Recent Advances in Cortical Representations of Speech using MEG

Jonathan Z. Simon

Department of Electrical & Computer Engineering

Department of Biology

Institute for Systems Research

University of Maryland

We are looking for a Post-doc!

Modern Signal Processing Techniques Applied to Analysis of Neural Recordings

University of Maryland, College Park
(Just outside of Washington DC)

Jonathan & Behtash

Acknowledgements

Lab Members & Affiliates

Ross Baehr
Christian Brodbeck
Proloy Das
Alex Jiao
Joshua Kulasingham
Sina Miran
David Nahmias
Jonas Vanthornhout
Peng Zan
Natalia Lapinskaya
Huan Luo
Mahshid Najafi
Alex Presacco
Krishna Puvvada
Ben Walsh
Yadong Wang
Juanjuan Xiang
Jiachen Zhuo
Mounya Elhilali
Tom Francart
Jonathan Fritz
Michael Fu
Stefanie Kuchinsky
Steven Marcus
Cindy Moss
David Poeppel
Shihab Shamma

Past Lab Members & Affiliates

Nayef Ahmar
Sahar Akram
Murat Aytekin
Francisco Cervantes Constantino
Maria Chait
Marisel Villafane Delgado
Kim Drnec
Nai Ding
Victor Grau-Serrat
Julian Jenkins
Pirazh Khorramshahi

Collaborators

Pamela Abshire
Samira Anderson
Behtash Babadi
Catherine Carr
Monita Chatterjee
Alain de Cheveigné
Stephen David
Didier Depireux

Past Undergraduate Students

Nicholas Asendorf
Anurupa Bhonsale
Sonja Bohr
Elizabeth Camenga
Julien Dagenais
Katya Dombrowski
Kevin Hogan
Andrea Shome
James Williams

Funding NIH (NIDCD, NIA, NIBIB); NSF; DARPA; USDA; UMD

Outline

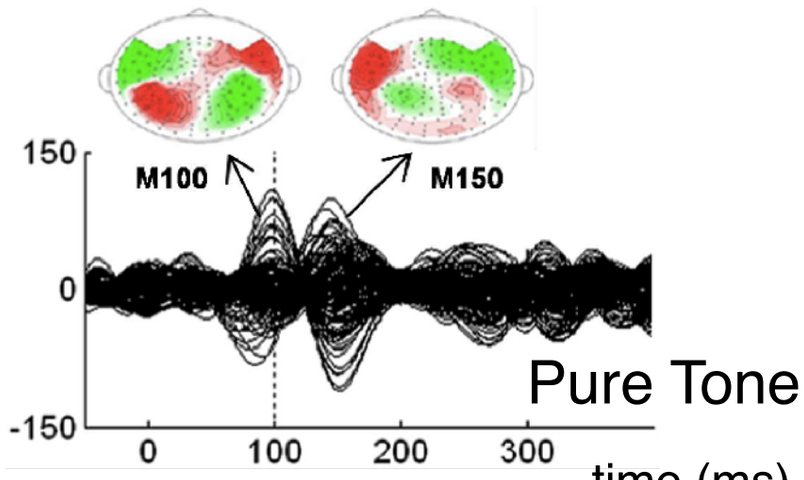
- Cortical Representations of Speech via MEG
- Missing speech / Familiar Speech
- Spatial Localization of Cortical Representations
  - Multiple Levels of Cortical Representations
- Representation Transition from Auditory/Phonetic → Lexical/Word-based
- Recent advances in Spatial Localization

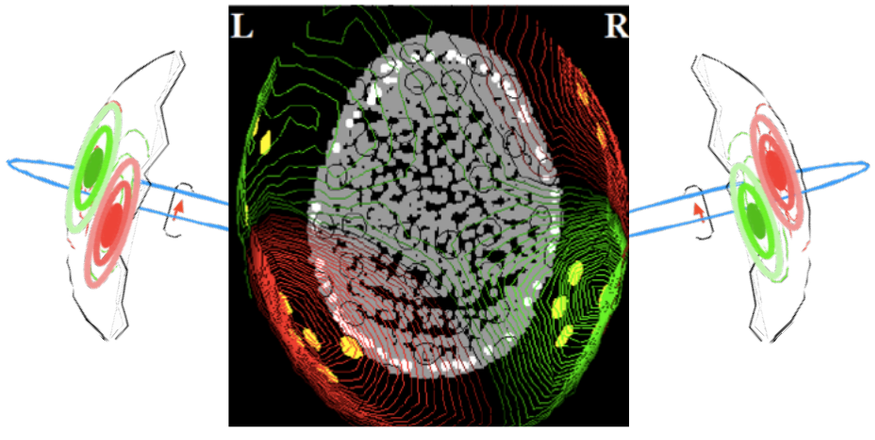
Outline

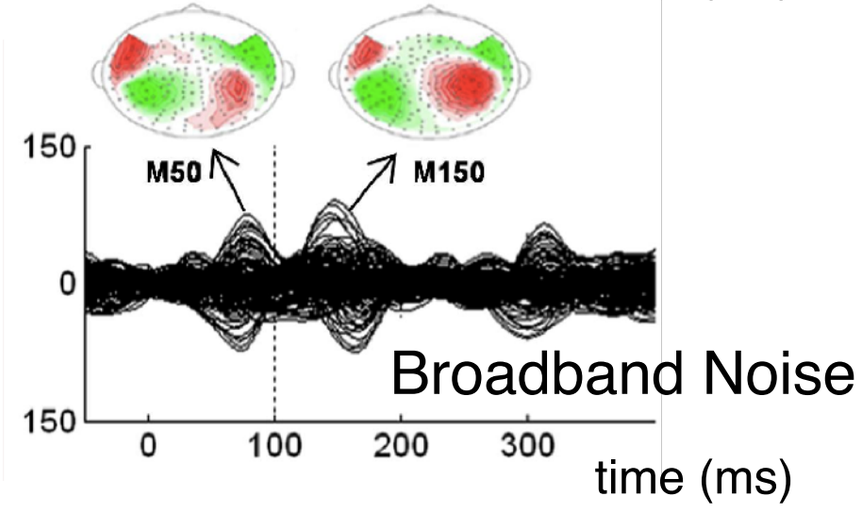
- Cortical Representations of Speech via MEG
- Missing speech / Familiar Speech
- Spatial Localization of Cortical Representations
  - Multiple Levels of Cortical Representations
- Representation Transition from Auditory/Phonetic → Lexical/Word-based
- Recent advances in Spatial Localization

MEG & Auditory Cortex

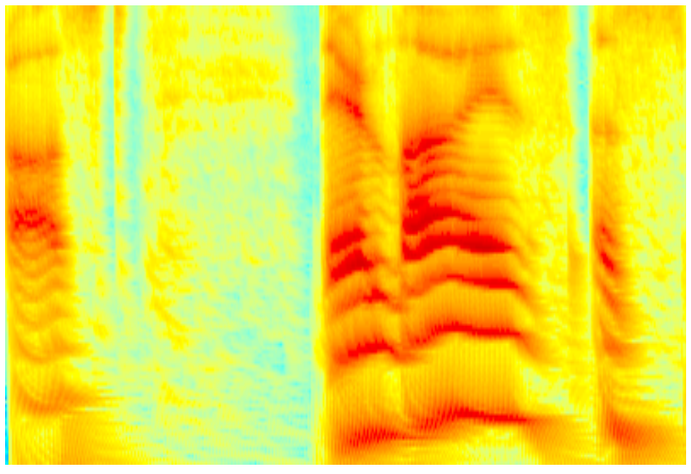
- Non-invasive, Passive, Silent Neural Recordings
- MEG Response Patterns Time-Locked to Stimulus Events
- Robust
- Strongly Lateralized
- Cortical Origin Only (some exceptions)

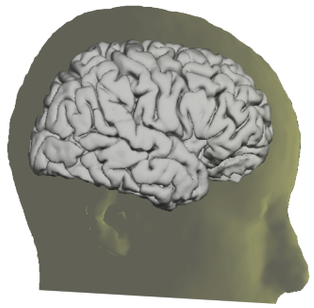




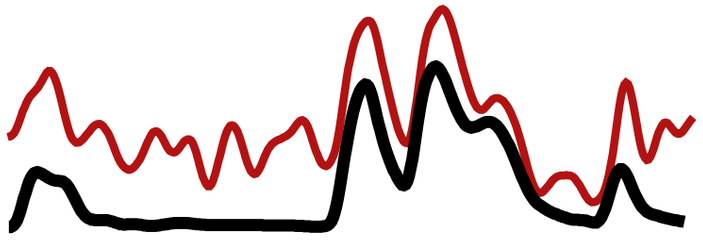


MEG Responses to Speech Modulations

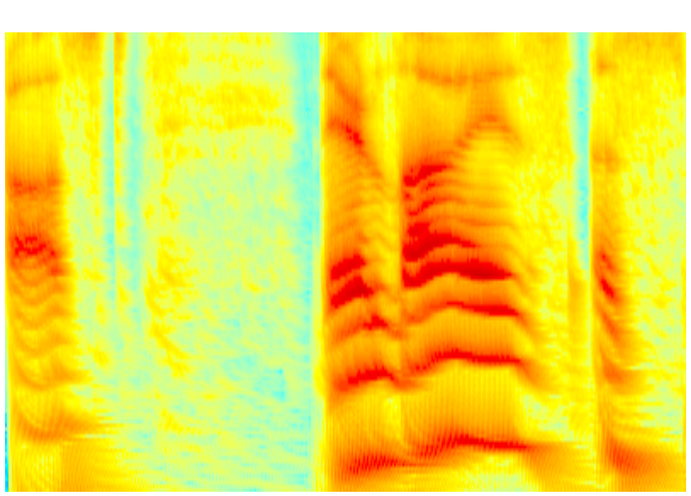


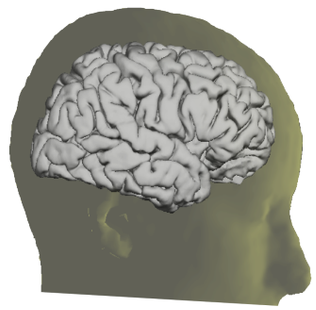


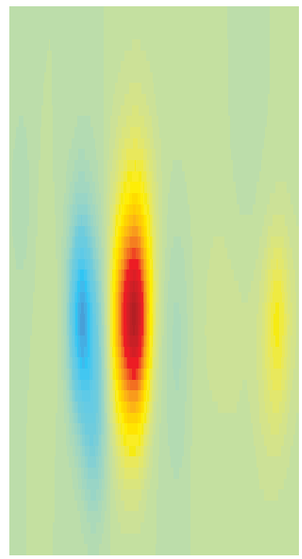
Auditory Model

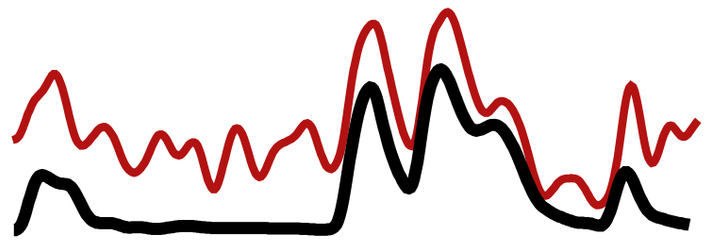


MEG Responses Predicted by STRF Model









Long duration speech: ~60 s

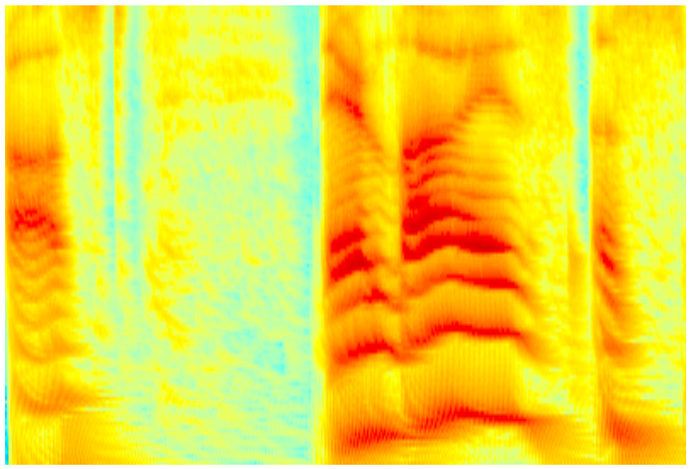
(up to ~10 Hz)

Linear Kernel = STRF

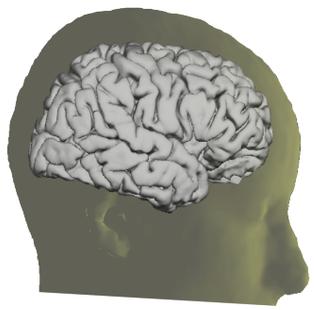
"Spectro-Temporal Response Function"

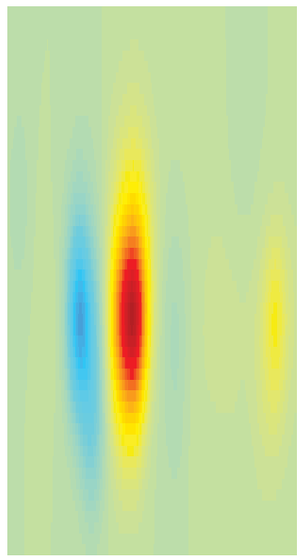
Ding & Simon, J Neurophysiol (2012)

MEG Responses Predicted by STRF Model

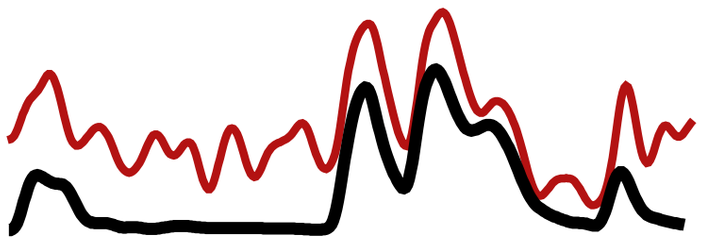


Long duration speech: ~60 s





(up to ~10 Hz)

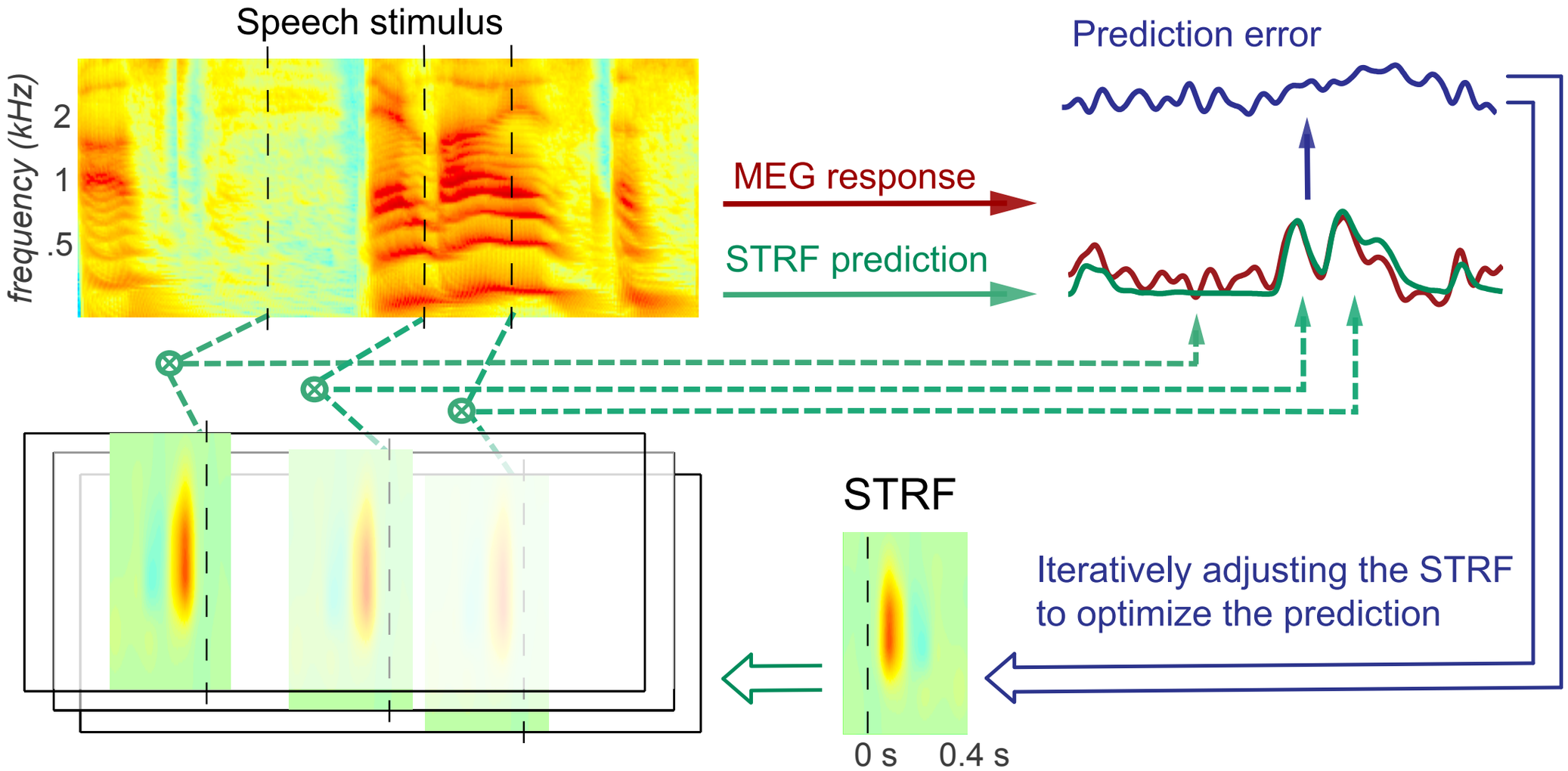


Linear Kernel = STRF

"Spectro-Temporal Response Function"

