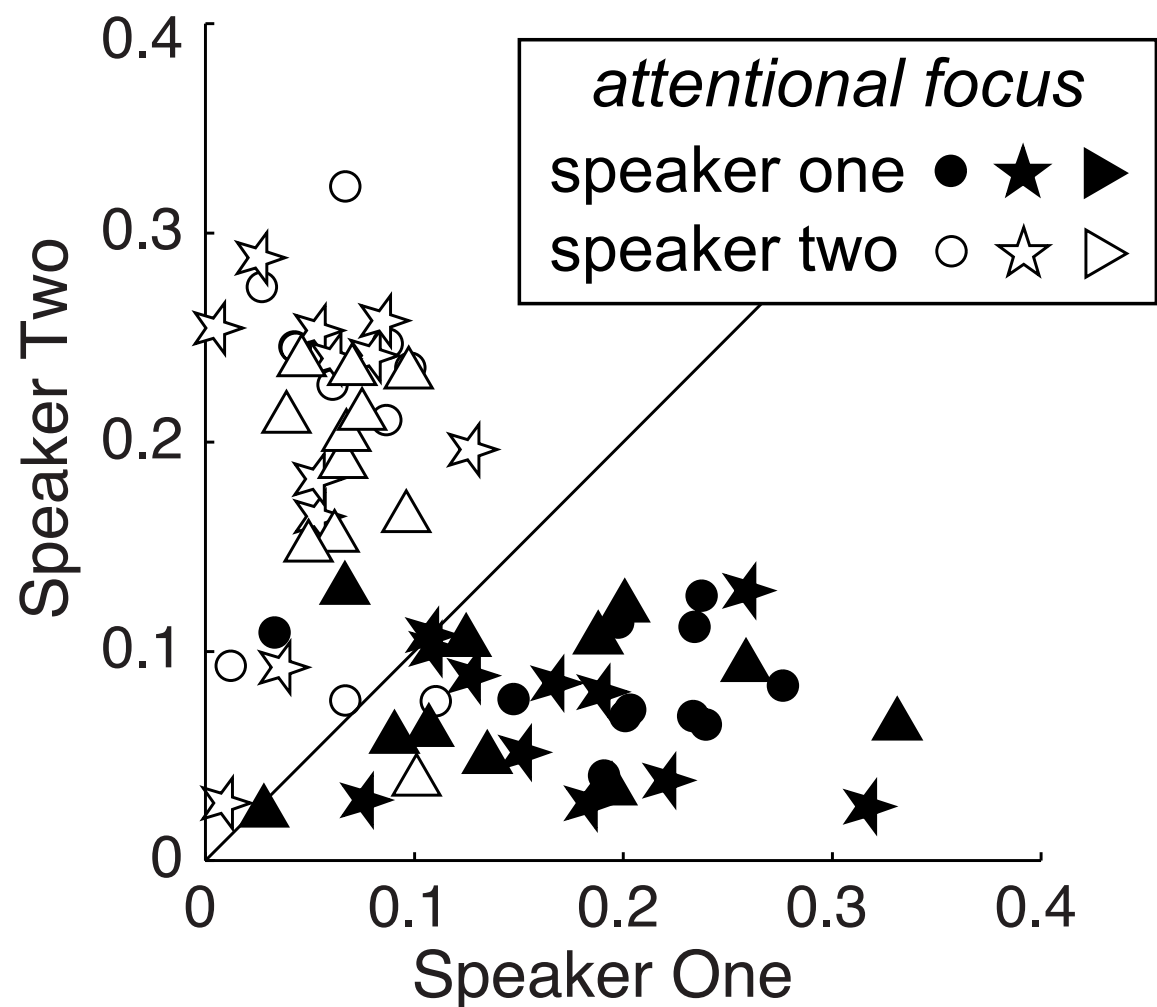


Single Trial Speech Reconstruction

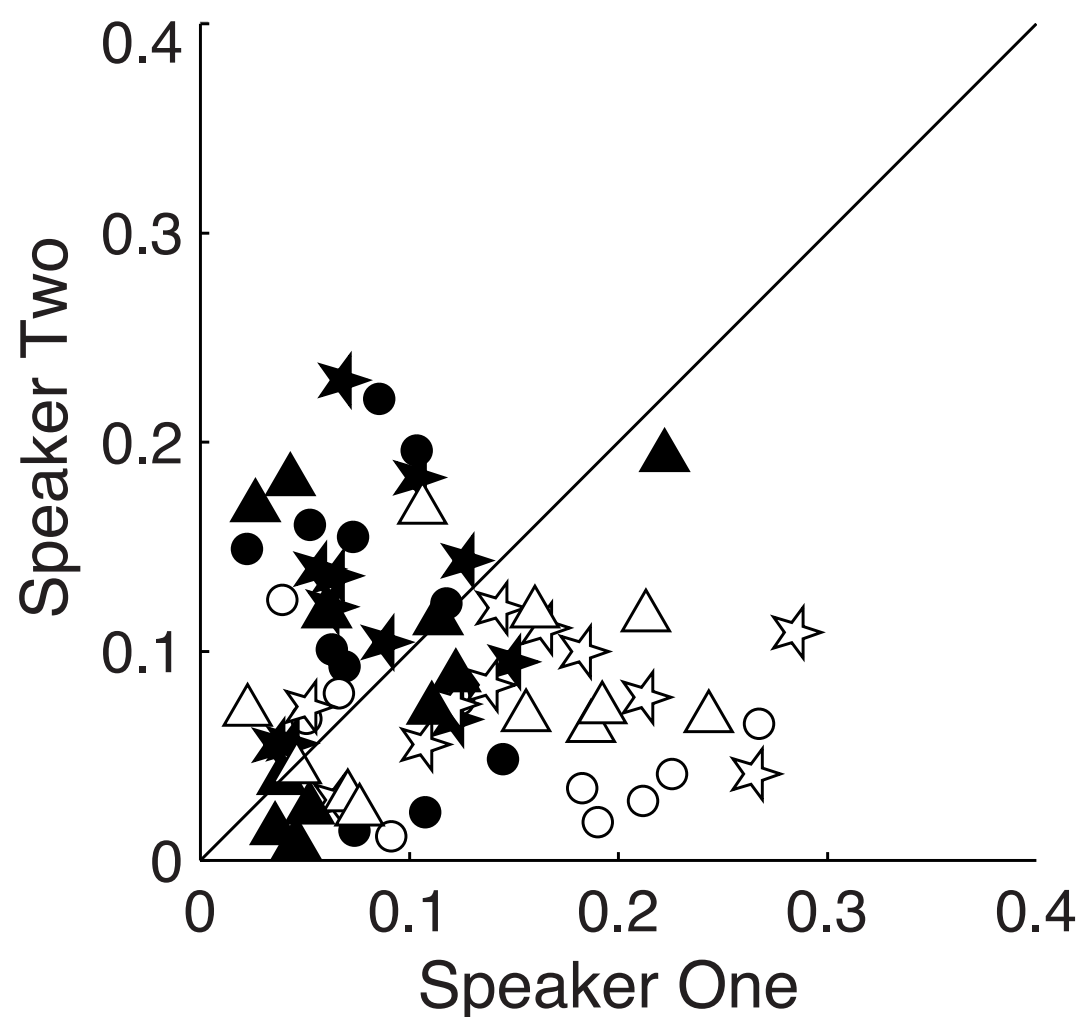


Single Trial Speech Reconstruction

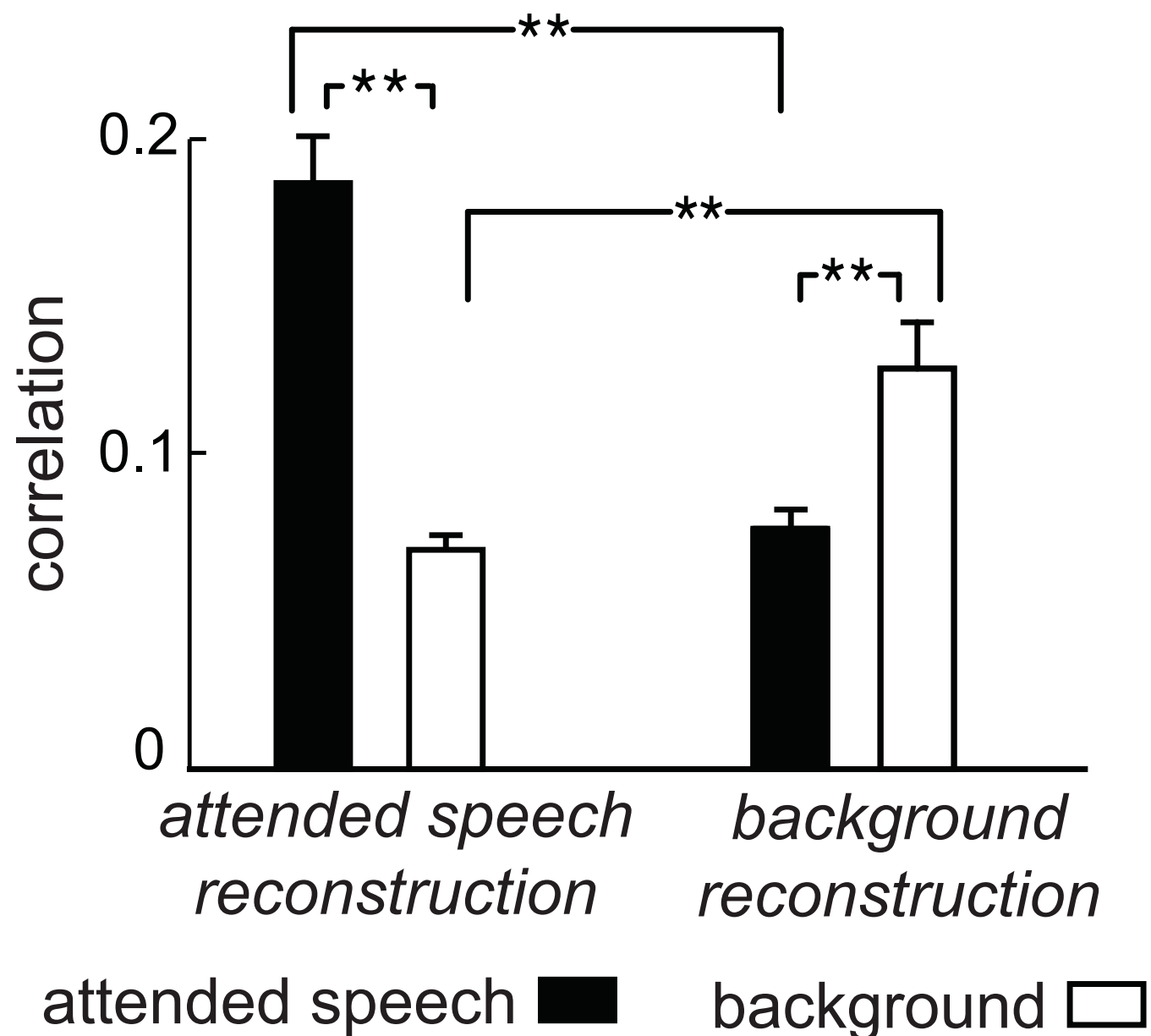