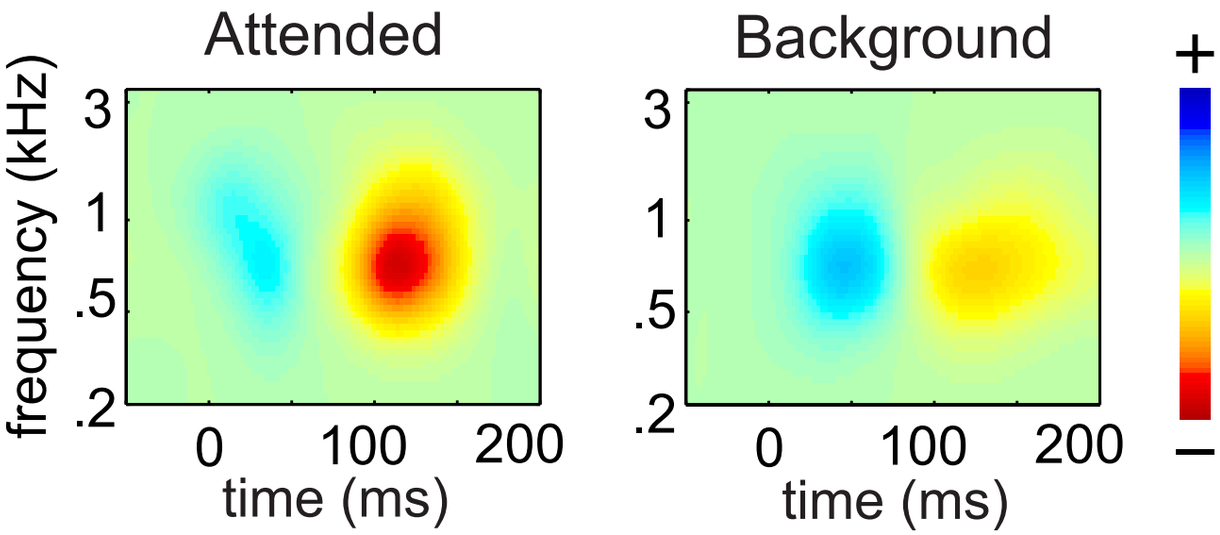
Ding & Simon, J Neurophysiol (2012)

STRF Estimation via Boosting



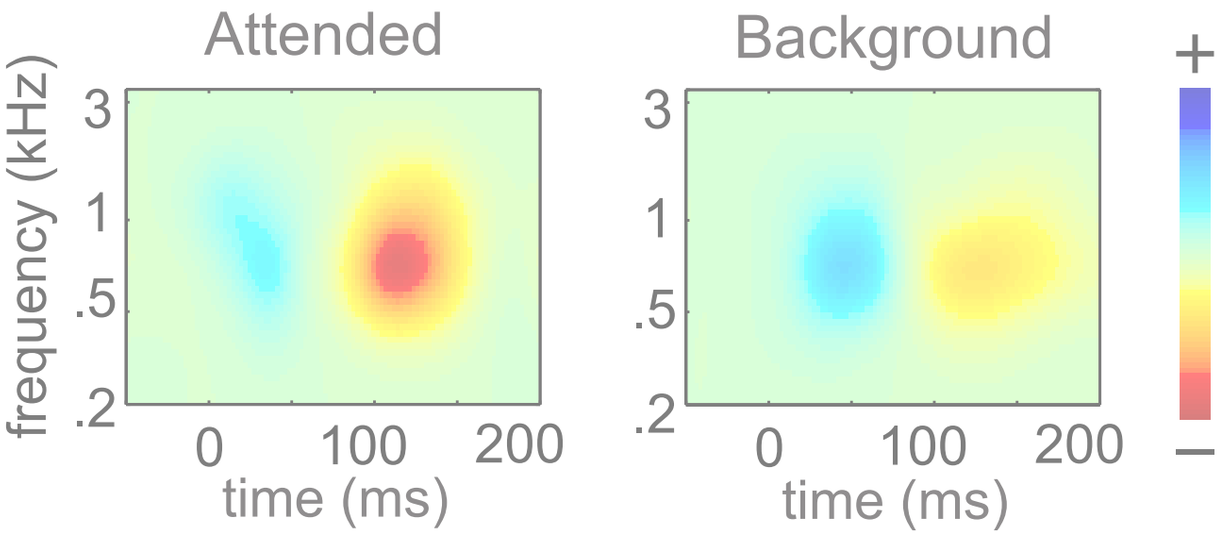
Ding & Simon, J Neurophysiol (2012)

STRF vs. TRF

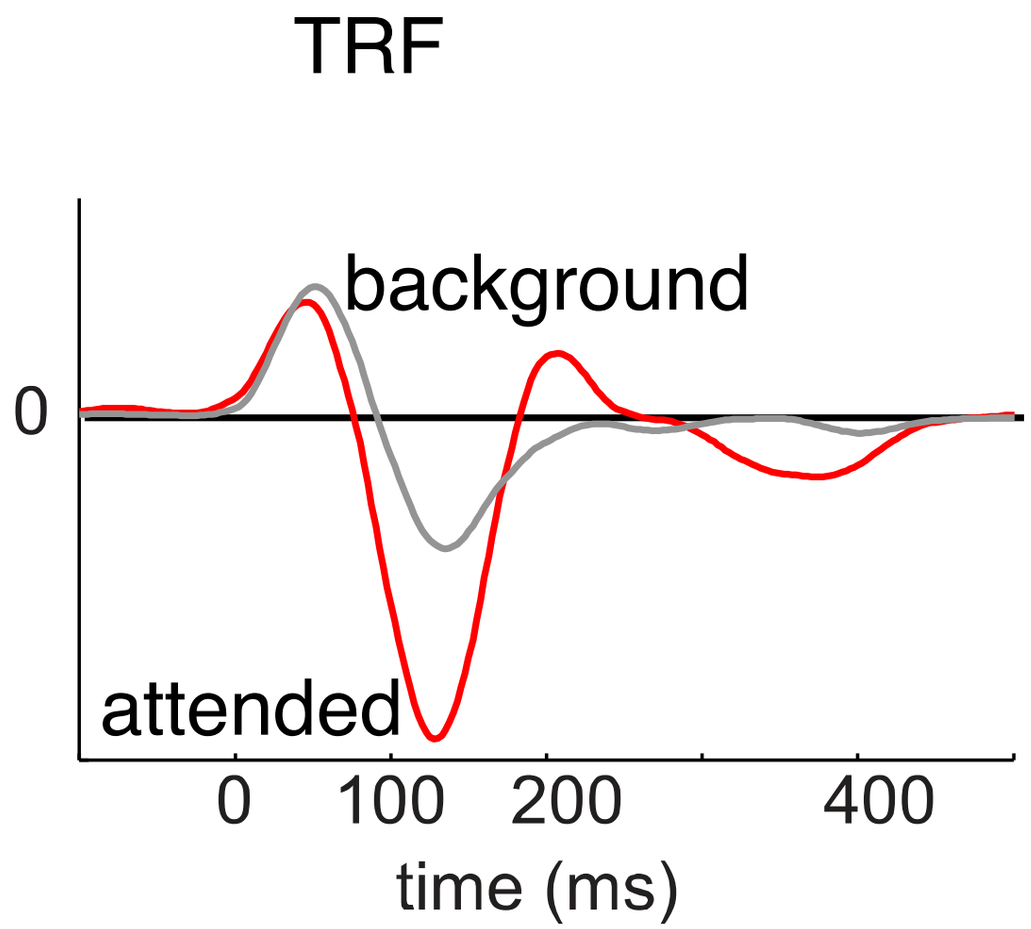


- STRF separable (time, frequency)
- 300 Hz - 2 kHz dominant carriers
- $M50_{STRF}$ positive peak
- $M100_{STRF}$ negative peak

STRF vs. TRF

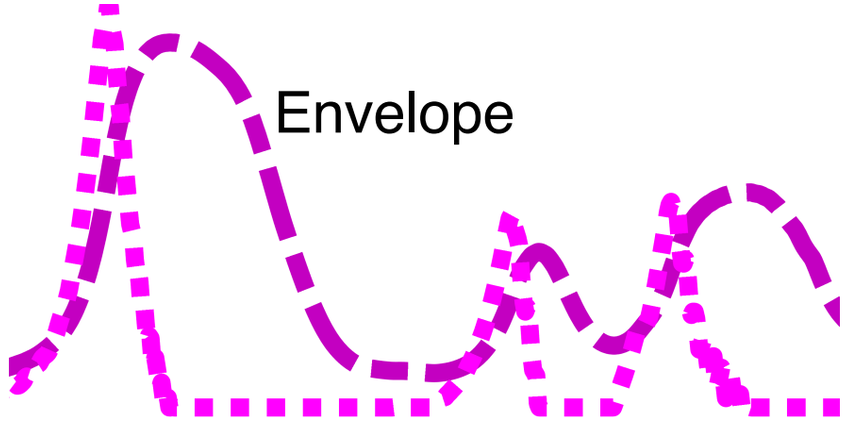


- STRF separable (time, frequency)
- 300 Hz - 2 kHz dominant carriers
- $M50_{STRF}$ positive peak
- $M100_{STRF}$ negative peak



TRF: Depends on Stimulus Representation

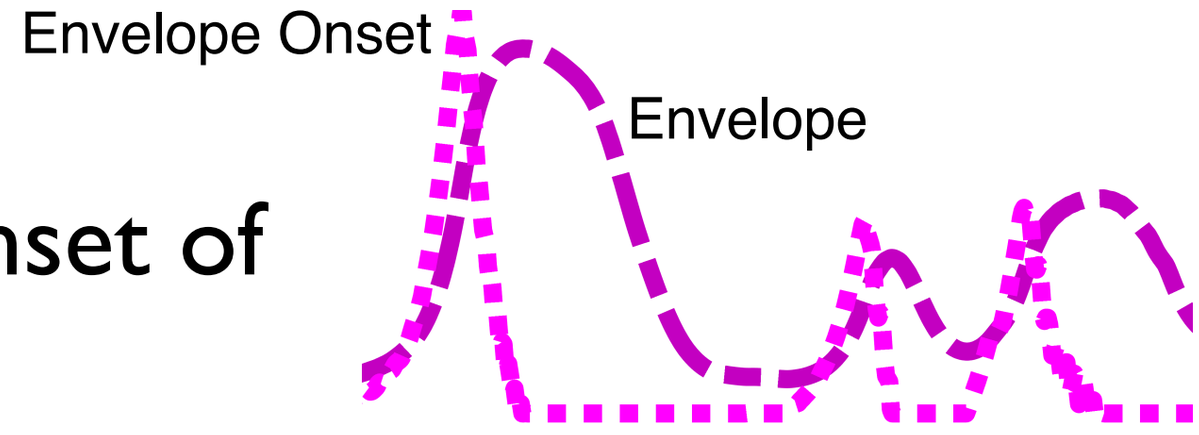
- Envelope vs. Onset of Envelope



Cervantes Constantino et al., bioRxiv 168997

TRF: Depends on Stimulus Representation

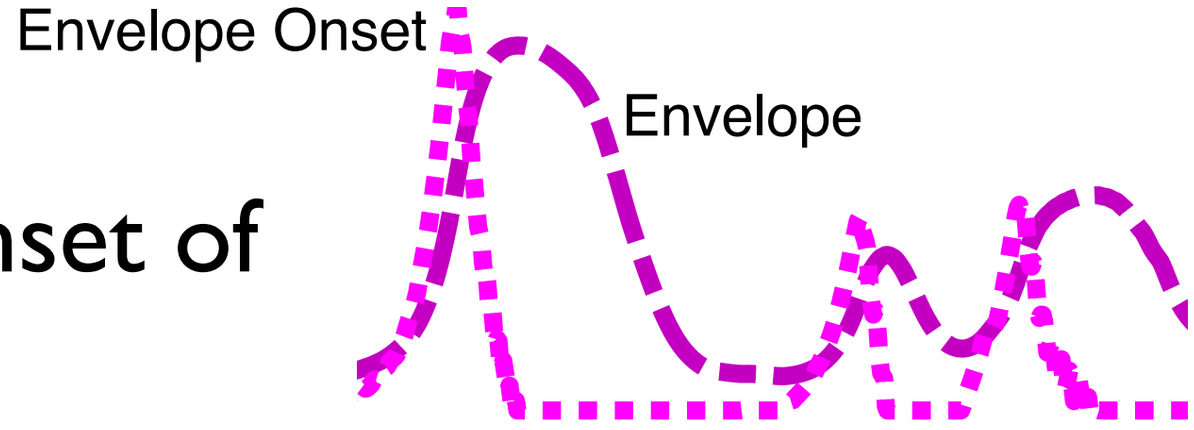
- Envelope vs. Onset of Envelope



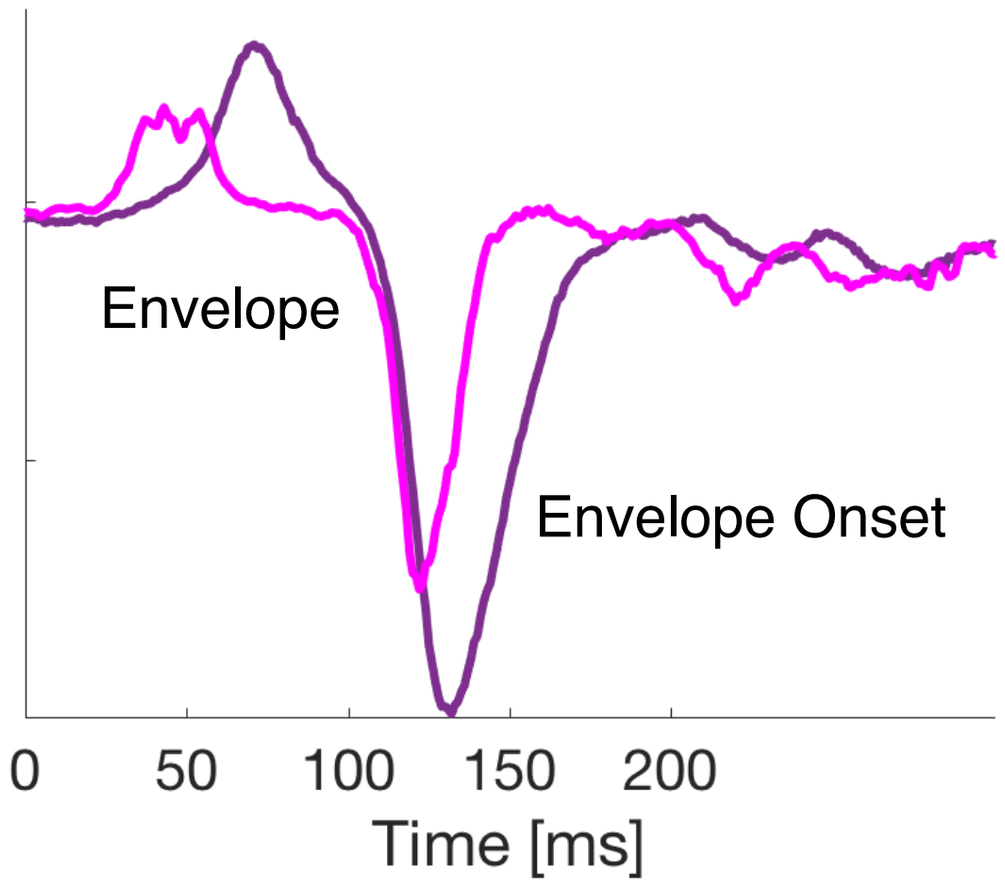
Cervantes Constantino et al., bioRxiv 168997

TRF: Depends on Stimulus Representation

- Envelope vs. Onset of Envelope



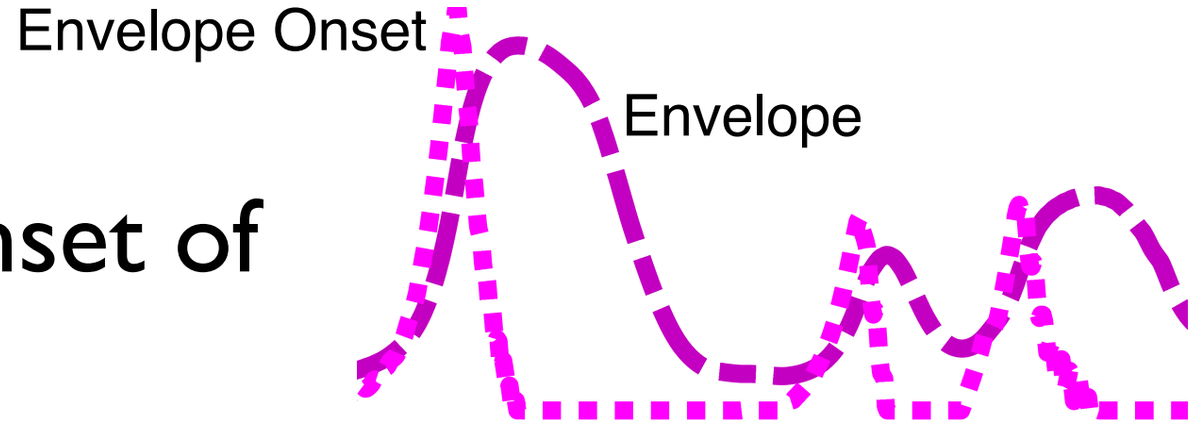
- Different TRFs



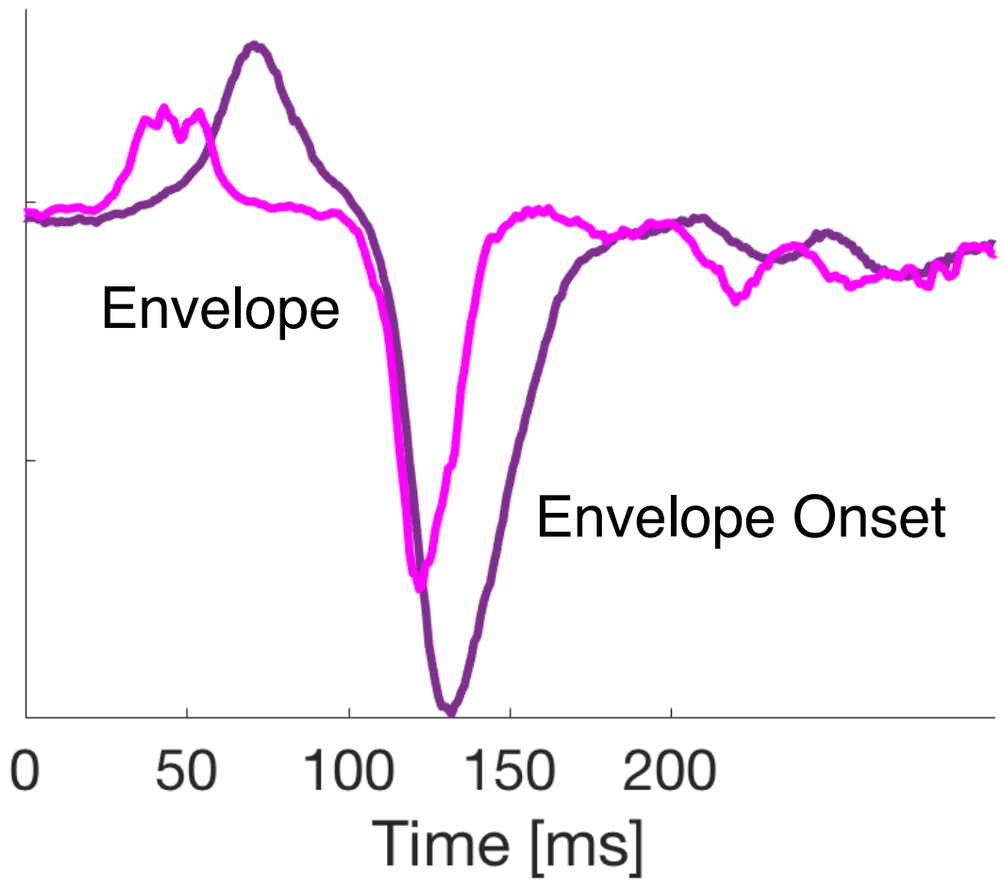
Cervantes Constantino et al., bioRxiv 168997

TRF: Depends on Stimulus Representation

- Envelope vs. Onset of Envelope

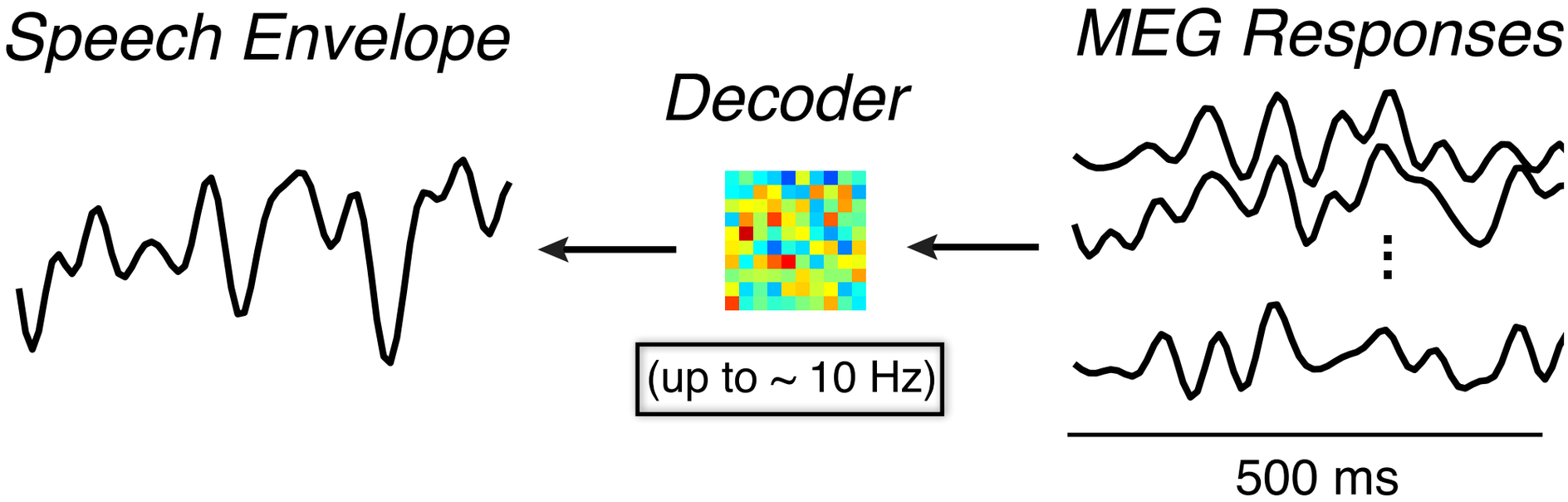


- Different TRFs
  - Latency (Shift in Stimulus Peak Timing)
  - Neural Processing

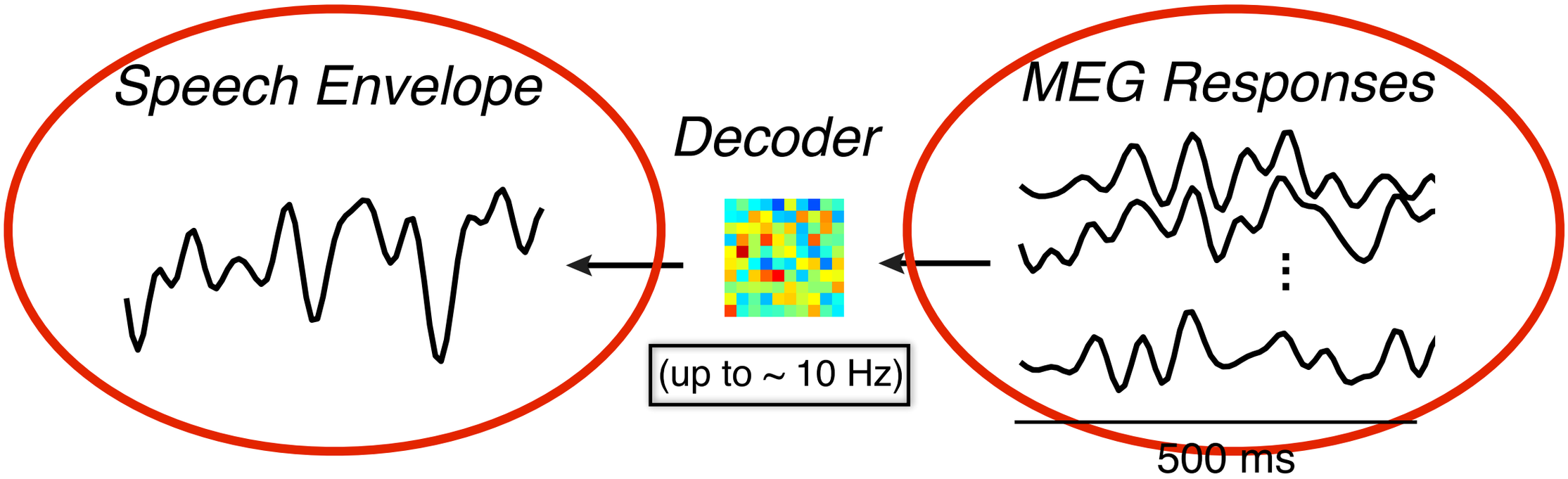


Cervantes Constantino et al., bioRxiv 168997

Neural Reconstruction of Speech Envelope

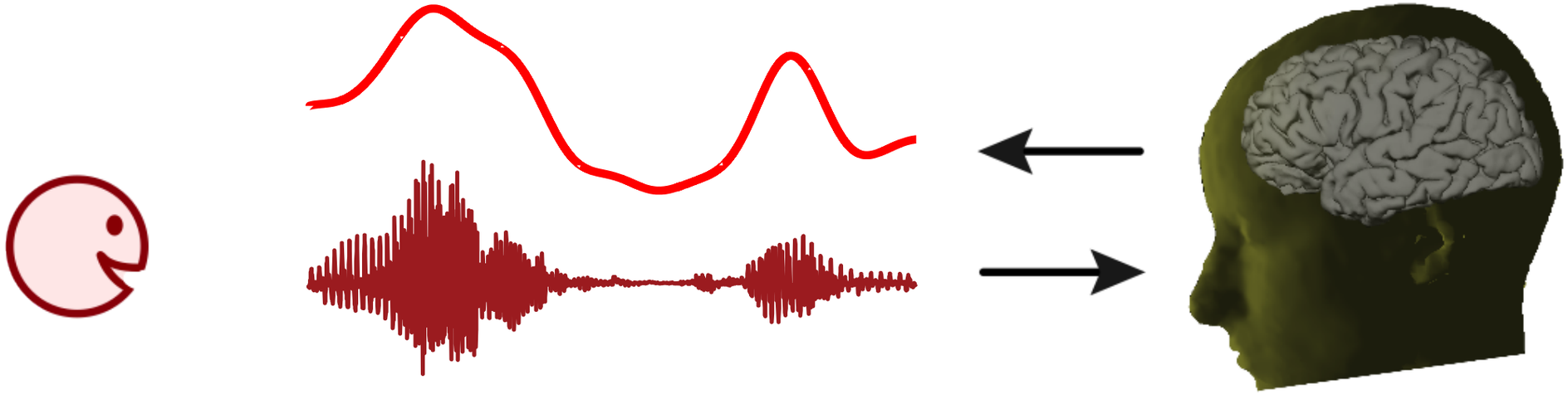


Ding & Simon, J Neurophysiol (2012)



Ding & Simon, J Neurophysiol (2012)

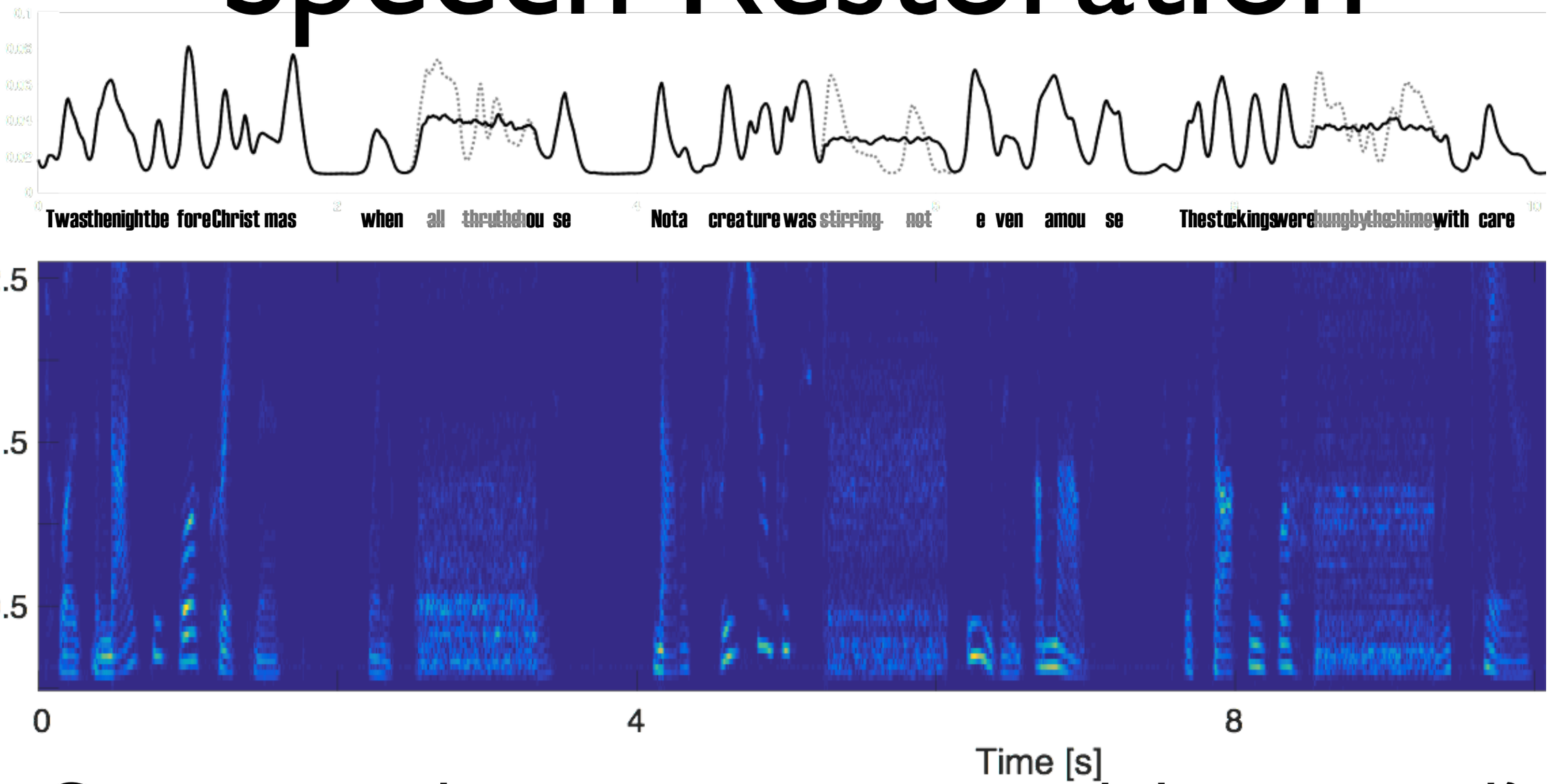
Neural Representation of Speech: Temporal



Outline

- Cortical Representations of Speech via MEG
- Missing speech / Familiar Speech
- Spatial Localization of Cortical Representations
  - Multiple Levels of Cortical Representations
- Representation Transition from Auditory/Phonetic → Lexical/Word-based
- Recent advances in Spatial Localization