Attended Speech Reconstruction



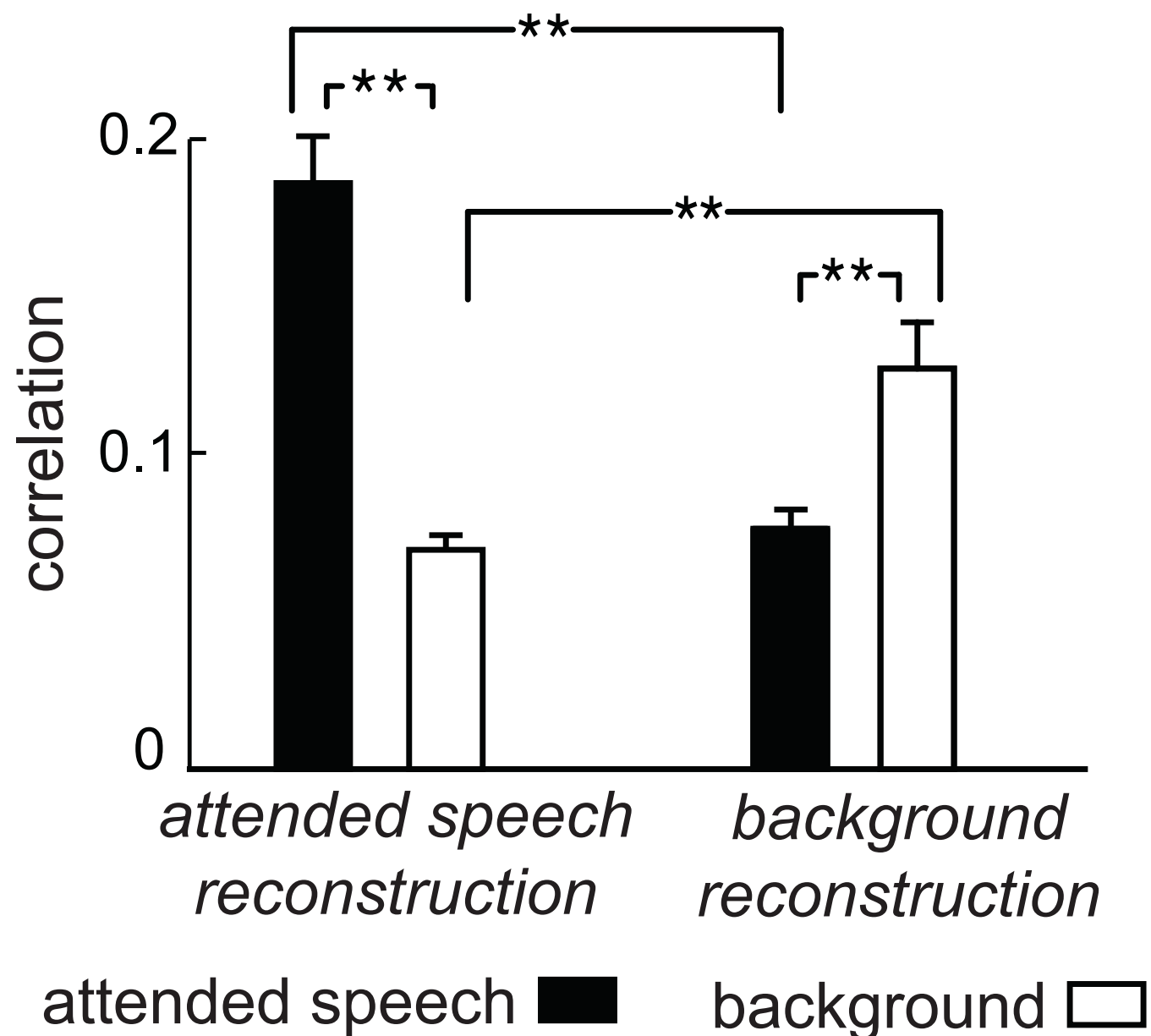
Background Speech Reconstruction



Overall Speech Reconstruction

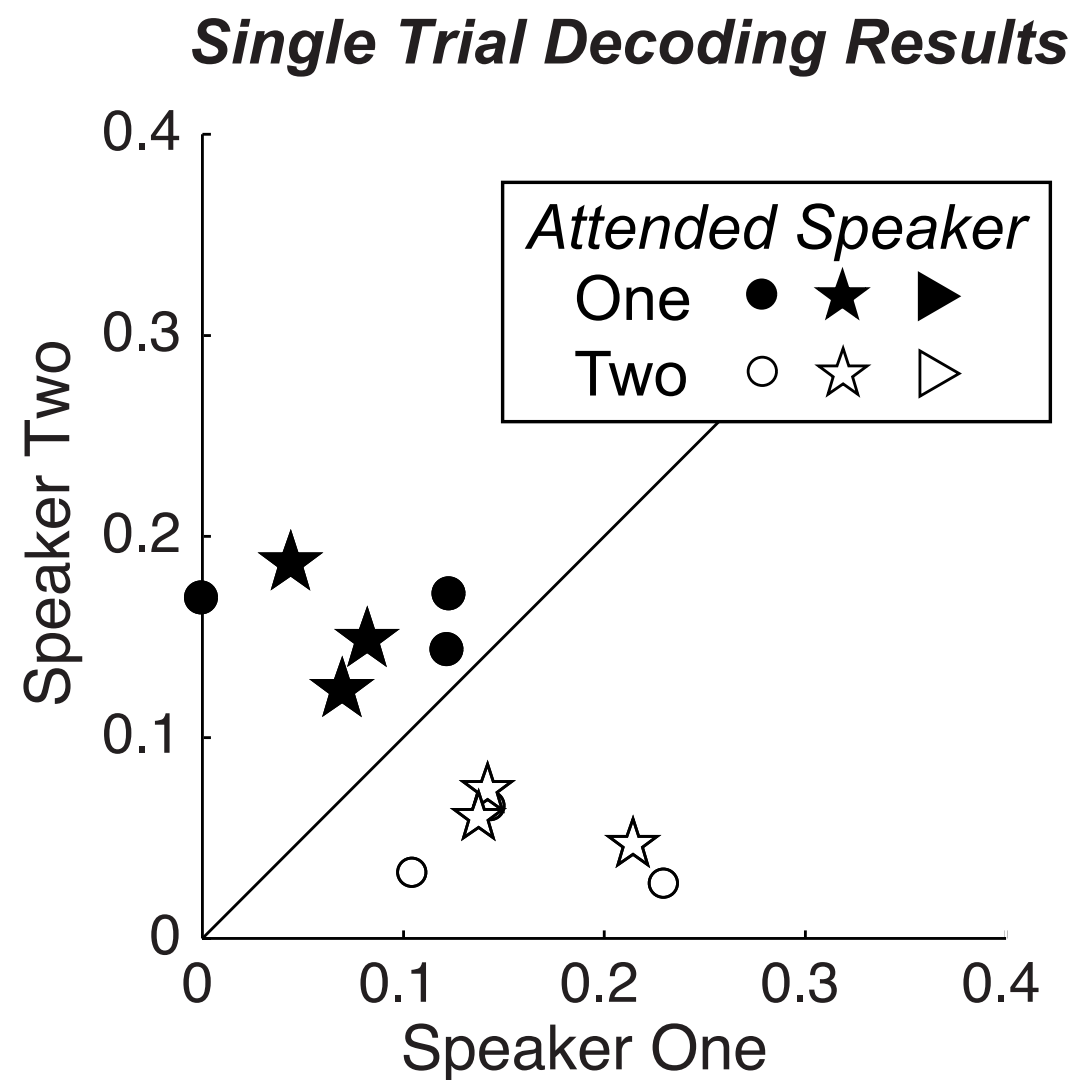
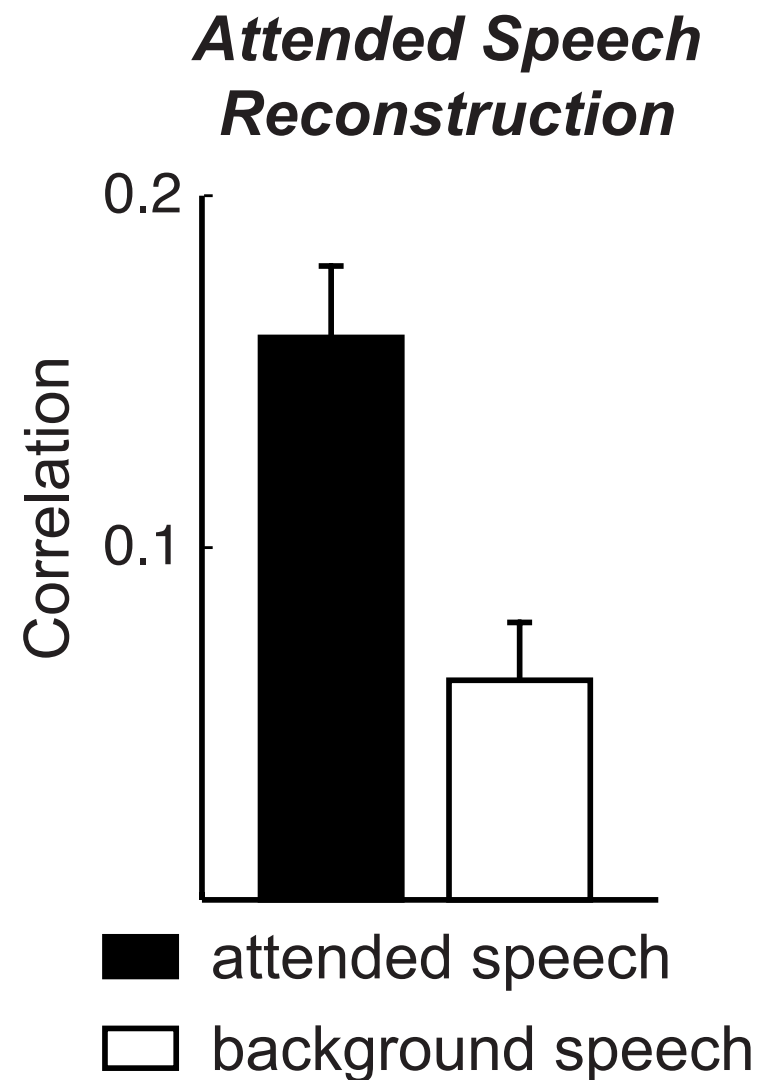


Overall Speech Reconstruction

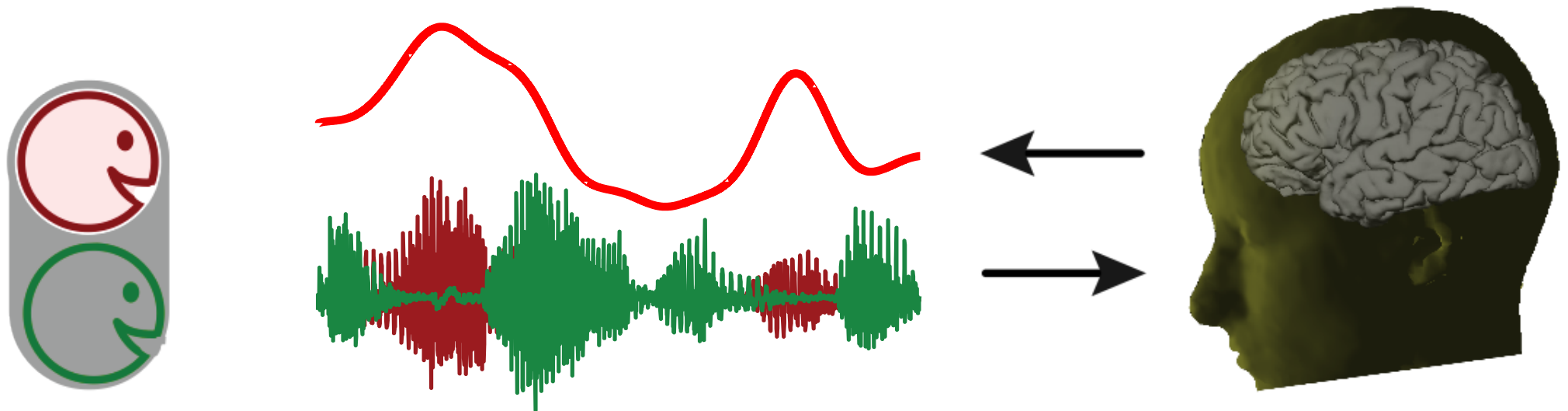


Distinct neural representations for different speech streams

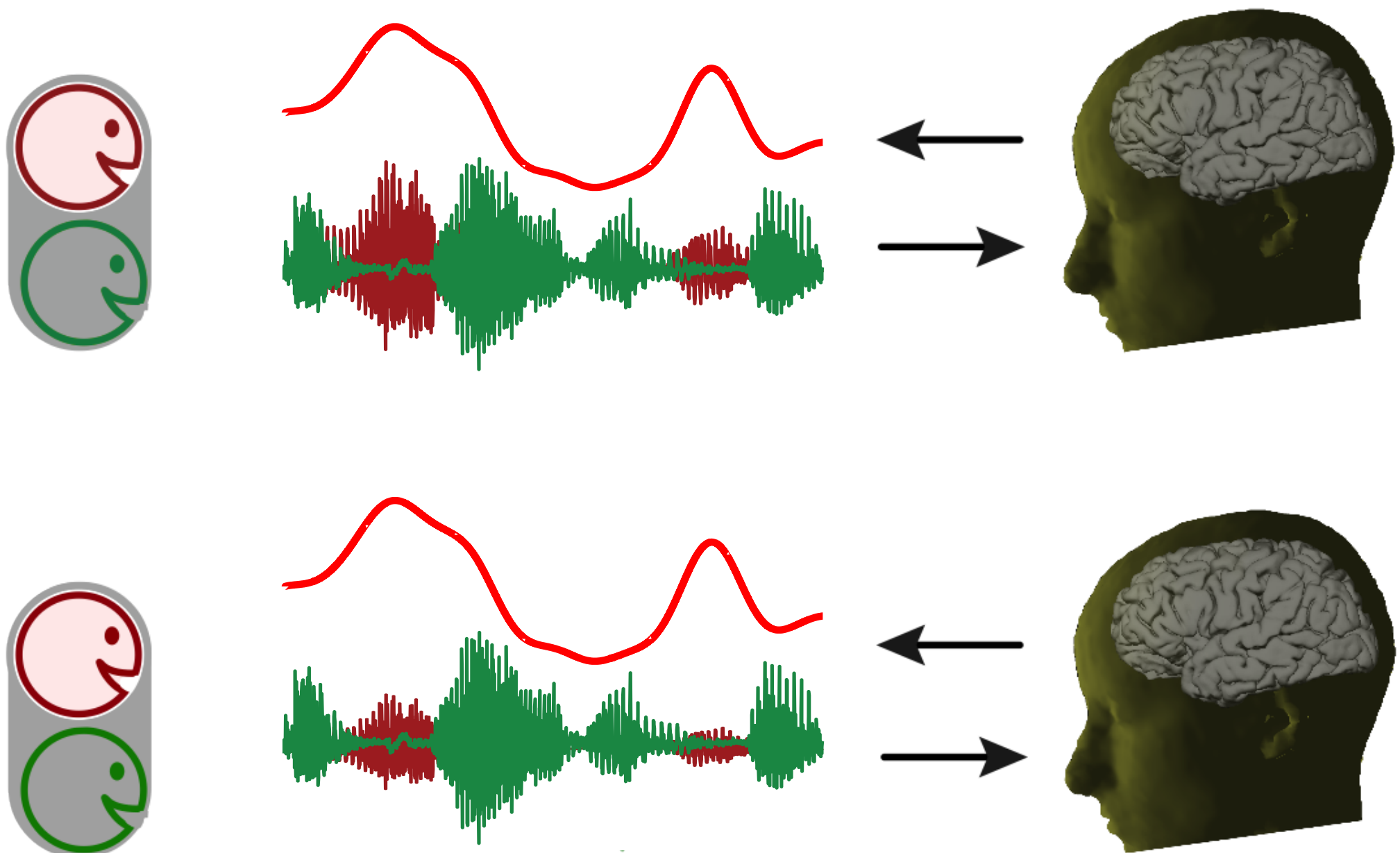
Reconstruction of Same-Sex Speech



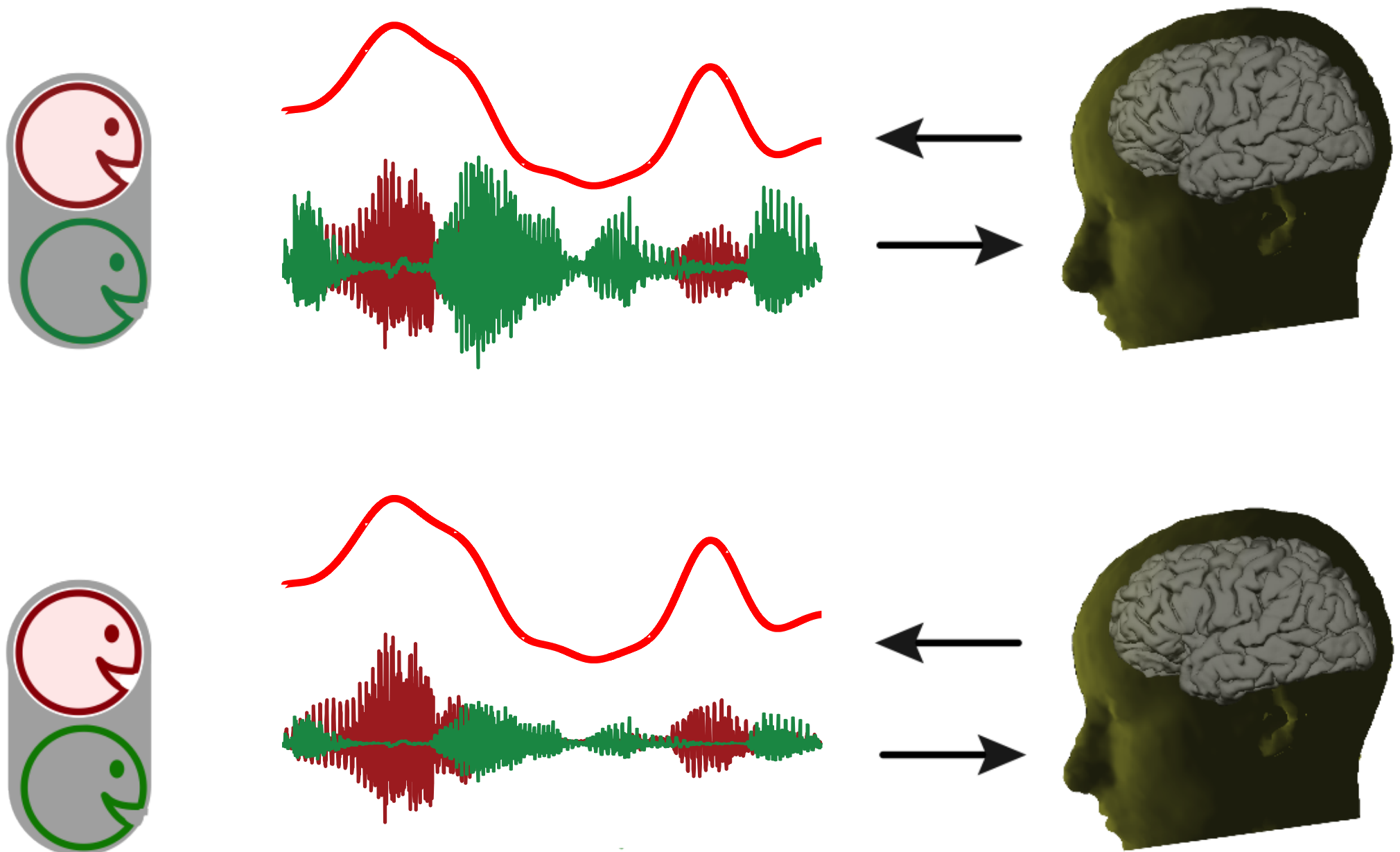
Invariance Under Acoustic Changes