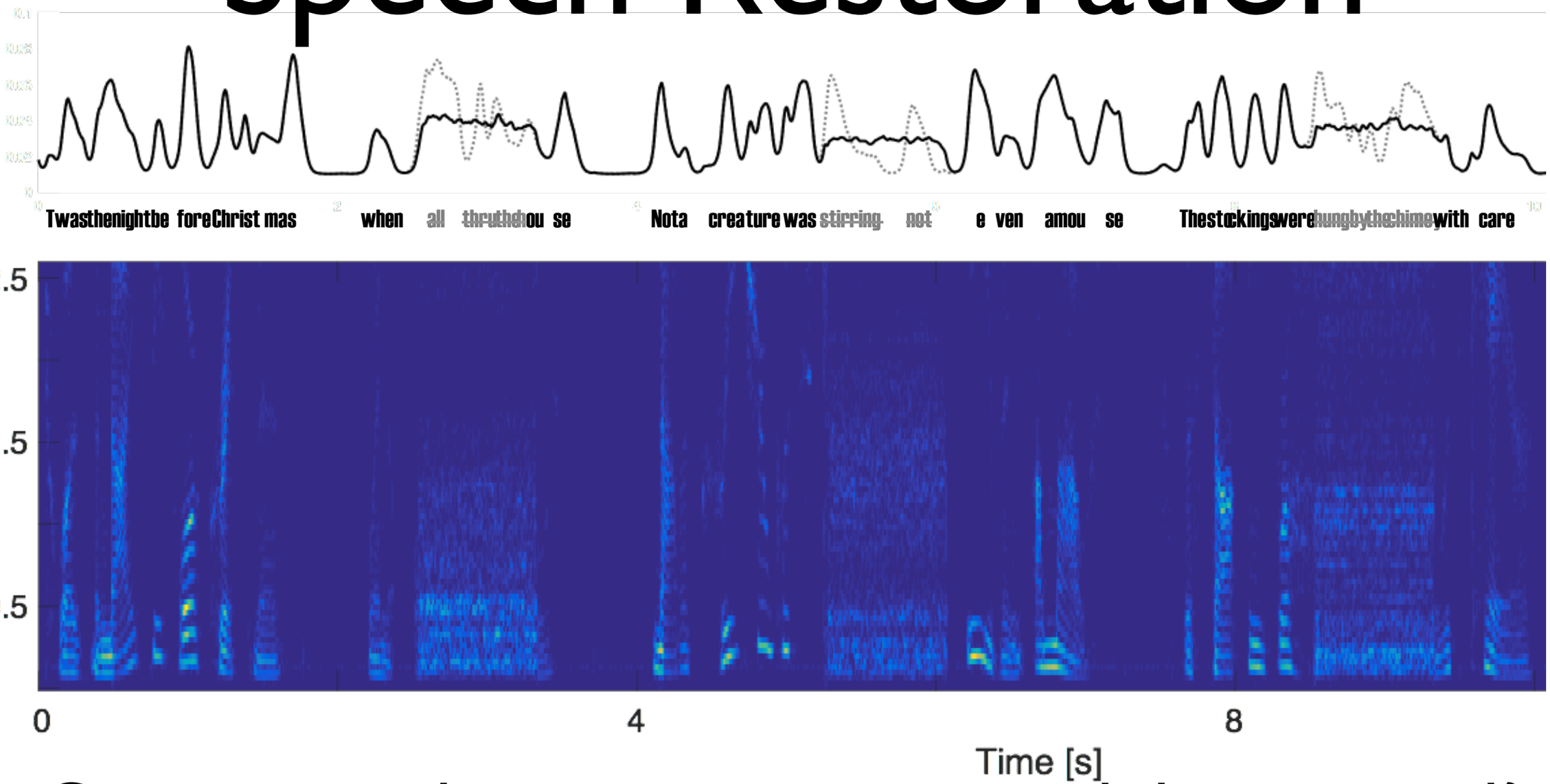
Speech Restoration



- Can sustained, non-stationary, speech be restored?
  - Might be aided by contextual knowledge/familiarity
  - Might be aided by strong rhythmicity

Cervantes Constantino & Simon, bioRxiv 251793

Speech Restoration



- Can sustained, non-stationary, speech be restored?
  - Might be aided by contextual knowledge/familiarity
  - Might be aided by strong rhythmicity

Cervantes Constantino & Simon, bioRxiv 251793

Speech Familiarity

Twas the night before Christmas, when all through the house
not a creature was stirring, not even a mouse.
The stockings were hung by the chimney with care,
in hopes that St. Nicholas soon would be there.

The children were nestled all snug in their beds,
while visions of sugar plums danced in their heads.
And Mama in her 'kerchief, and I in my cap,
had just settled our brains for a long winter's nap.

When out on the lawn there arose such a clatter,
I sprang from my bed to see what was the matter.
Away to the window I flew like a flash,
tore open the shutter, and threw up the sash.

The moon on the breast of the new-fallen snow
gave the lustre of midday to objects below,
when, what to my wondering eyes should appear,
but a miniature sleigh and eight tiny reindeer.

With a little old driver, so lively and quick,
I knew in a moment it must be St. Nick.
More rapid than eagles, his coursers they came,
and he whistled and shouted and called them by name.

"Now Dasher! Now Dancer! Now, Prancer and Vixen!
On, Comet! On, Cupid! On, Donner and Blitzen!
To the top of the porch! To the top of the wall!
Now dash away! Dash away! Dash away all!"

As dry leaves that before the wild hurricane fly,
when they meet with an obstacle, mount to the sky
so up to the house-top the coursers they flew,
with the sleigh full of toys, and St. Nicholas too.

And then, in a twinkling, I heard on the roof
the prancing and pawing of each little hoof.
As I drew in my head and was turning around,
down the chimney St. Nicholas came with a bound.

He was dressed all in fur, from his head to his foot,
and his clothes were all tarnished with ashes and soot.
A bundle of toys he had flung on his back,
and he looked like a peddler just opening his pack.

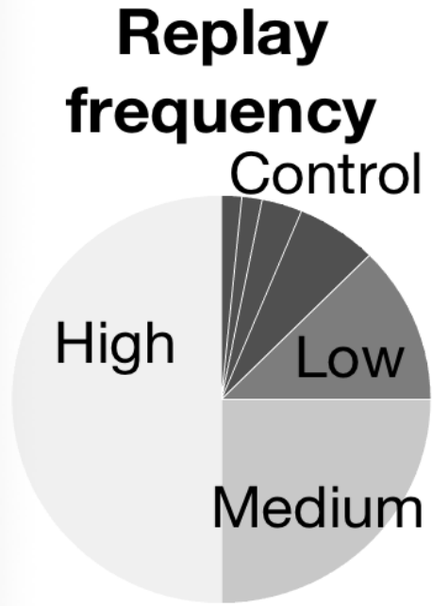
His eyes--how they twinkled! His dimples, how merry!
His cheeks were like roses, his nose like a cherry!
His droll little mouth was drawn up like a bow,
and the beard on his chin was as white as the snow.

The stump of a pipe he held tight in his teeth,
and the smoke it encircled his head like a wreath.
He had a broad face and a little round belly,
that shook when he laughed, like a bowl full of jelly.

He was chubby and plump, a right jolly old elf,
and I laughed when I saw him, in spite of myself.
A wink of his eye and a twist of his head
soon gave me to know I had nothing to dread.

He spoke not a word, but went straight to his work,
and filled all the stockings, then turned with a jerk.
And laying his finger aside of his nose,
and giving a nod, up the chimney he rose.

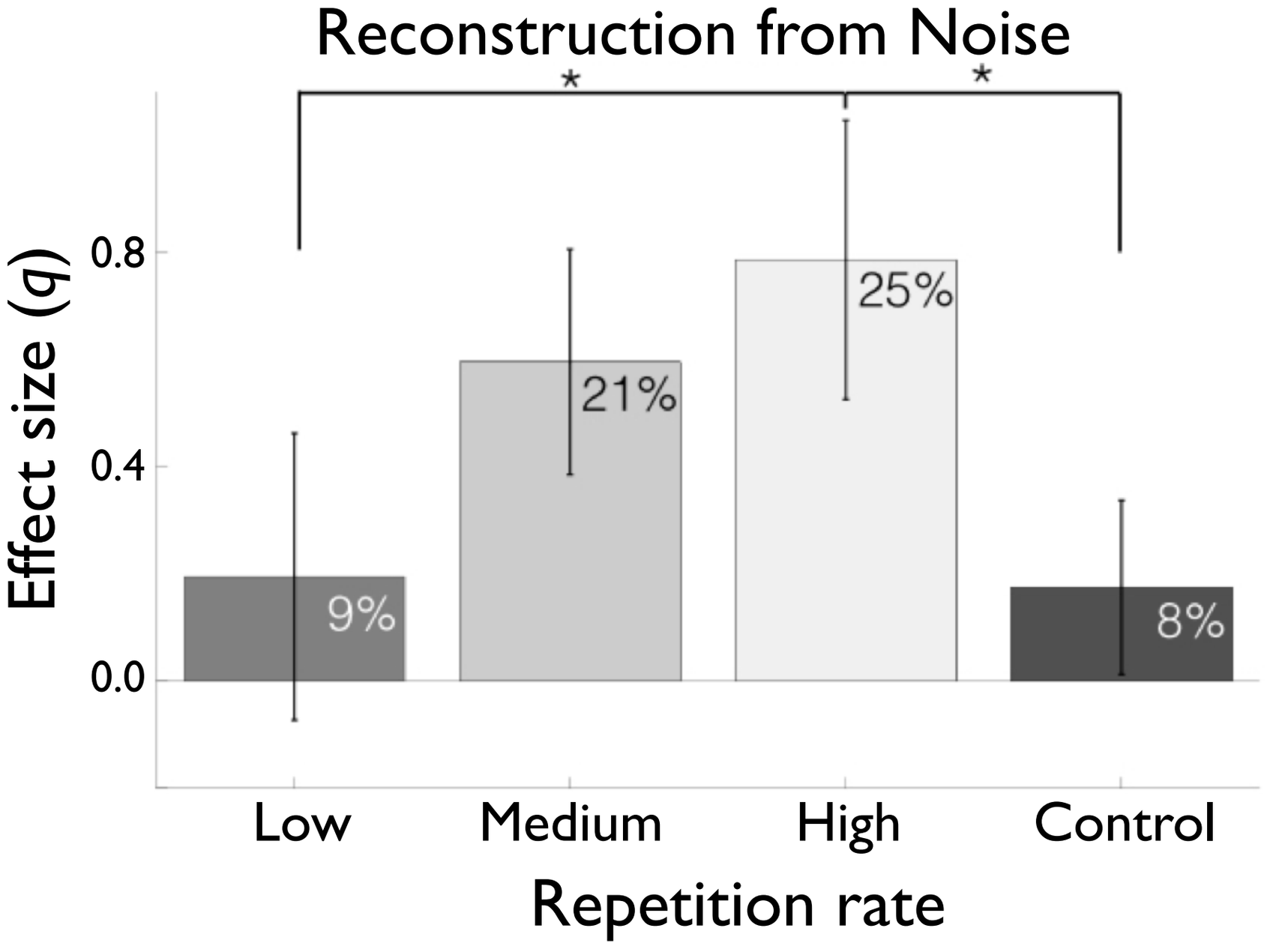
He sprang to his sleigh, to his team gave a whistle,
And away they all flew like the down of a thistle.
But I heard him exclaim, 'ere he drove out of sight,
"Happy Christmas to all, and to all a good night!"



- Hypothesis: contextual knowledge of missing speech can be controlled by exposure to the speech

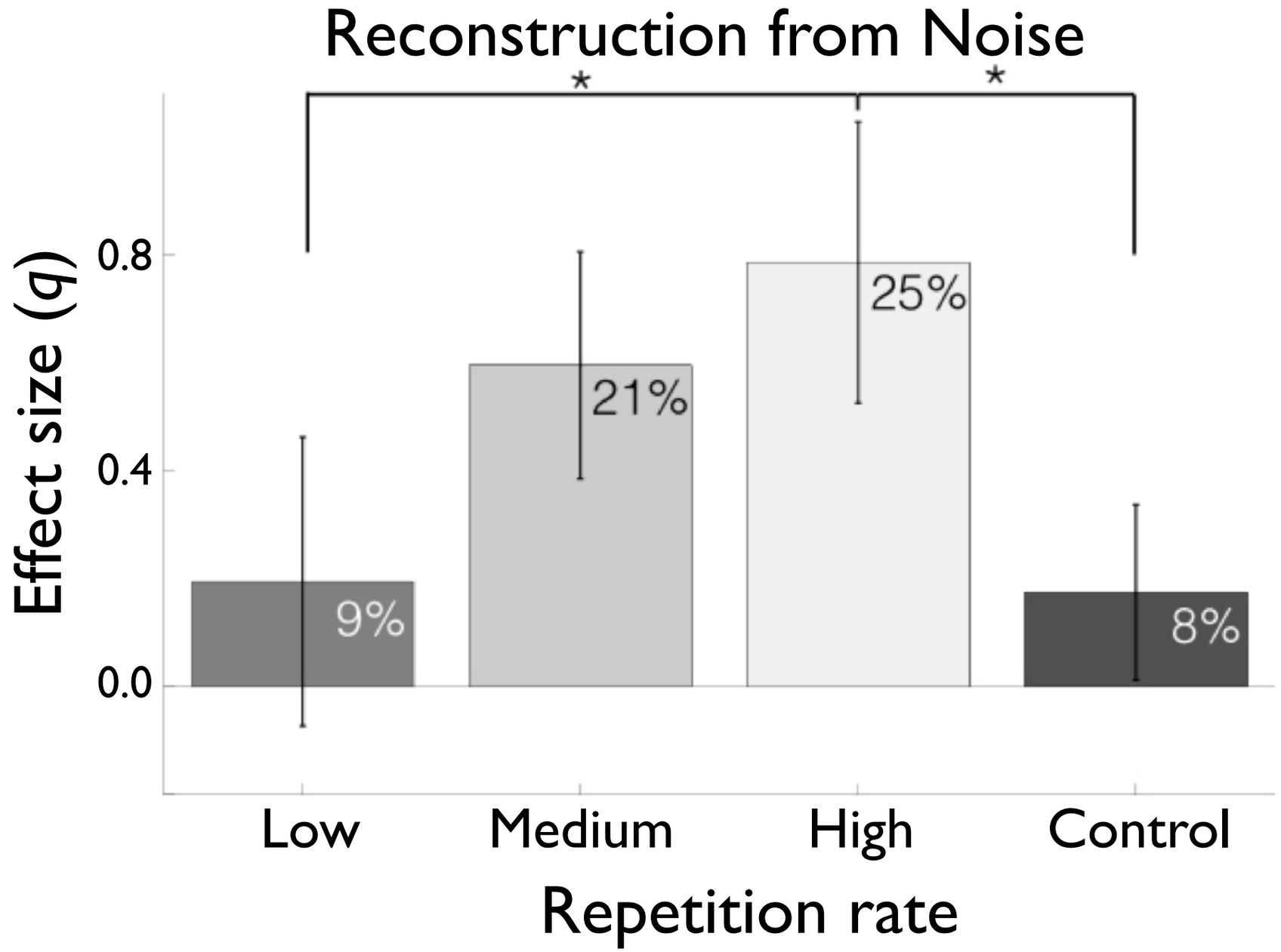
Cervantes Constantino & Simon, bioRxiv 251793