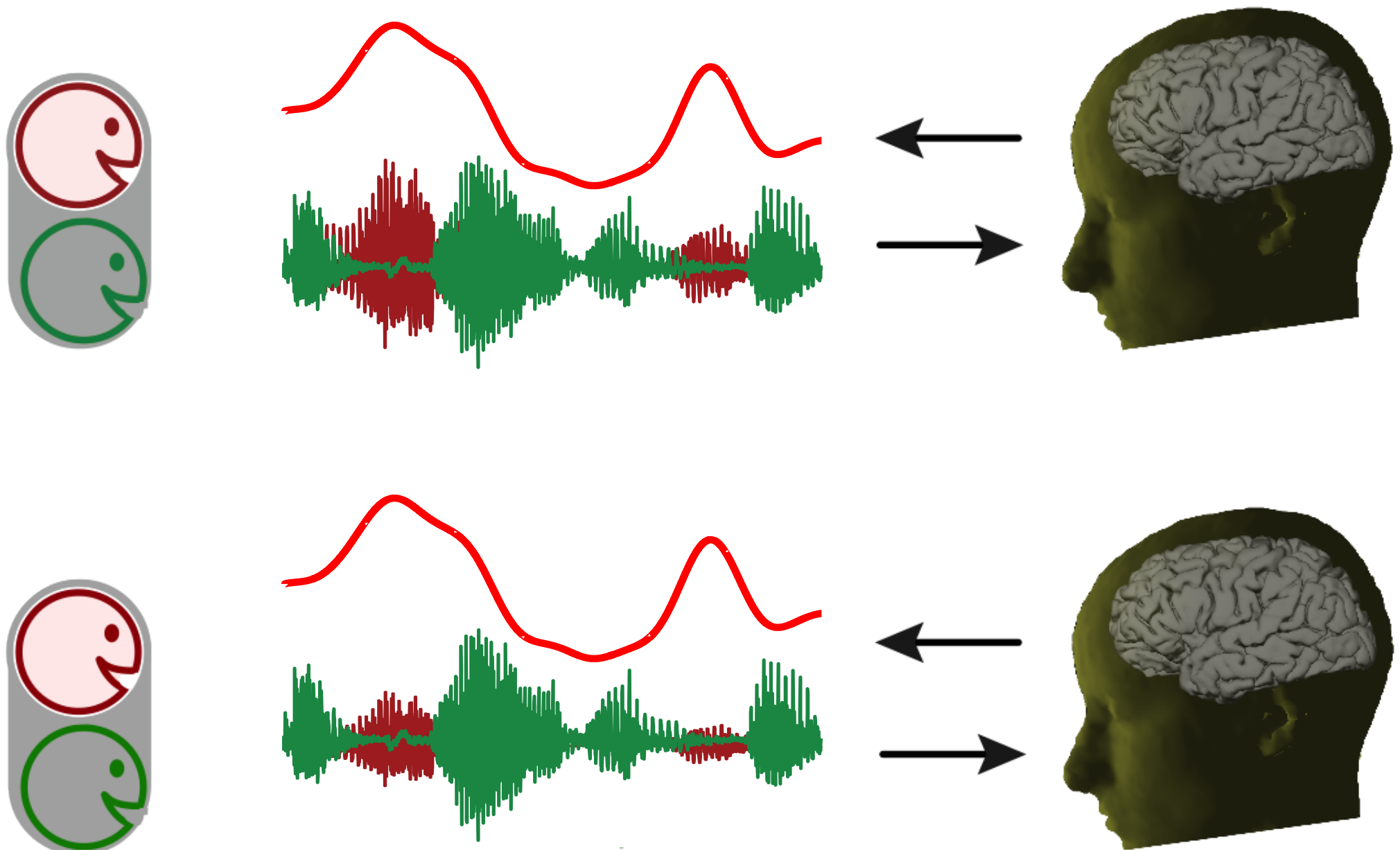
Invariance Under Acoustic Changes



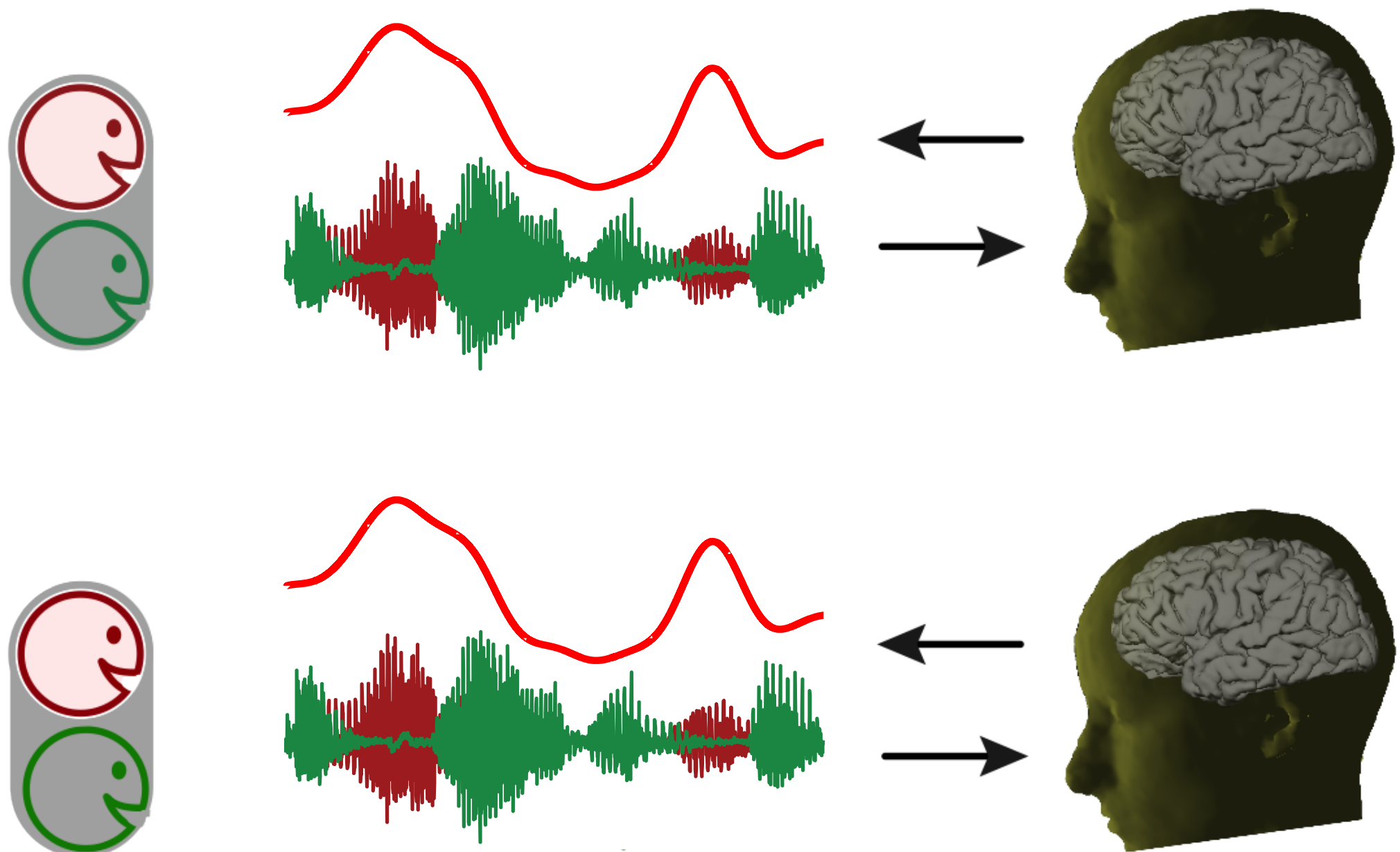
Invariance Under Acoustic Changes



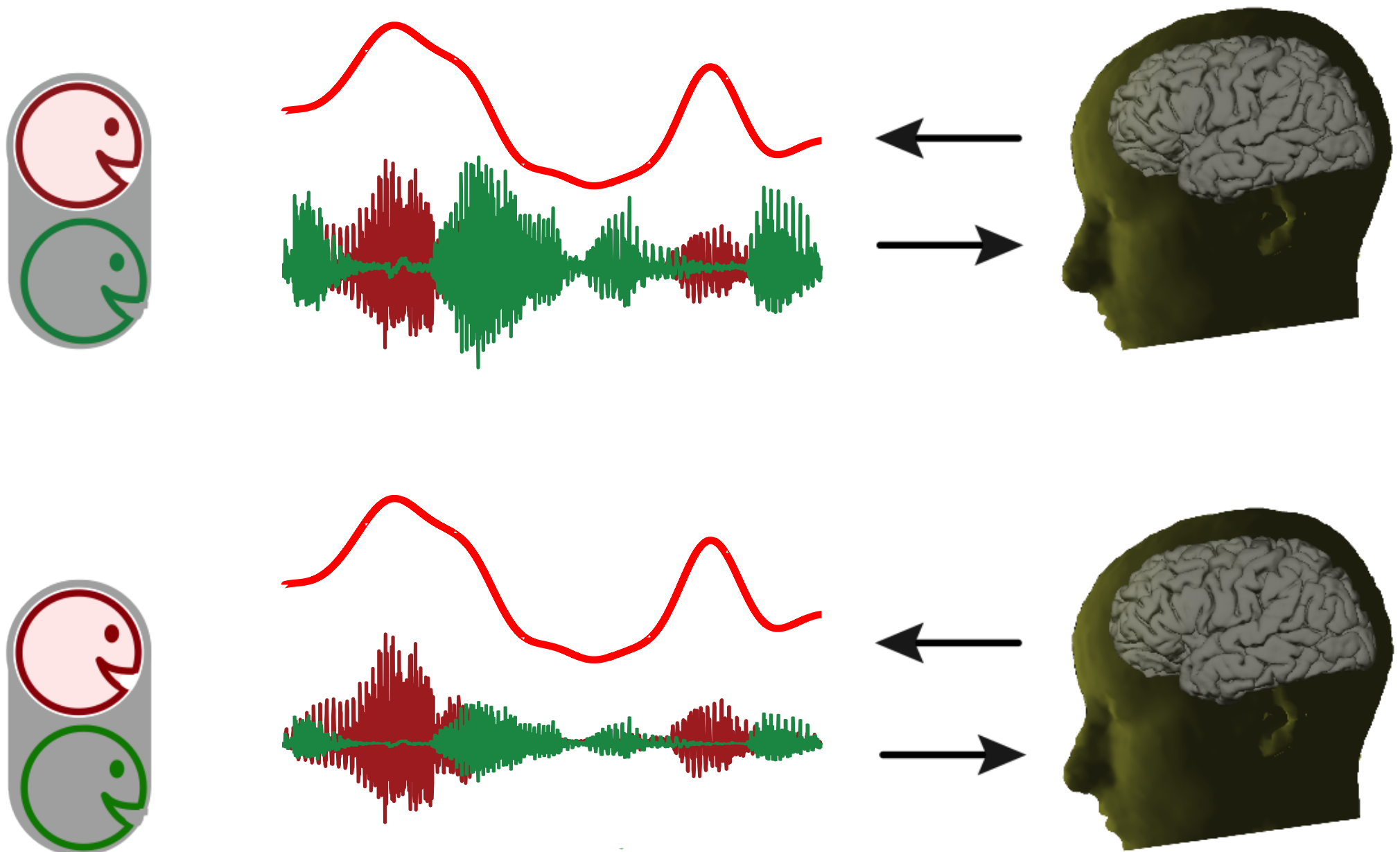
Invariance Under Acoustic Changes



Invariance Under Acoustic Changes



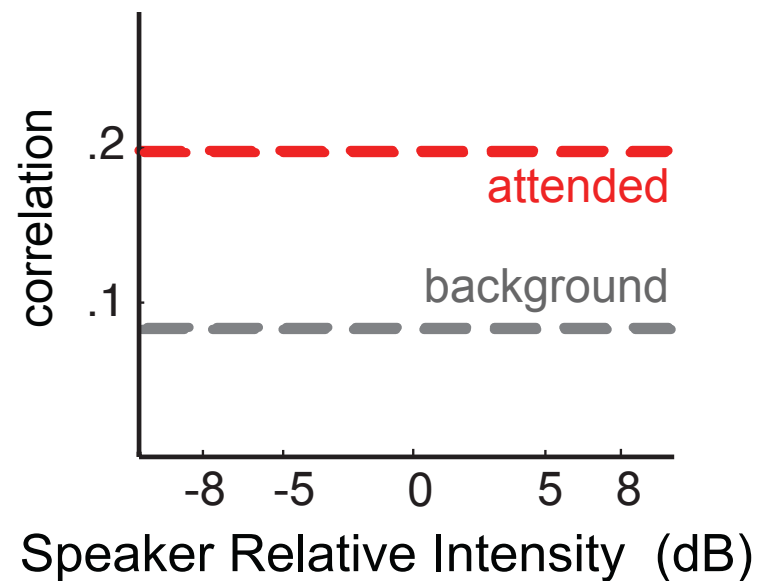
Invariance Under Acoustic Changes



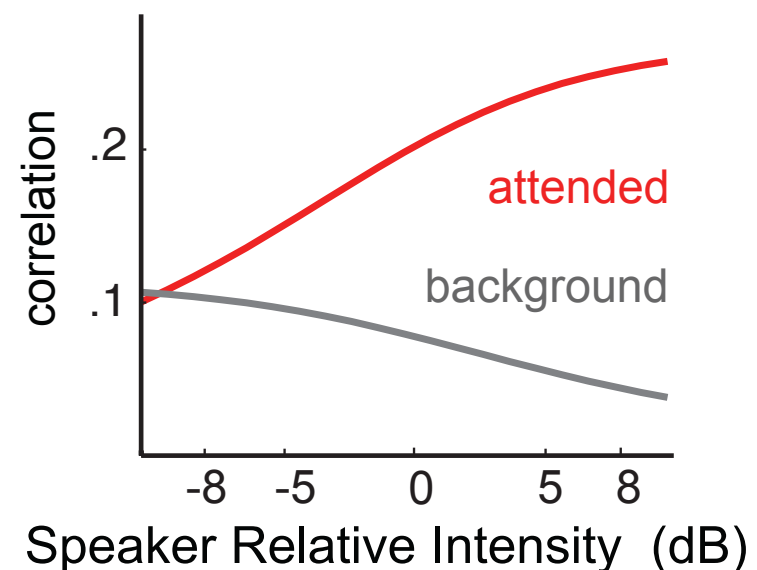
Object-Based Gain Control?

Gain-Control Models

Object-Based



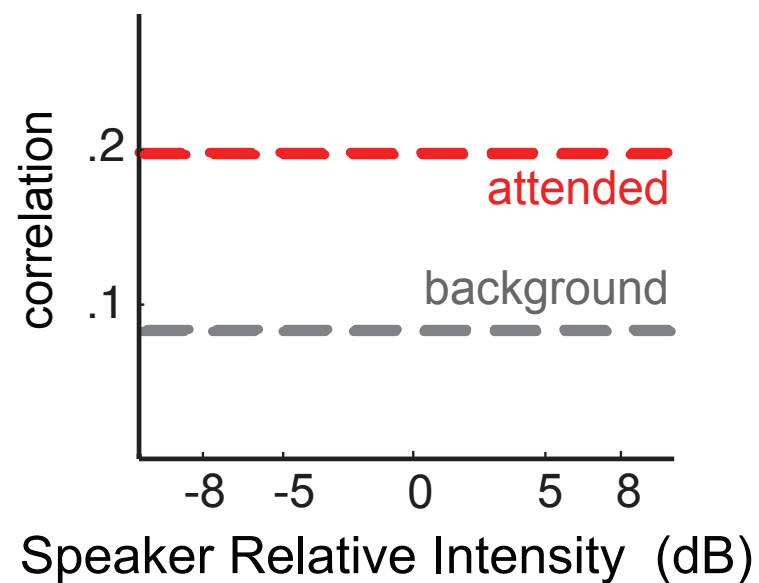
Globally Based



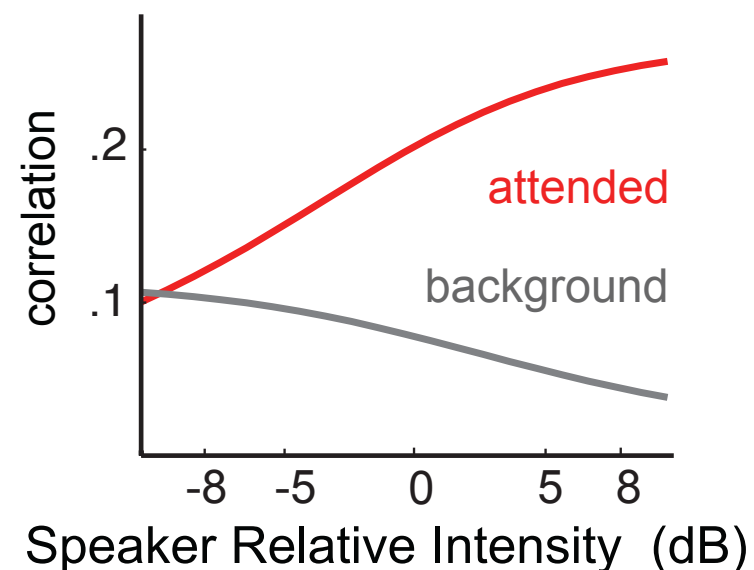
Object-Based Gain Control?