Missing Speech Reconstruction



Cervantes Constantino & Simon, bioRxiv 251793

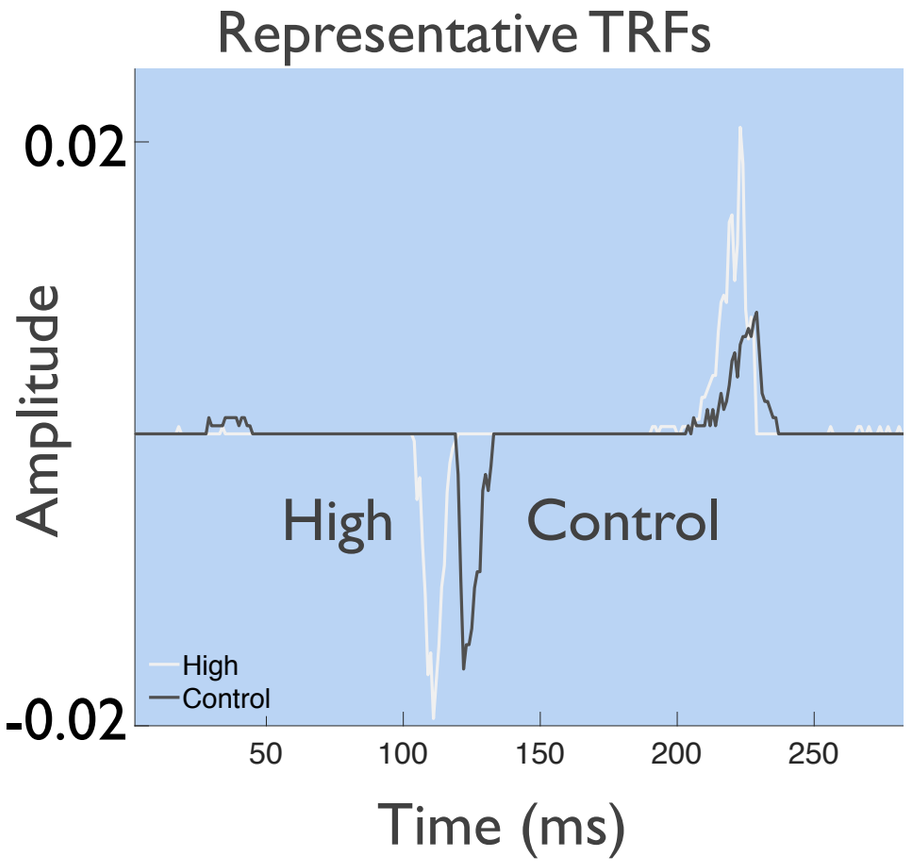
Missing Speech Reconstruction

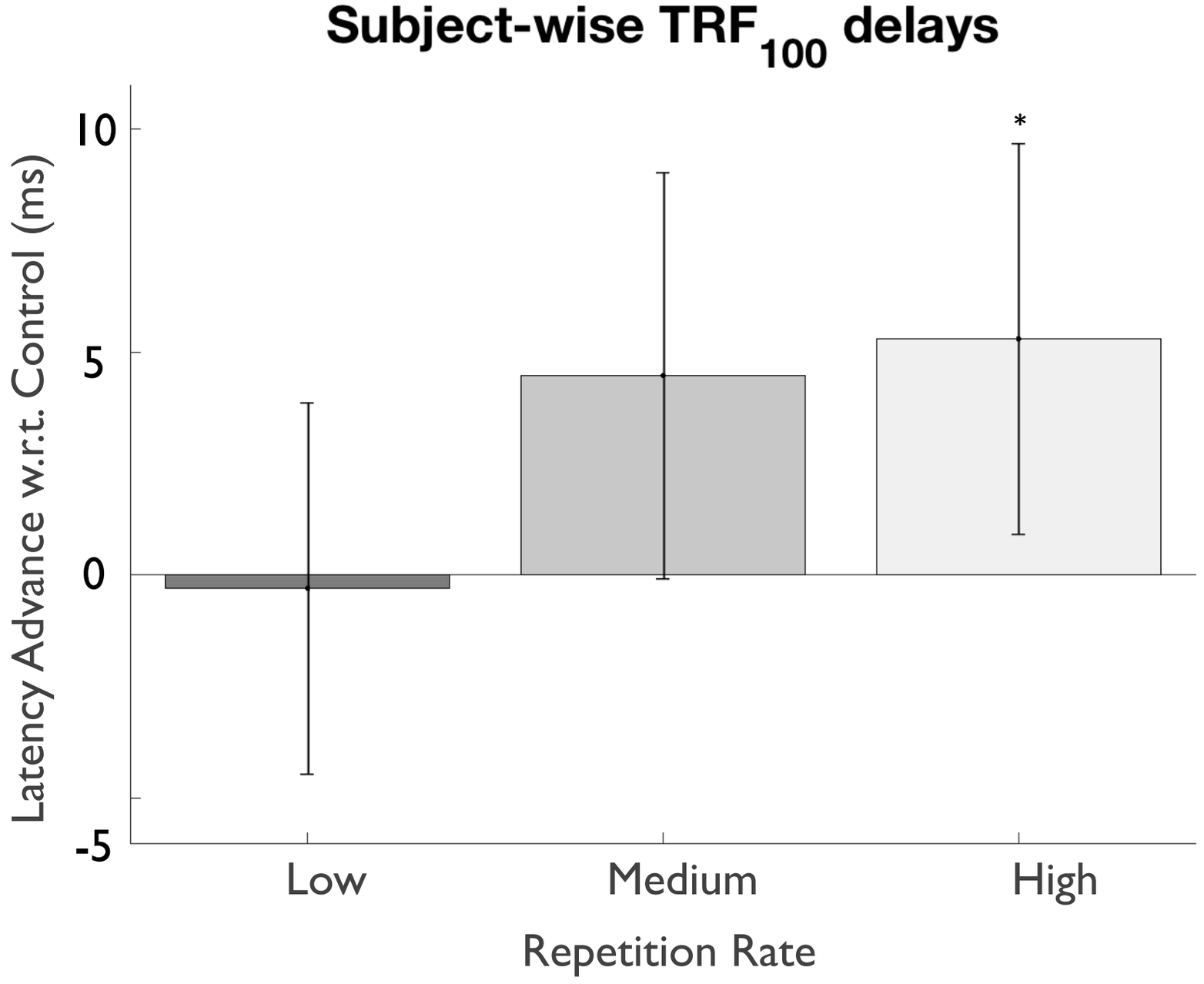


- Decoding of the ***missing*** speech token improves with prior experience
- Performance is a considerable fraction of that for clean speech

Cervantes Constantino & Simon, bioRxiv 251793

Speech Anticipation





- Prior experience speeds subsequent responses

Cervantes Constantino & Simon, bioRxiv 251793

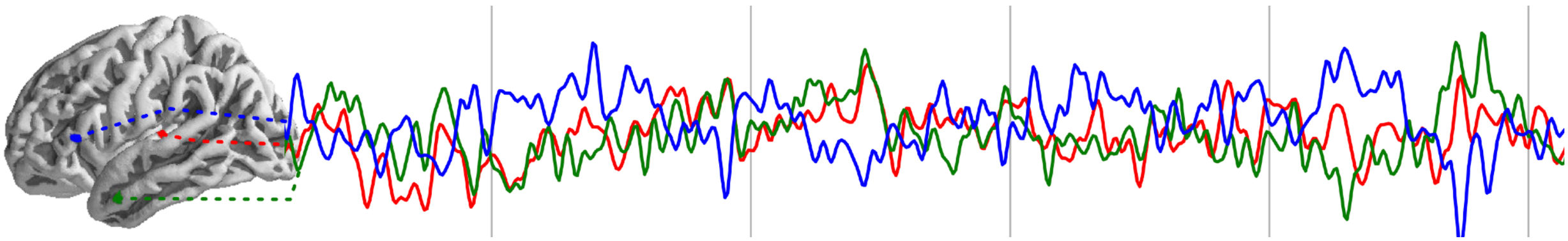
Outline

- Cortical Representations of Speech via MEG
- Missing speech / Familiar Speech
- Spatial Localization of Cortical Representations
  - Multiple Levels of Cortical Representations
- Representation Transition from Auditory/Phonetic → Lexical/Word-based
- Recent advances in Spatial Localization

TRF: One Per Response

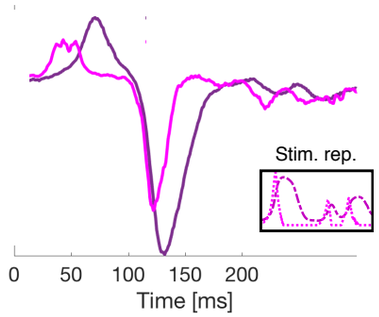
- But multiple channels
- Option: TRF for each channel
- Another Option: TRF for each *component*
- Component = linear superposition of channels
- Other linear superpositions of channels?

TRF: One Per Neural Source



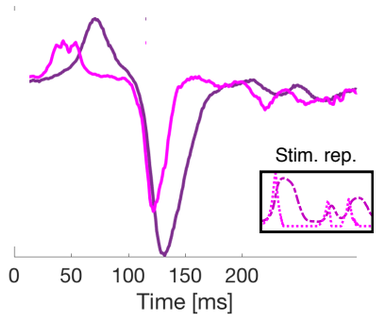
- Another option: TRF for each *neural source*
  - Each neural source current maps linearly into a magnetic field distribution (via the "Lead Field" matrix)
  - Every potential neural source is a linear superposition of channels

TRF: One Per Representation



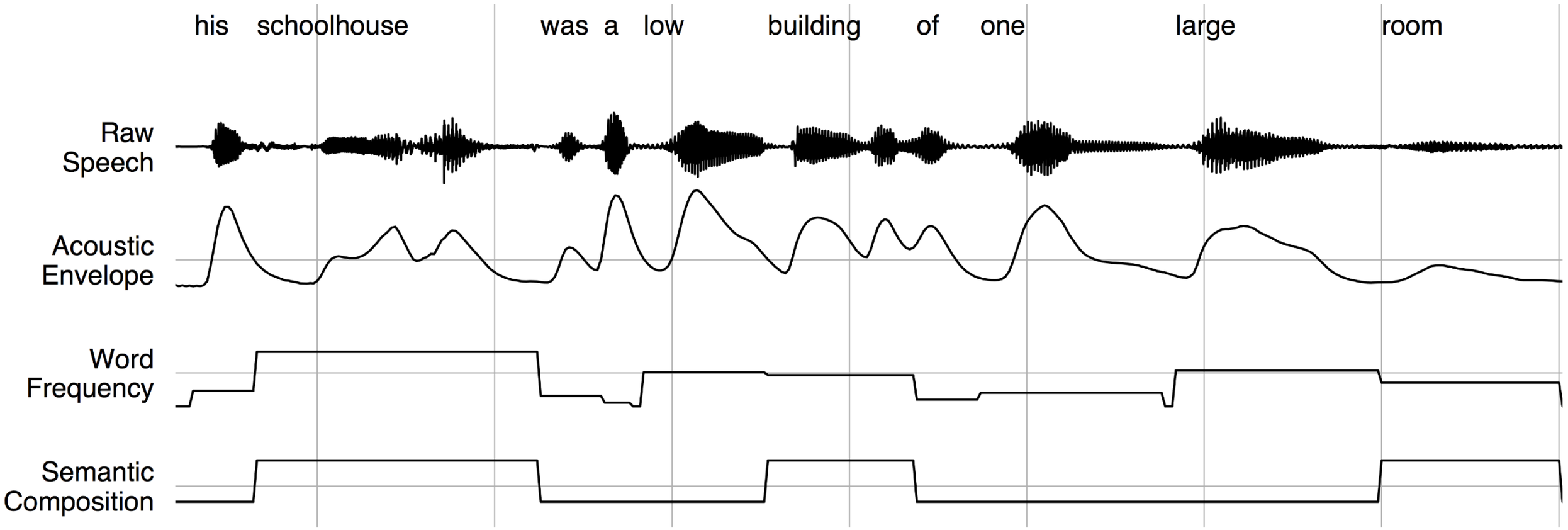
TRF: One Per Representation

- Different cortical areas should encode different representations



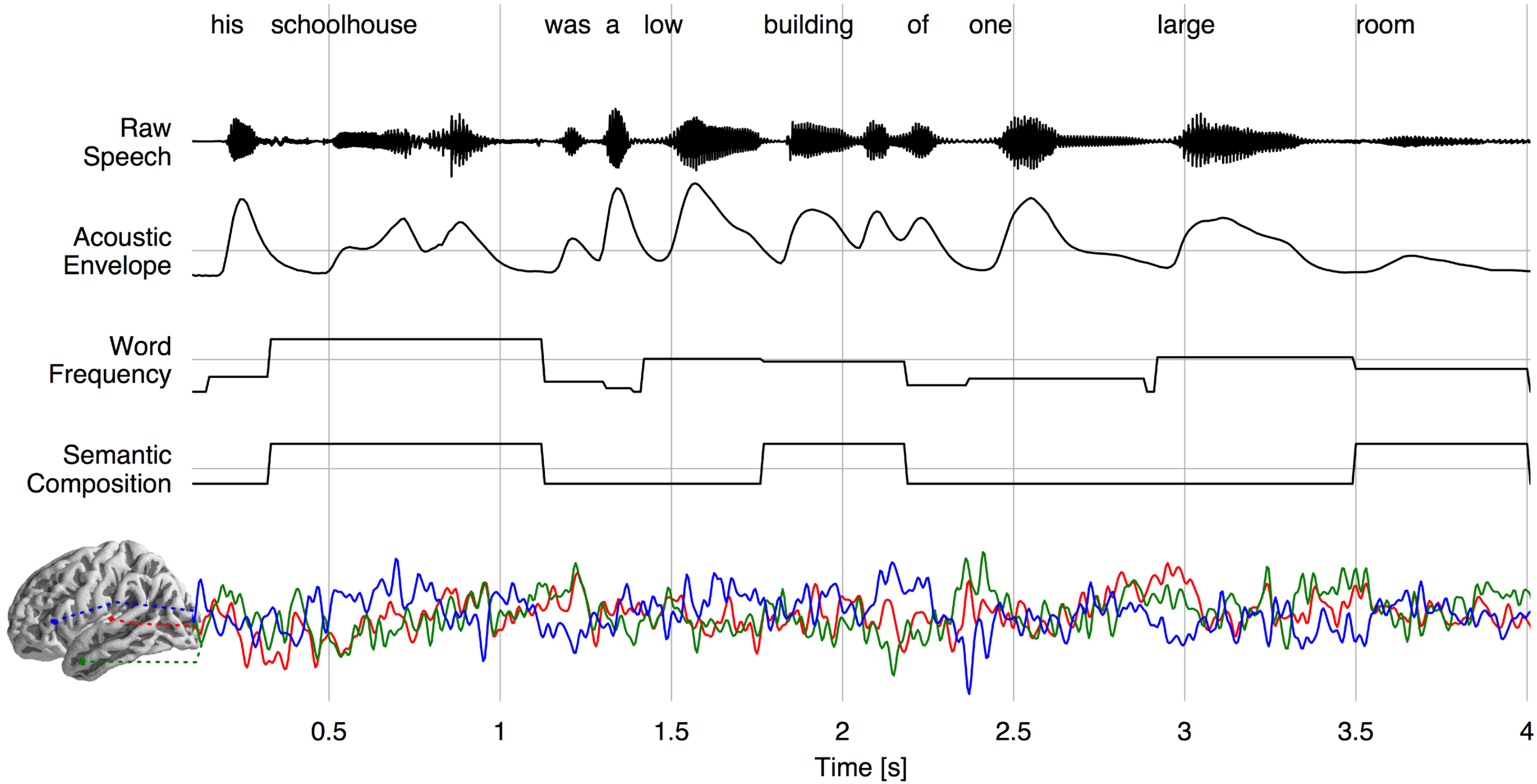
- When considering entire brain, should consider non-auditory representations also, e.g.,
  - Lexical (word-based) variables
  - Semantic variables
  - Other…

TRF: One Per Representation



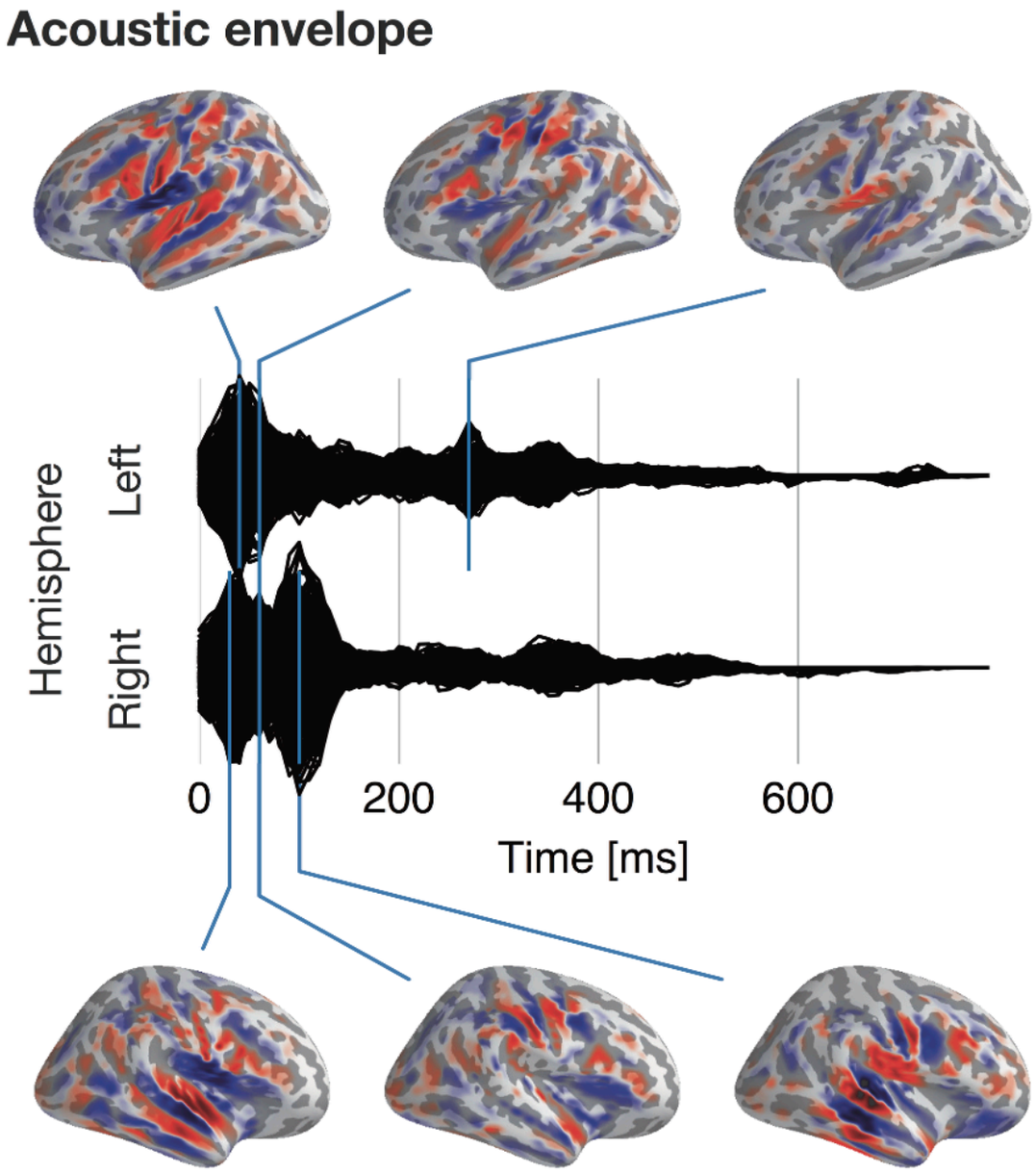
Brodbeck & Simon, NeuroImage (2017)

TRF: One Per Representation Per Neural Source



Brodbeck & Simon, NeuroImage (2017)