Gain-Control Models

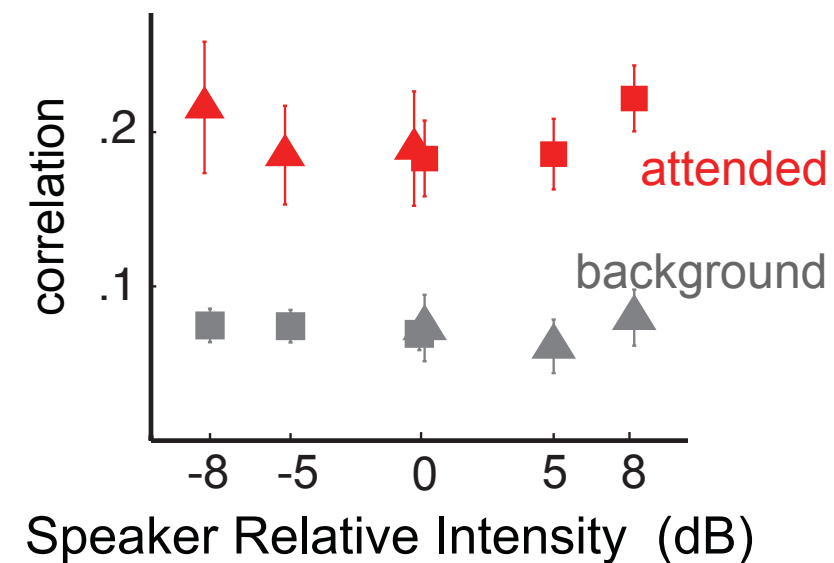
Object-Based



Globally Based



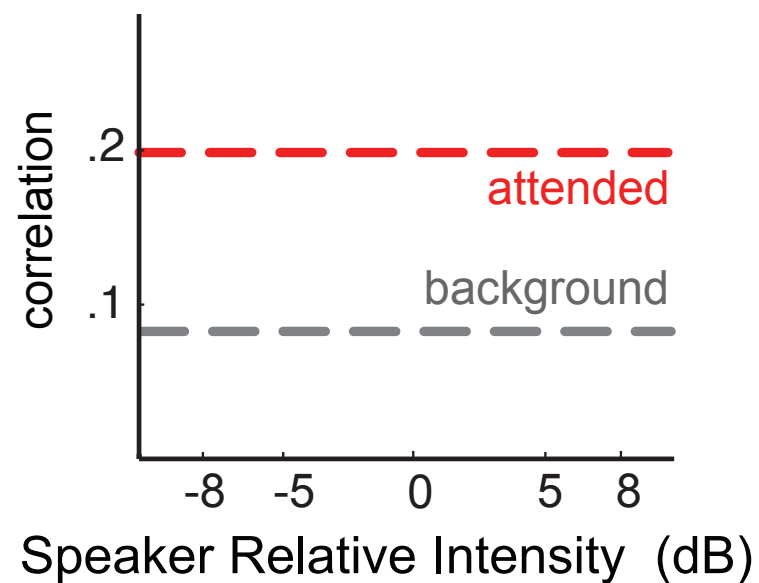
Neural Results



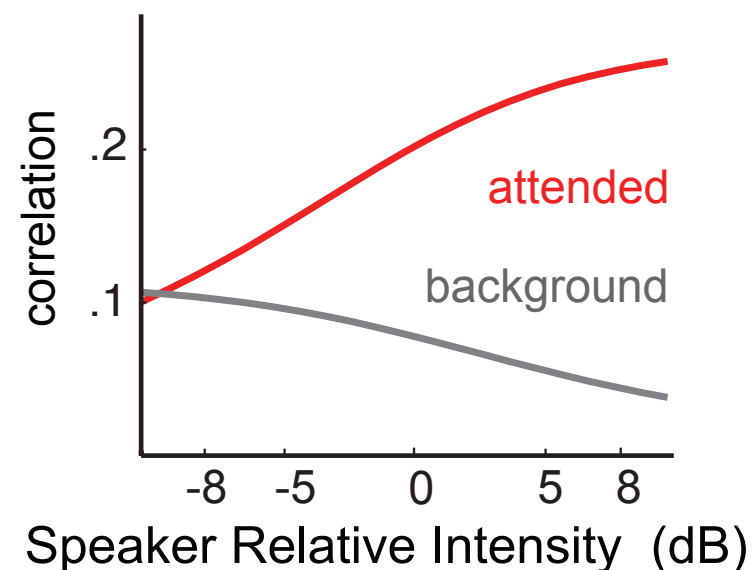
Object-Based Gain Control?

Gain-Control Models

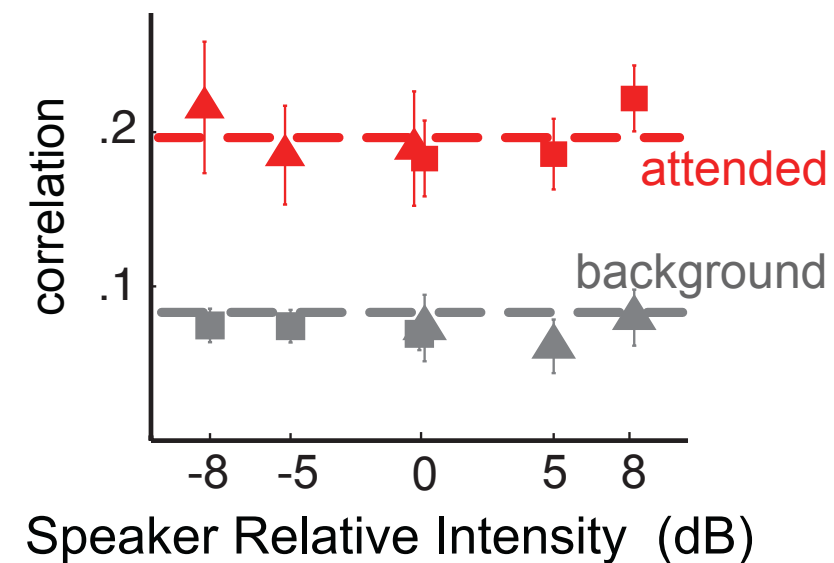
Object-Based



Globally Based



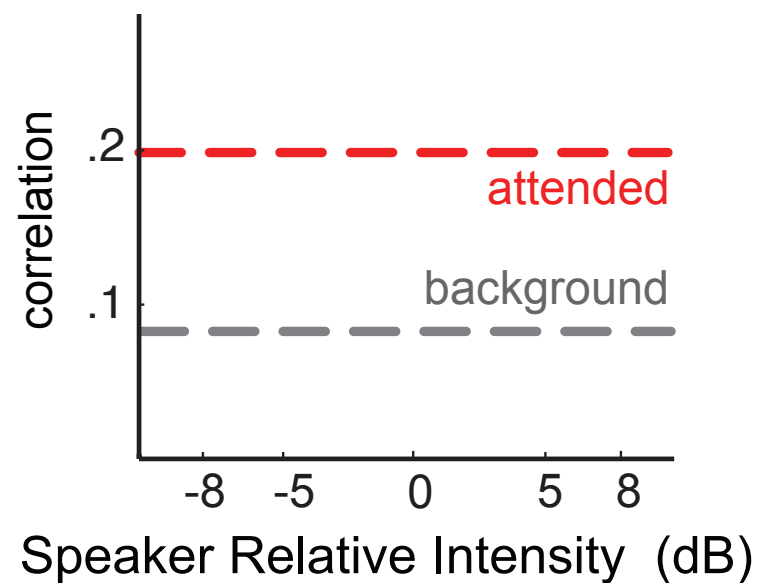
Neural Results



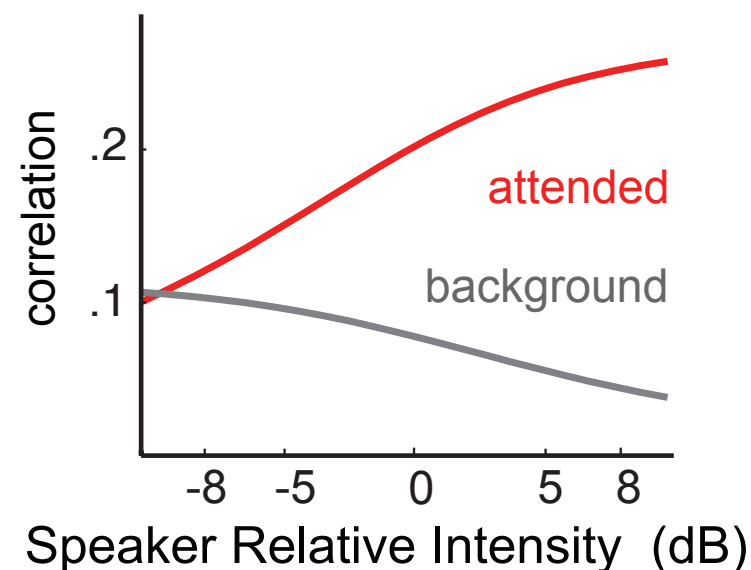
Object-Based Gain Control?