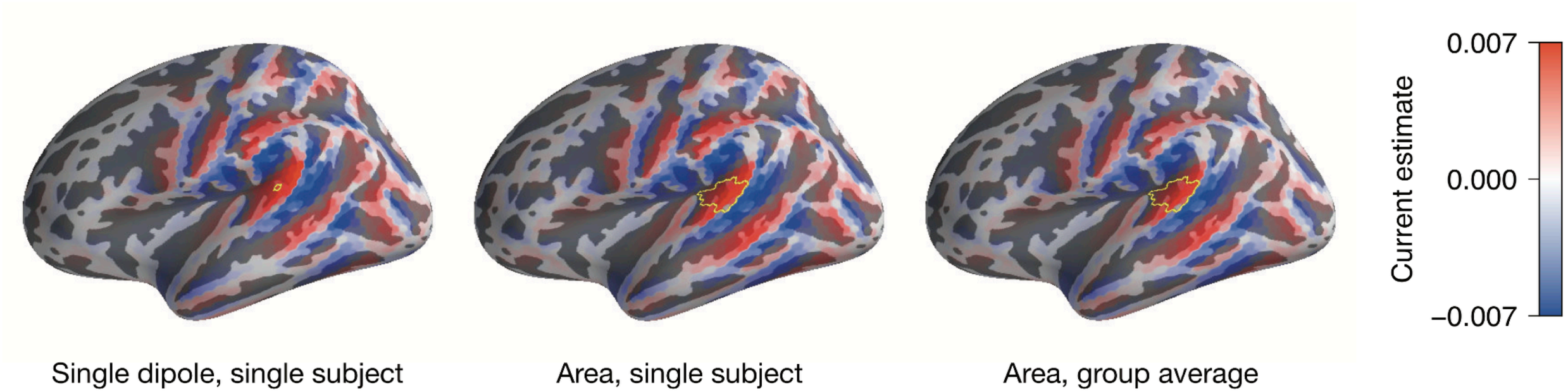
TRF: One Per Neural Source



- Bilateral early response
  - auditory cortex (~30 ms)
  - sensorimotor parietal and frontal cortices (~50 ms)
- Right-lateralized later response
  - auditory cortex (~100 ms)

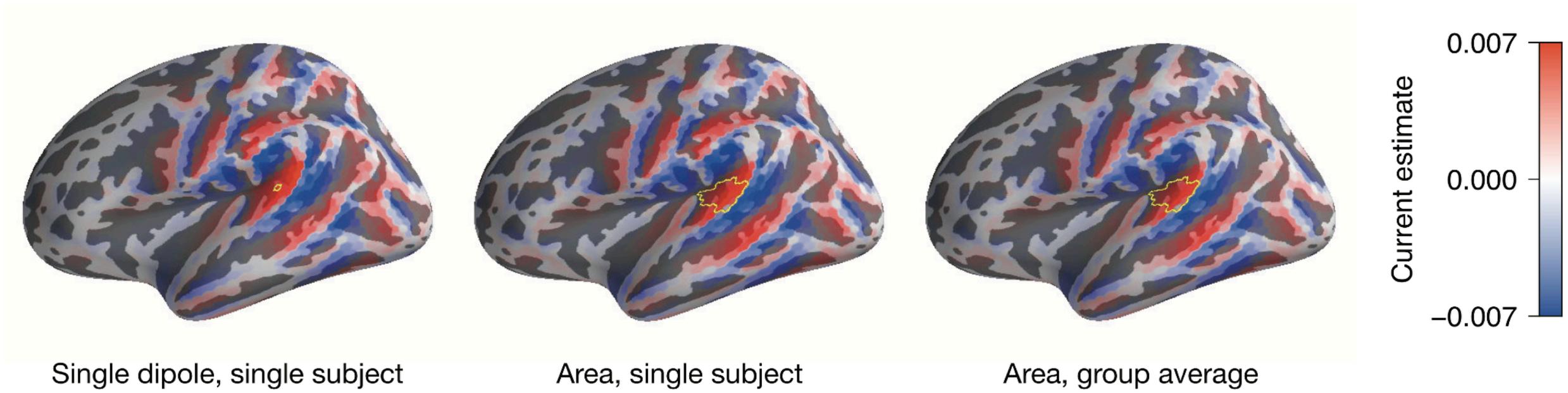
Brodbeck & Simon, NeuroImage (2017)

MEG Source Spread



Brodbeck & Simon, NeuroImage (2017)

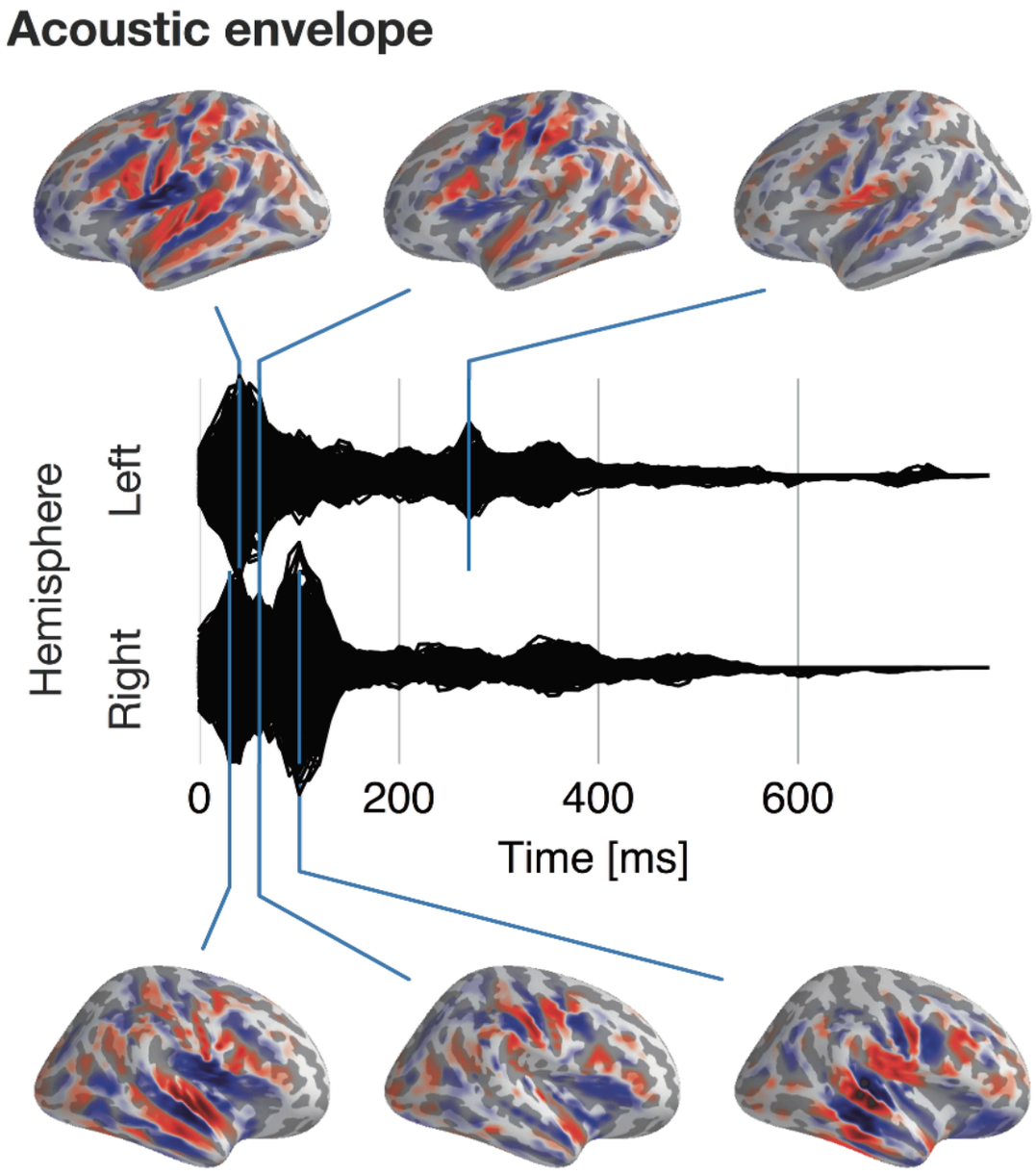
MEG Source Spread



- Apparent Broad Areas of Activation: Artifact
- Due to *Point Spread Function*
- Difficult to Avoid for Single Source Analysis

Brodbeck & Simon, NeuroImage (2017)

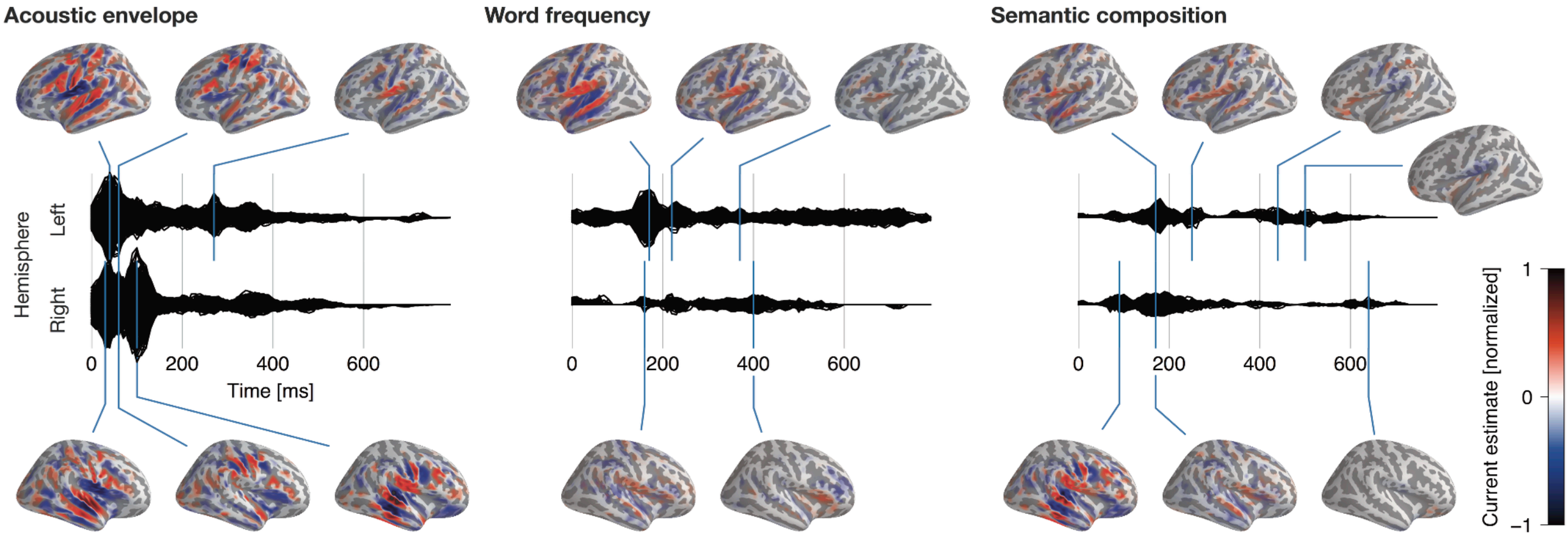
TRF: One Per Neural Source



- Bilateral early response
  - auditory cortex (~30 ms)
  - sensorimotor parietal and frontal cortices (~50 ms)
- Right-lateralized later response
  - auditory cortex (~100 ms)

Brodbeck & Simon, NeuroImage (2017)

TRF: One Per Representation Per Neural Source



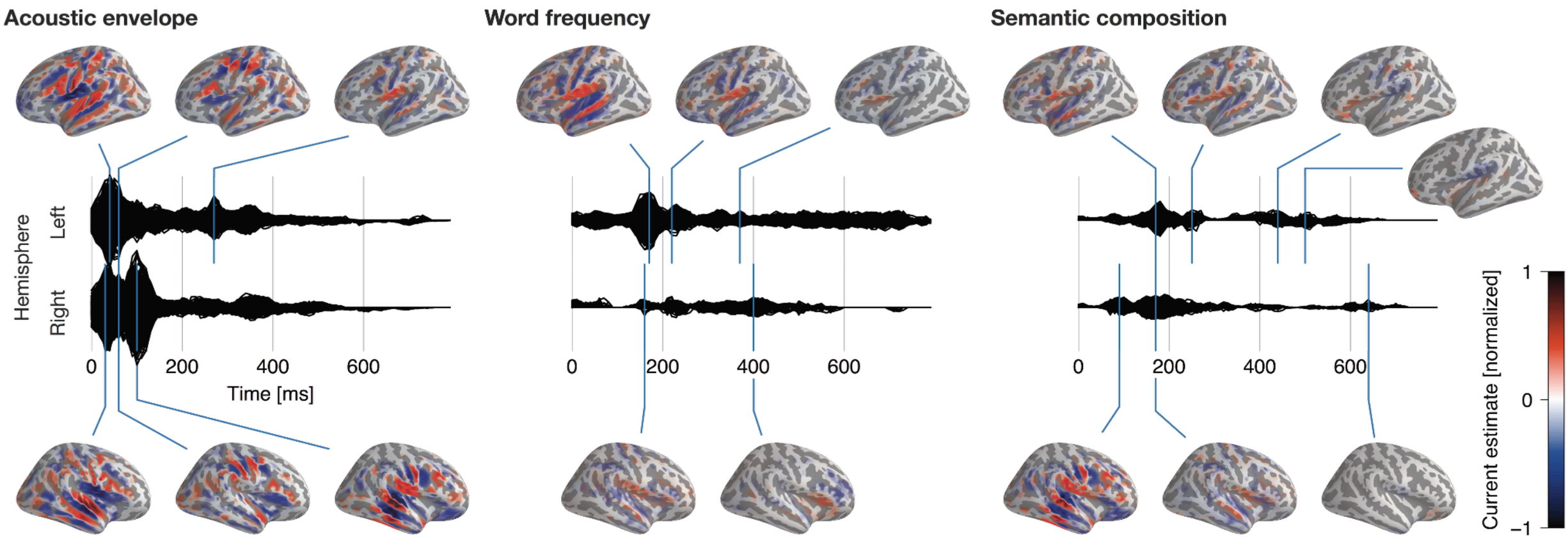
- Bilateral early response
  - auditory cortex (~30 ms)
  - sensorimotor parietal and frontal cortices (~50 ms)
- Right-lateralized later response
  - auditory cortex (~100 ms)

- Strong left-lateralized response in auditory cortex (~170 ms)
- Later, weaker bilateral frontal response

- Left hemisphere:
  - temporal progression from anterior temporal lobe to inferior frontal gyrus activation
- Right-hemisphere
  - similarly localized, temporally more diffuse

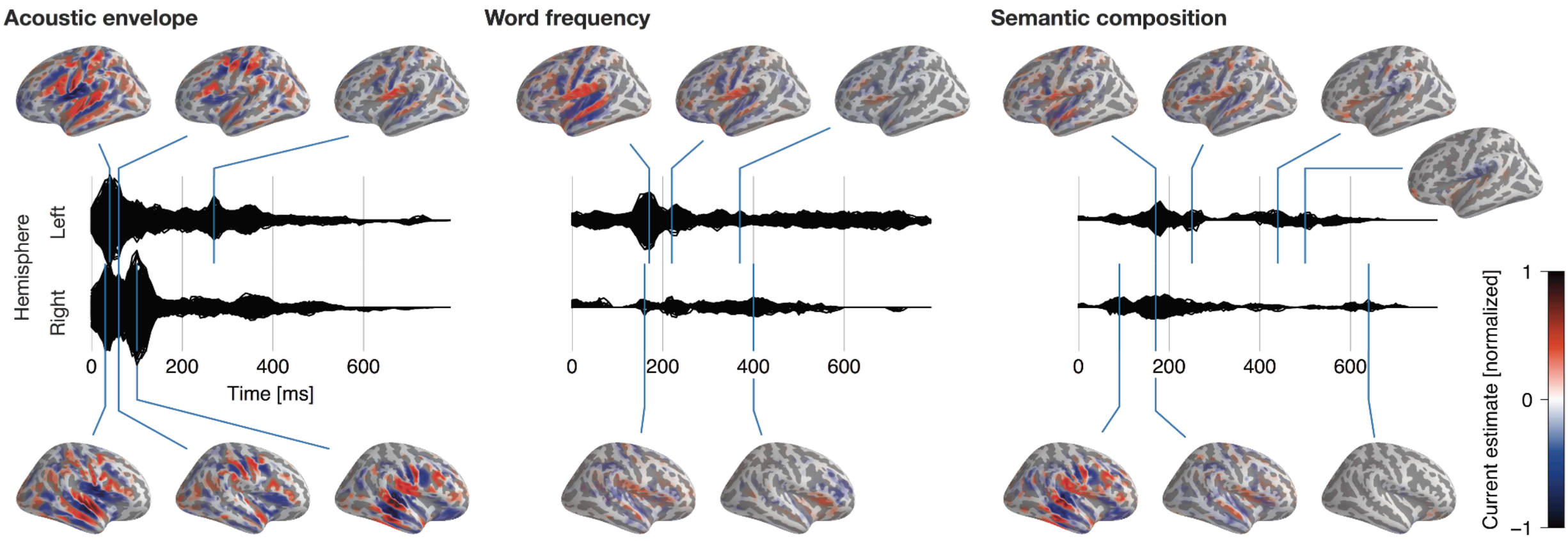
Brodbeck & Simon, NeuroImage (2017)

Too Much Data
Not Enough Neuroscience?