Gain-Control Models

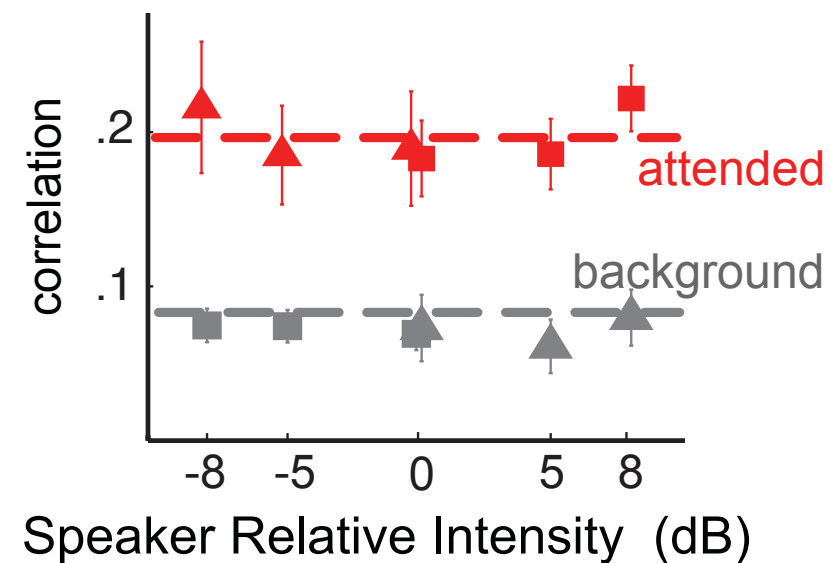
Object-Based



Globally Based

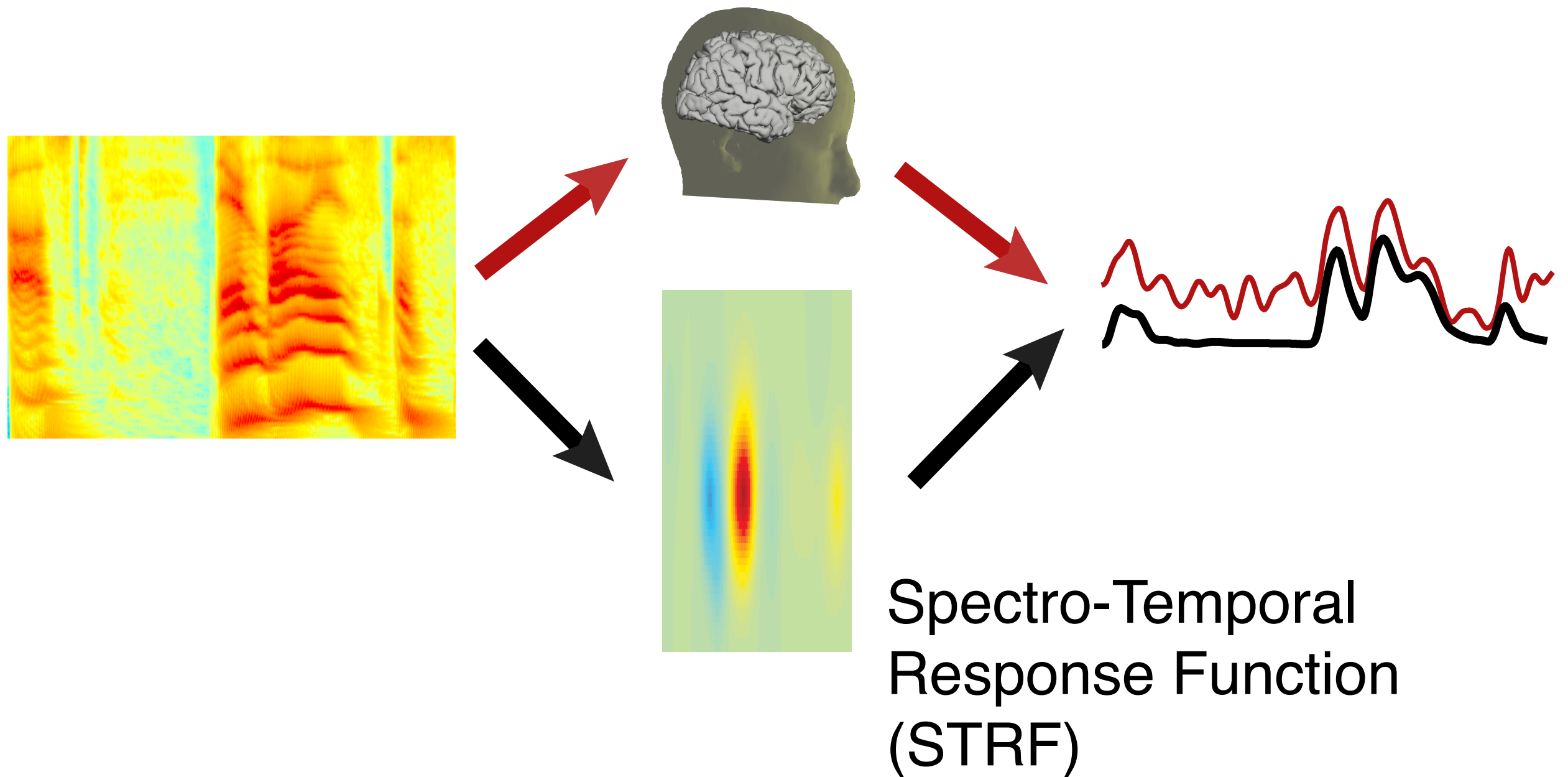


Neural Results

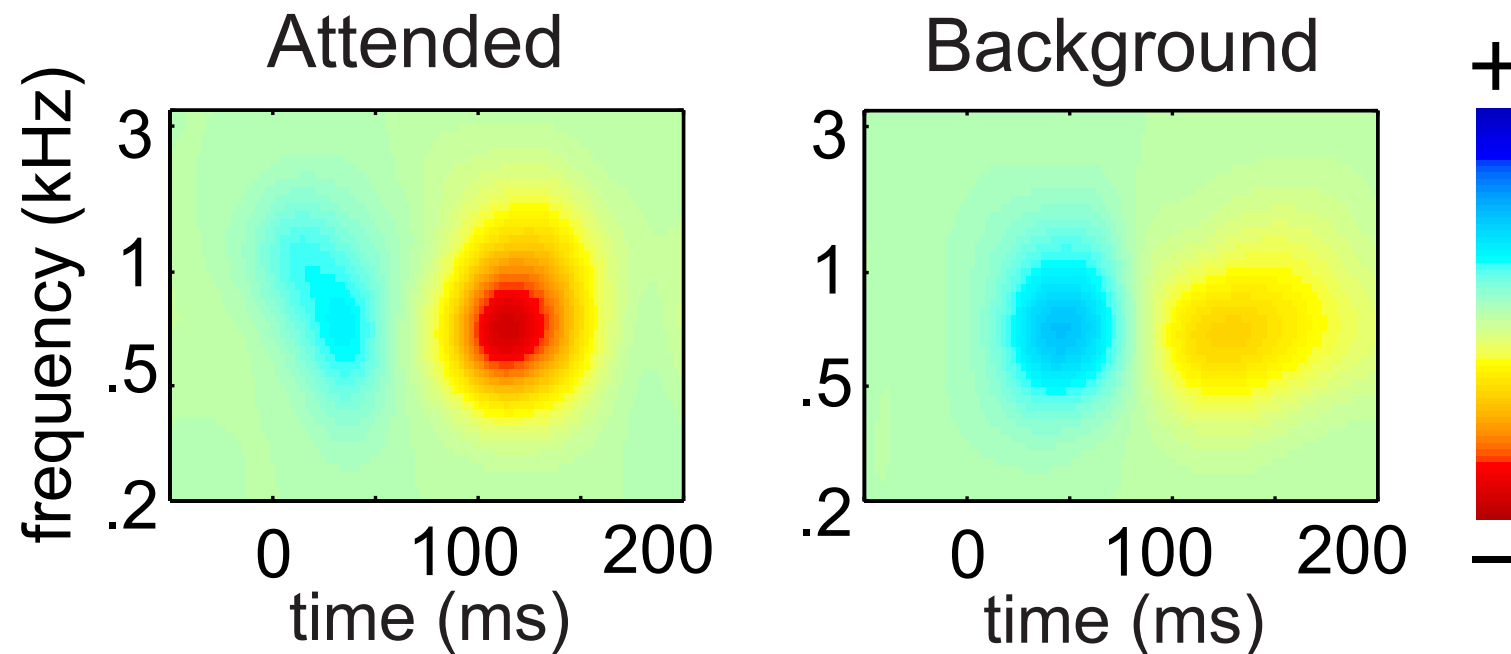


- Gain Control is Object-Based
- Neural representation is invariant to acoustic changes.