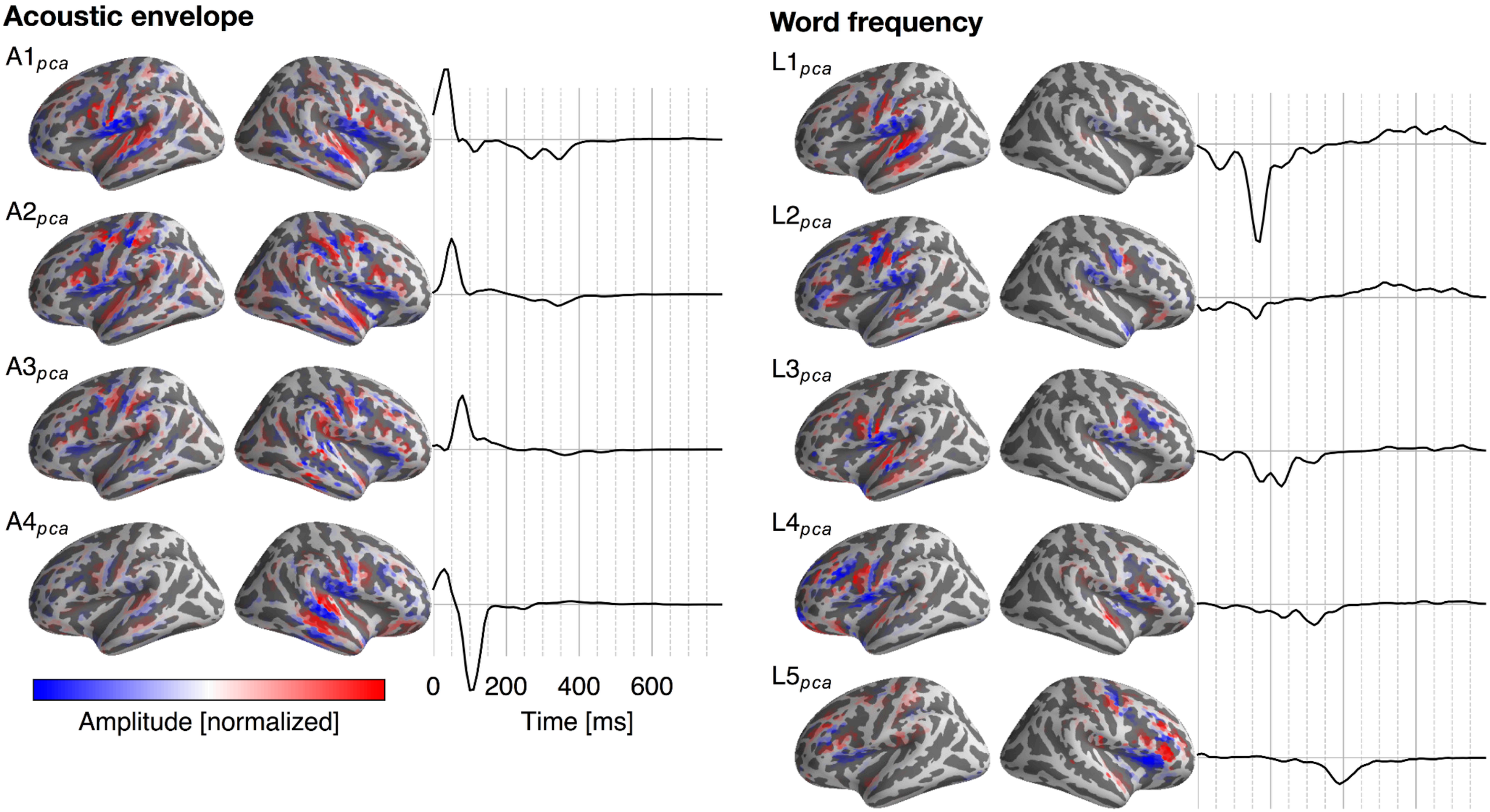
- Common Neural Response Patterns?
- Separable Cortical Processing Stages?

Too Much Data
Not Enough Neuroscience?



- Employ Cluster Analysis
- Estimate all TRF sources simultaneously(?)

TRFs: Per Processing Stage



Brodbeck & Simon, NeuroImage (2017)

- Cluster Analysis Helps

Outline

- Cortical Representations of Speech via MEG
- Missing speech / Familiar Speech
- Spatial Localization of Cortical Representations
  - Multiple Levels of Cortical Representations
- Representation Transition from Auditory/Phonetic → Lexical/Word-based
- Recent advances in Spatial Localization

Auditory/Phonetic Transition to Lexical/Word Representations

- Transition from acoustic/phonetic representations to symbolic, lexical representations?
  - Investigate: Responses representing *integration of phonetic information* towards *word identification*

Brodbeck & Simon, bioRxiv 326785

Stimulus Representation: Auditory vs. Phonemic Context within Words

Acoustic Envelope (8 bands)

Acoustic Onset (8 bands)

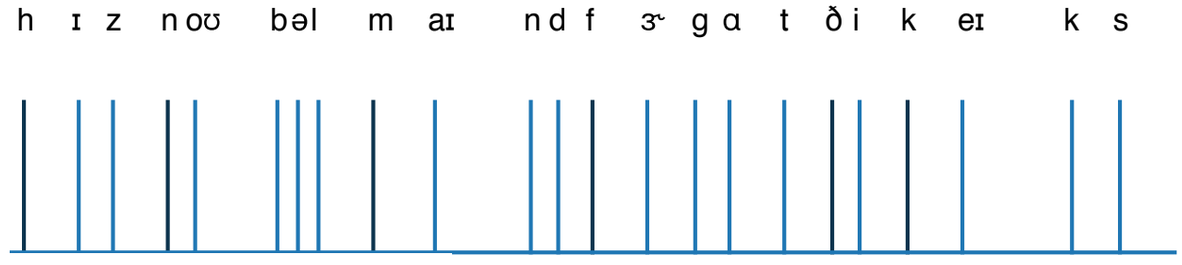
Different Stimulus Representations compete to explain variance

Brodbeck & Simon, bioRxiv 326785

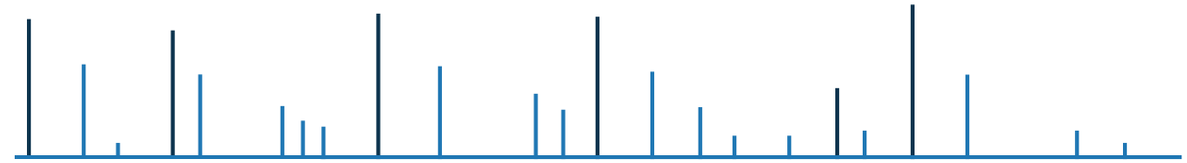
Stimulus Representation: Auditory vs. Phonemic Context within Words

Acoustic Envelope (8 bands)

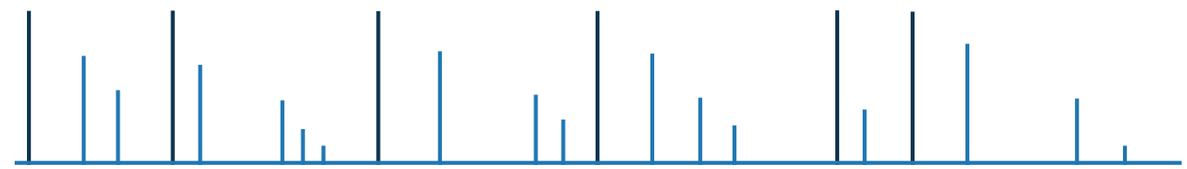
Acoustic Onset (8 bands)