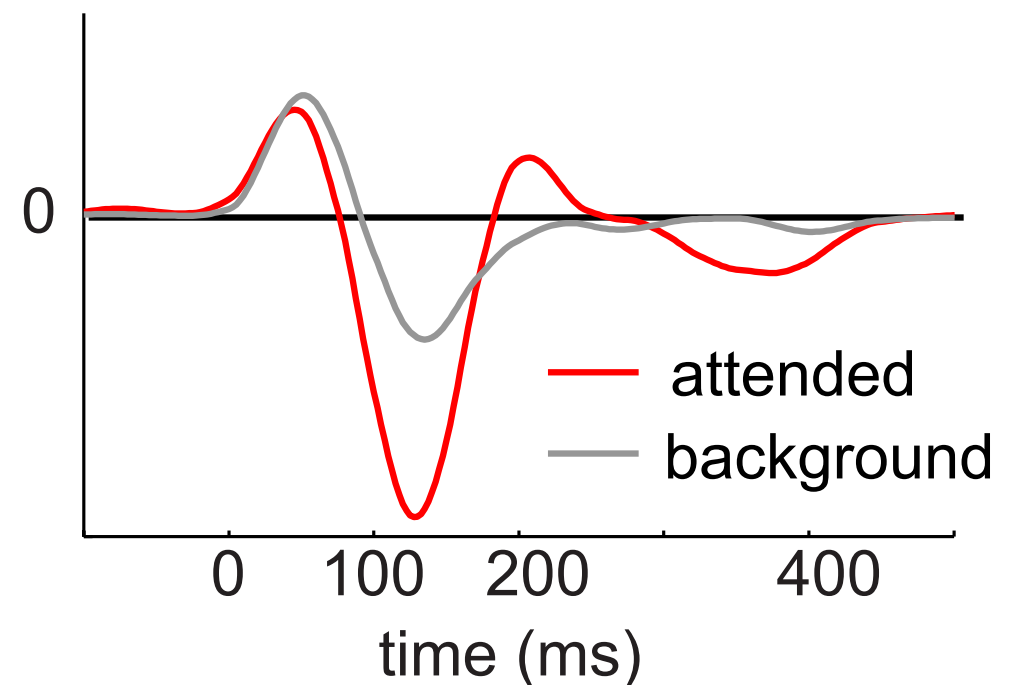
STRF model



STRF Results

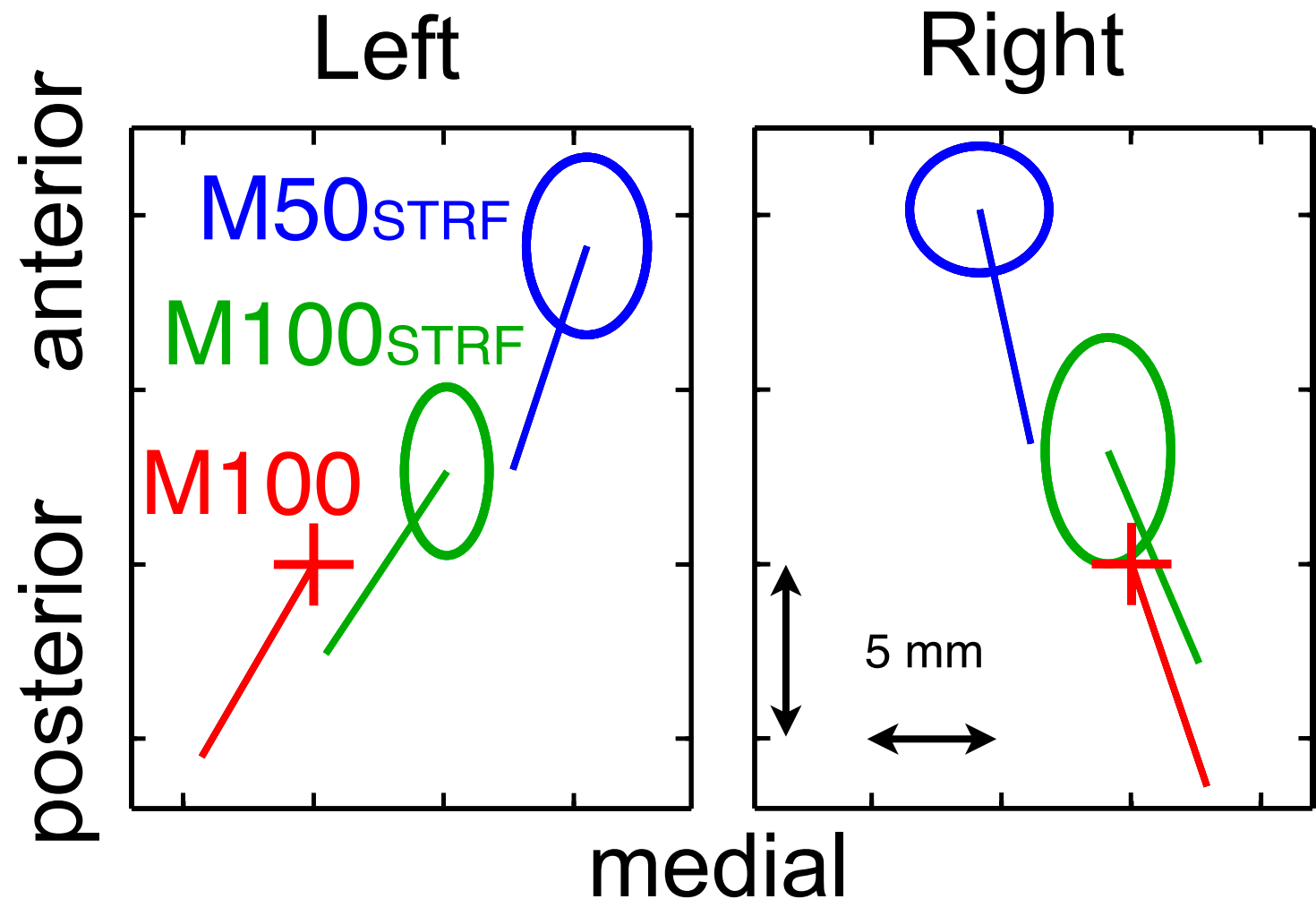


- STRF separable (time, frequency)
- 300 Hz - 2 kHz dominant carrier
- M50_{STRF} positive peak
- M100_{STRF} negative peak
- M100_{STRF} strongly modulated, but not M50_{STRF}

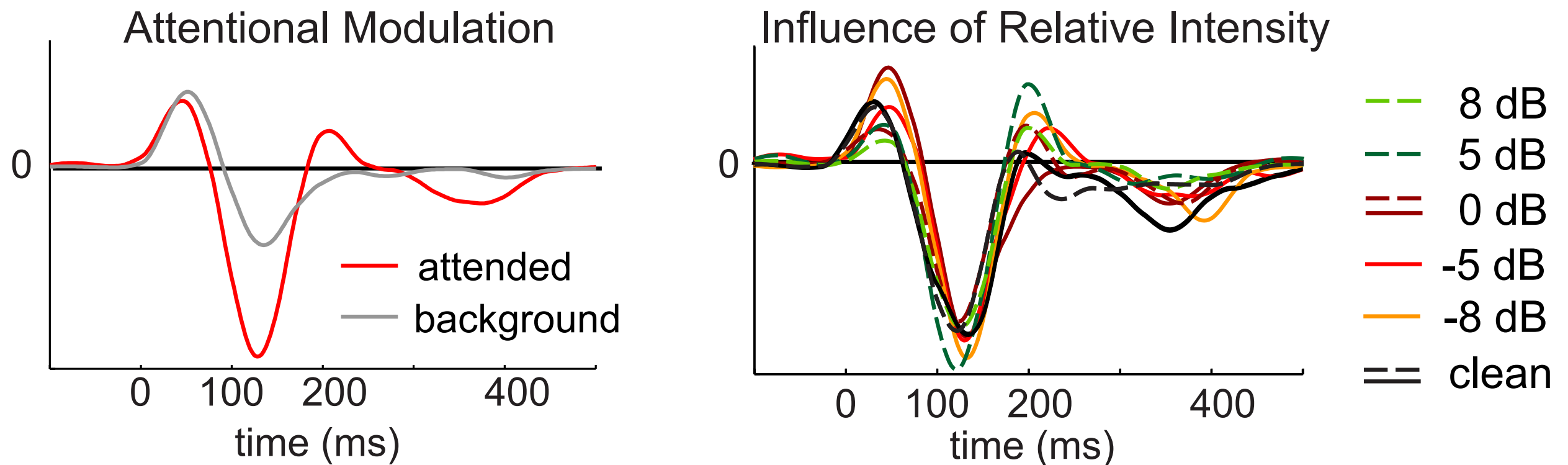


Neural Sources

- M100_{STRF} source near (same as?) M100 source: STG/PT
- M50_{STRF} source is anterior and medial to M100 (same as M50?): HG



Cortical Object-Processing Hierarchy



- $M100_{STRF}$ strongly modulated by attention, but not $M50_{STRF}$.
- $M100_{STRF}$ invariant against acoustic changes (but not $M50_{STRF}$?).
- Objects well-neurally represented at 100 ms, but not 50 ms.

Summary

- Cortical representations of speech show properties consistent with being neural representations of auditory objects
 - Meet three formal criteria
- Object representation well-formed at 100 ms latency (STG, PT), but not at 50 ms (HG)

Acknowledgements

Collaborators

Catherine Carr
Alain de Cheveigné
Didier Depireux
Mounya Elhilali
Jonathan Fritz
Cindy Moss
David Poeppel
Shihab Shamma

Past Postdocs

Dan Hertz
Yadong Wang

Current Grad Students

Francisco Cervantes
Marisel Villafane Delgado
Kim Drnec
Krishna Puvvada

Past Grad Students

Nayef Ahmar
Claudia Bonin
Maria Chait

Nai Ding

Victor Grau-Serrat
Ling Ma
Raul Rodriguez
Juanjuan Xiang
Kai Sum Li
Jiachen Zhuo

Collaborators' Students

Murat Aytekin
Julian Jenkins
David Klein
Huan Luo

Undergraduate Students

Abdulaziz Al-Turki
Nicholas Asendorf
Sonja Bohr
Elizabeth Camenga
Corinne Cameron
Julien Dagenais
Katya Dombrowski
Kevin Kahn
Andrea Shome
Ben Walsh

Funding

NIH R03 DC 004382
NIH R01 EB 004750
NIH R01 AG 027573
NIH R01 DC 007657
NIH R01 DC 000436

NIH R01 DC 008342

NIH R01 DC 005660
NIH F31 NS 055589
USDA 20096512005791