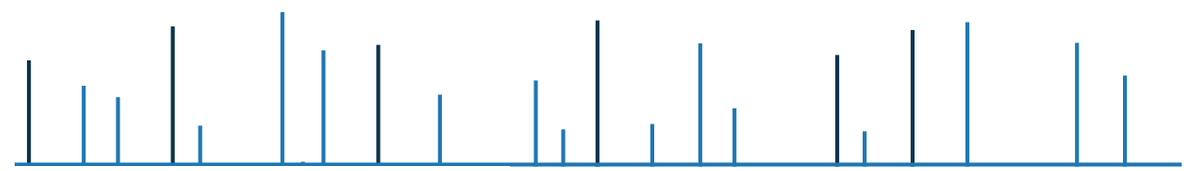
Phonemic Onset



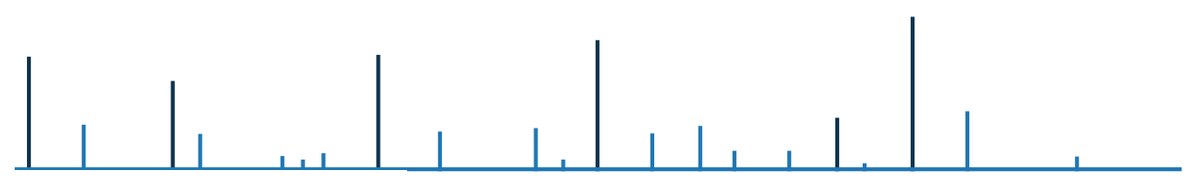
Cohort Size



Cohort Reduction



Phoneme Surprisal

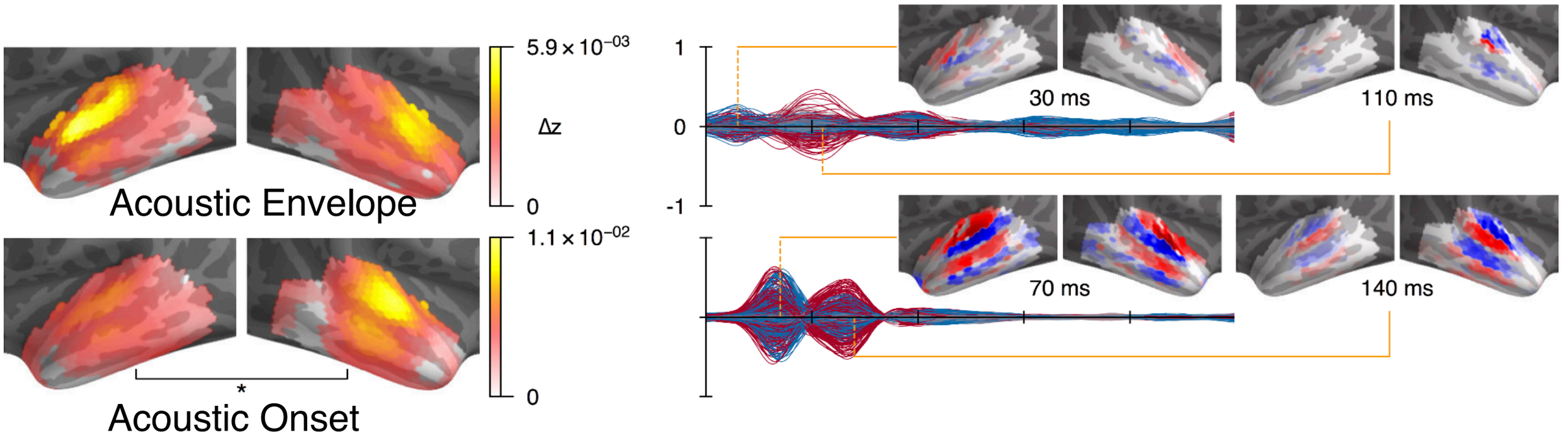


Cohort Entropy

Different Stimulus Representations compete to explain variance

Brodbeck & Simon, bioRxiv 326785

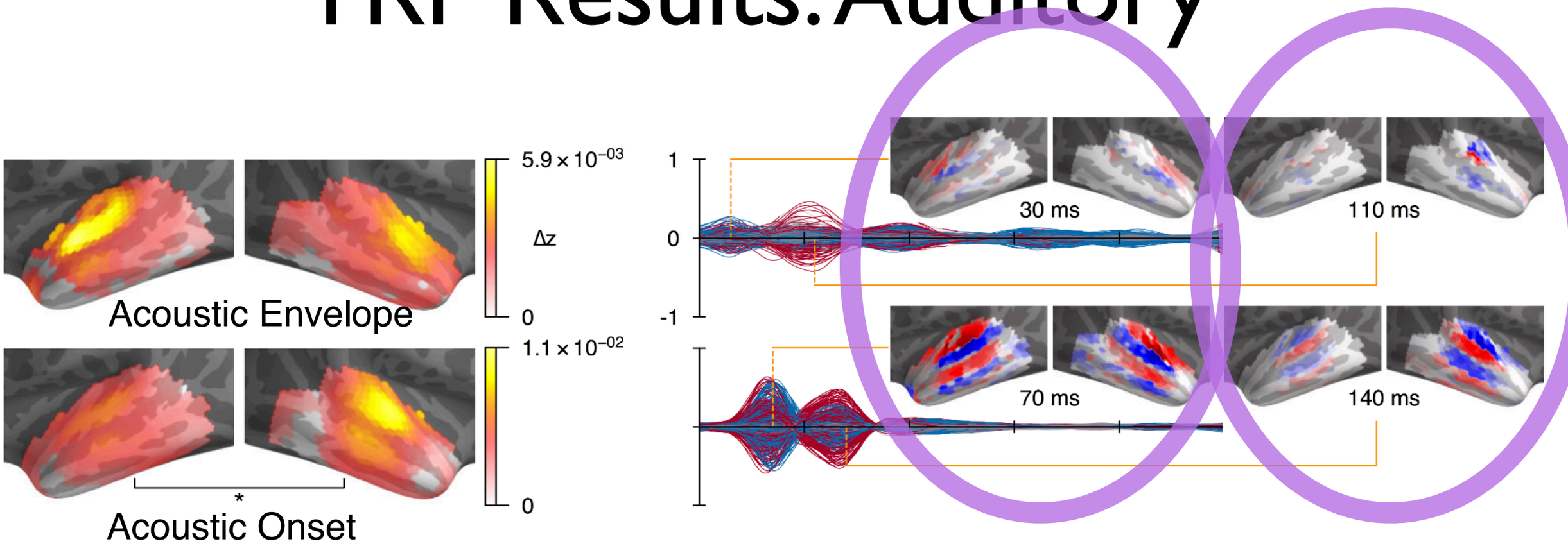
TRF Results: Auditory



- Strong, Bilateral Auditory Representations

Brodbeck & Simon, bioRxiv 326785

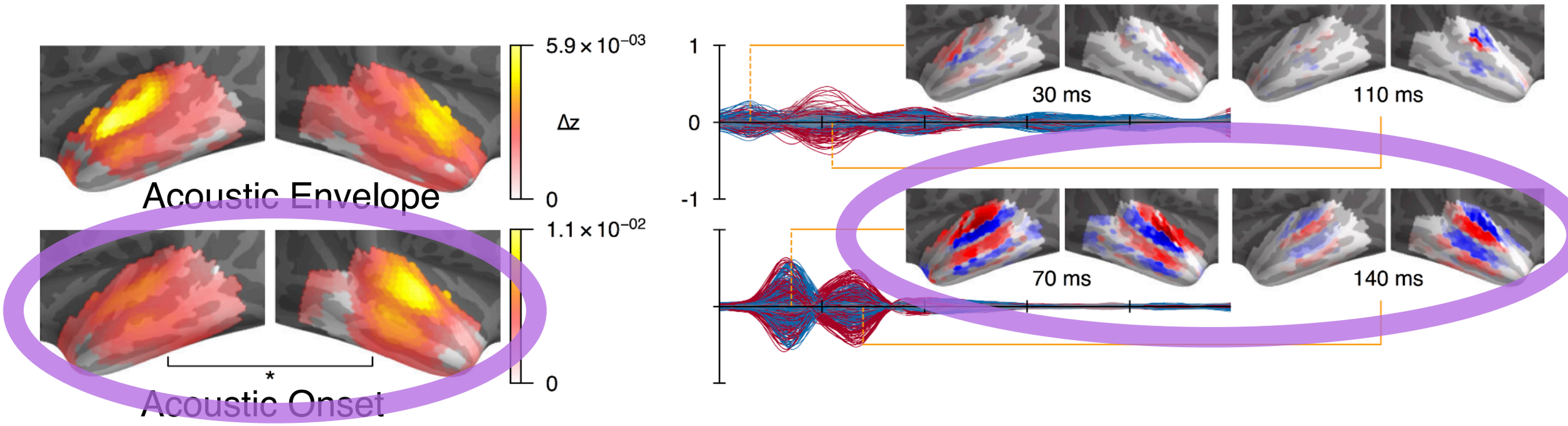
TRF Results: Auditory



- Strong, Bilateral Auditory Representations
- Acoustic Onset stronger

Brodbeck & Simon, bioRxiv 326785

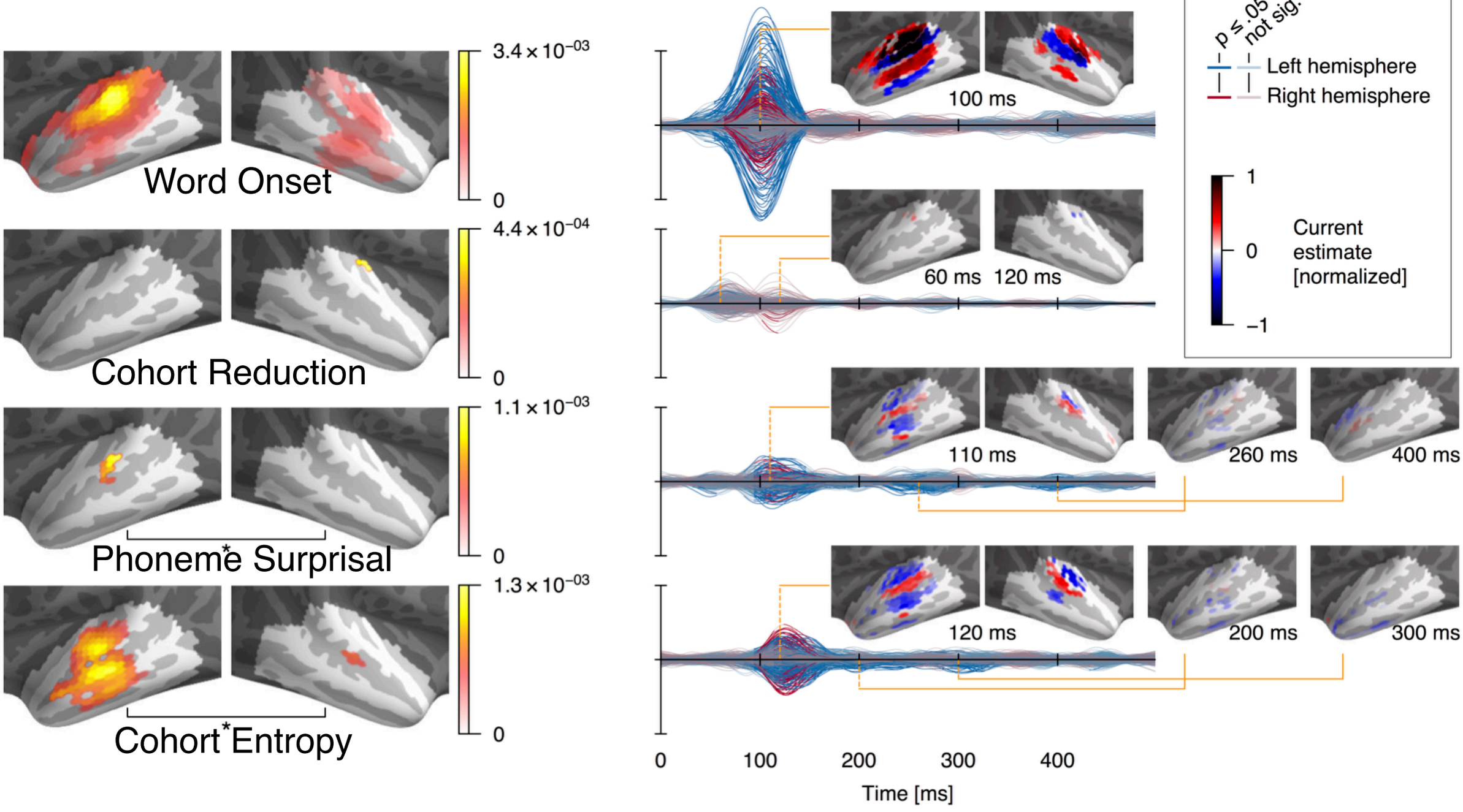
TRF Results: Auditory



- Strong, Bilateral Auditory Representations
- Acoustic Onset stronger
- Acoustic Onset Right-Lateralized

Brodbeck & Simon, bioRxiv 326785

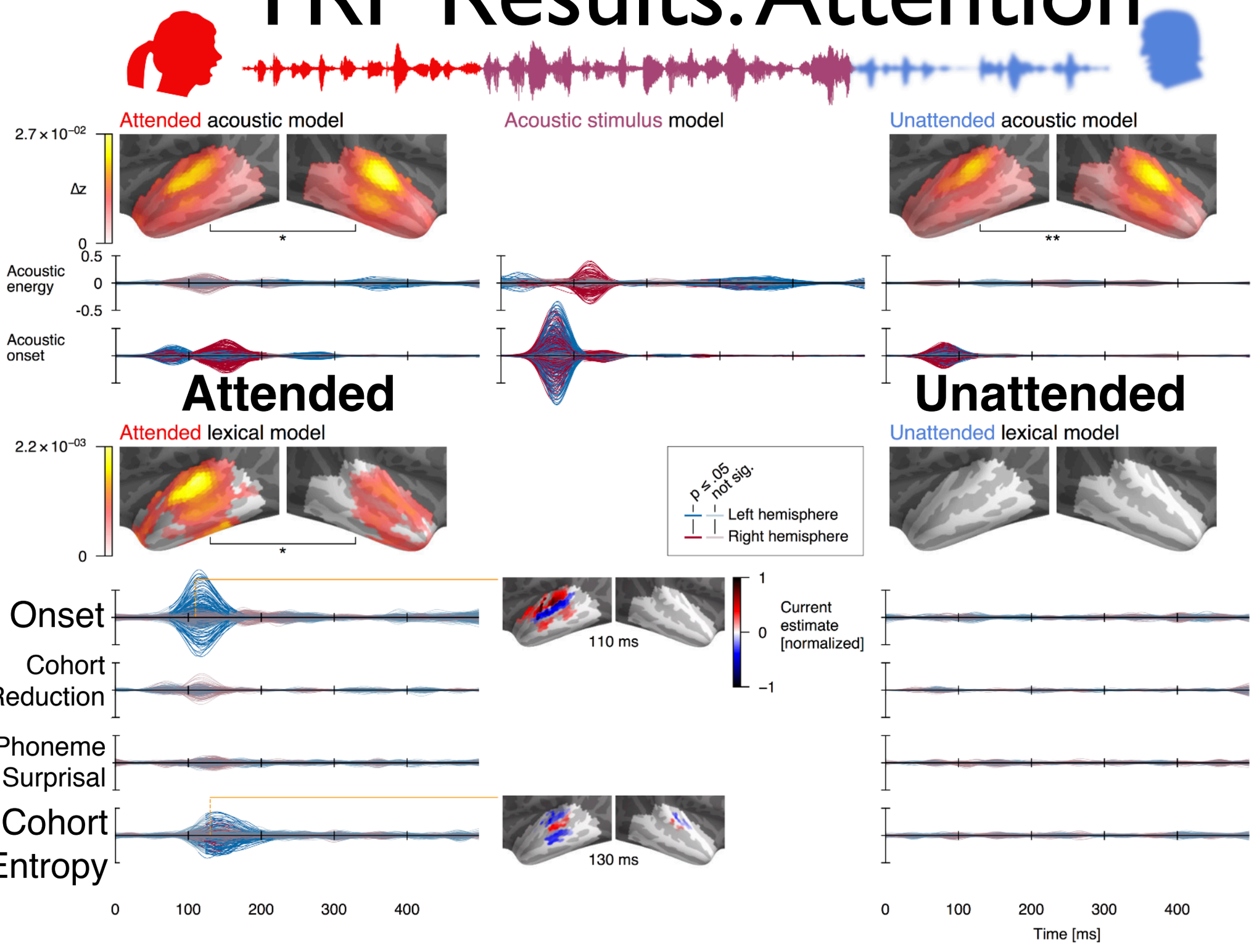
TRF Results: Lexical



- Robust, Left-Lateralized, starting at ~110 ms

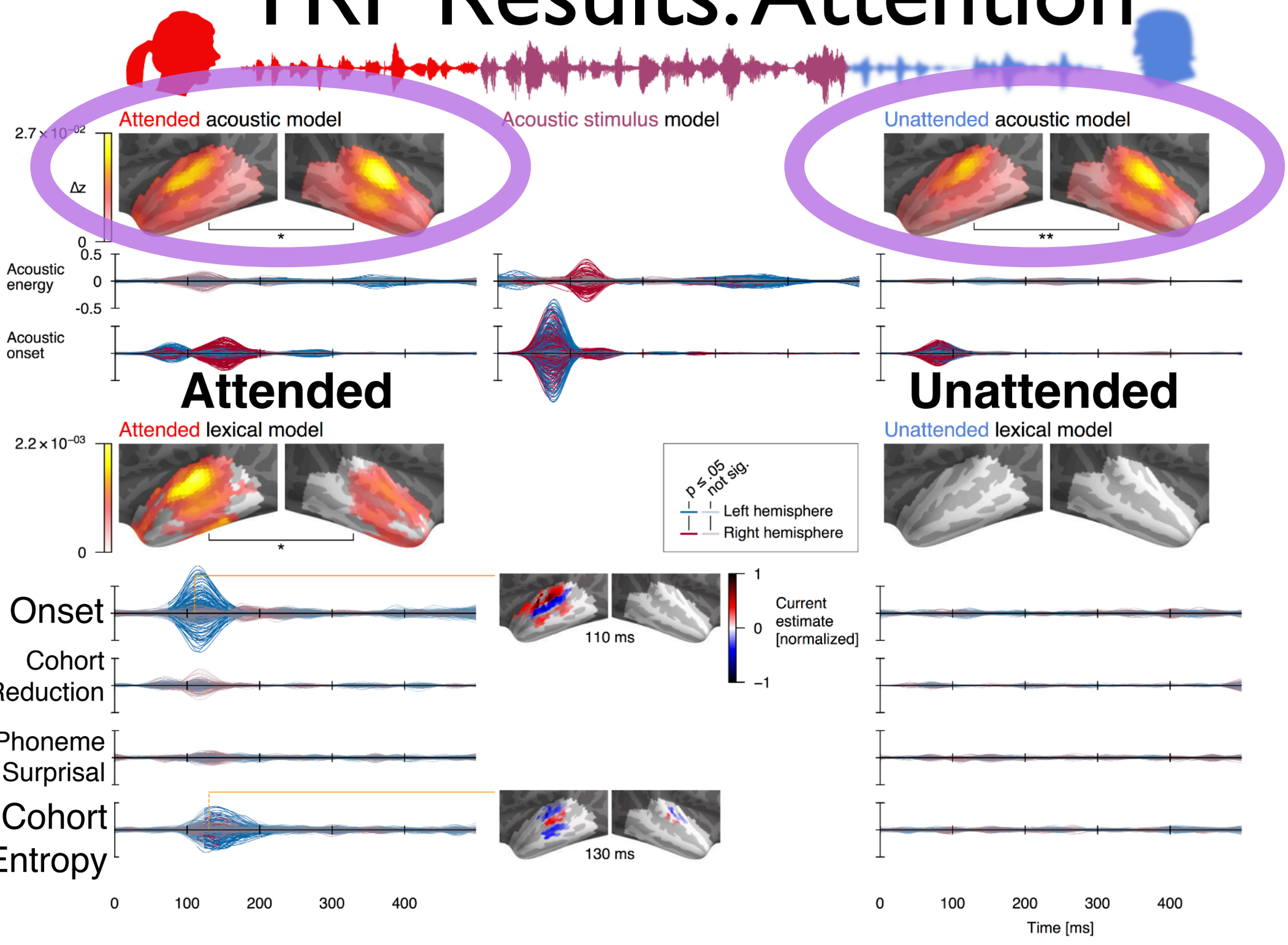
Brodbeck & Simon, bioRxiv 326785

TRF Results: Attention



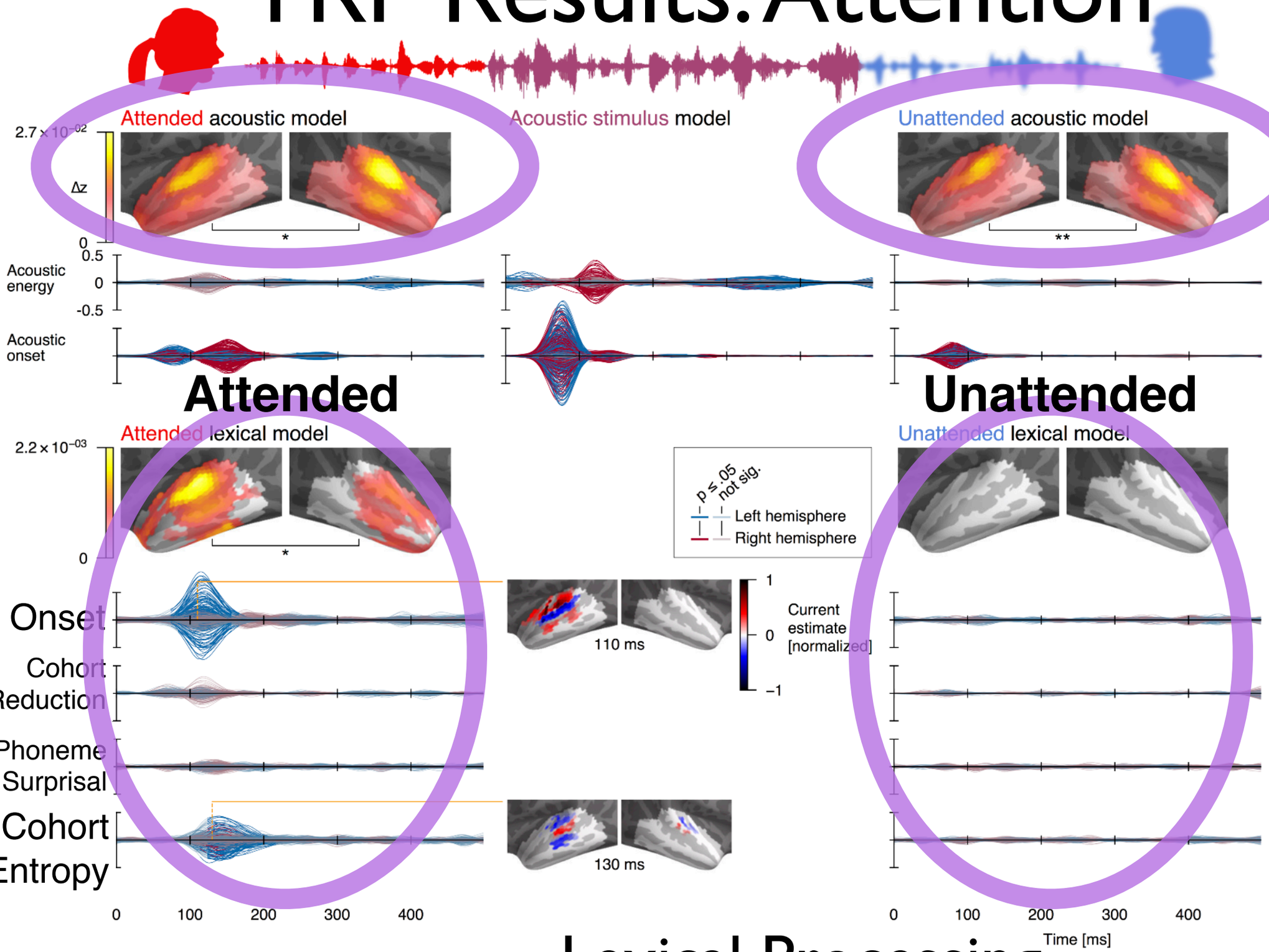
Brodbeck & Simon, bioRxiv 326785

TRF Results: Attention



Brodbeck & Simon, bioRxiv 326785

TRF Results: Attention



Brodbeck & Simon, bioRxiv 326785

Lexical Processing only for Attended Speech

Lexical Processing Summary

- Lexical processing: incremental integration of phonetic information for word identification
  - Transition from acoustic to linguistic information
  - Fast: ~110 ms post phoneme onset
- Responses also tracked word boundaries
- Left hemisphere dominant
- Lexical processing of only attended speech

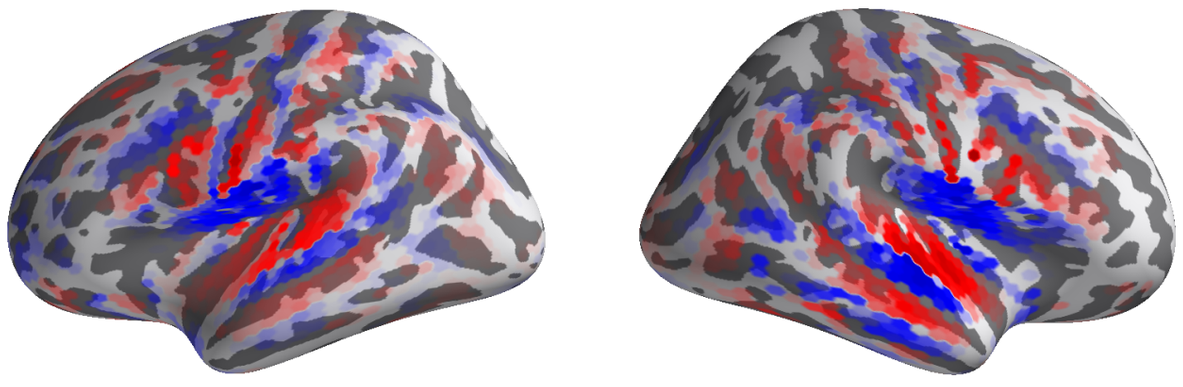
Outline

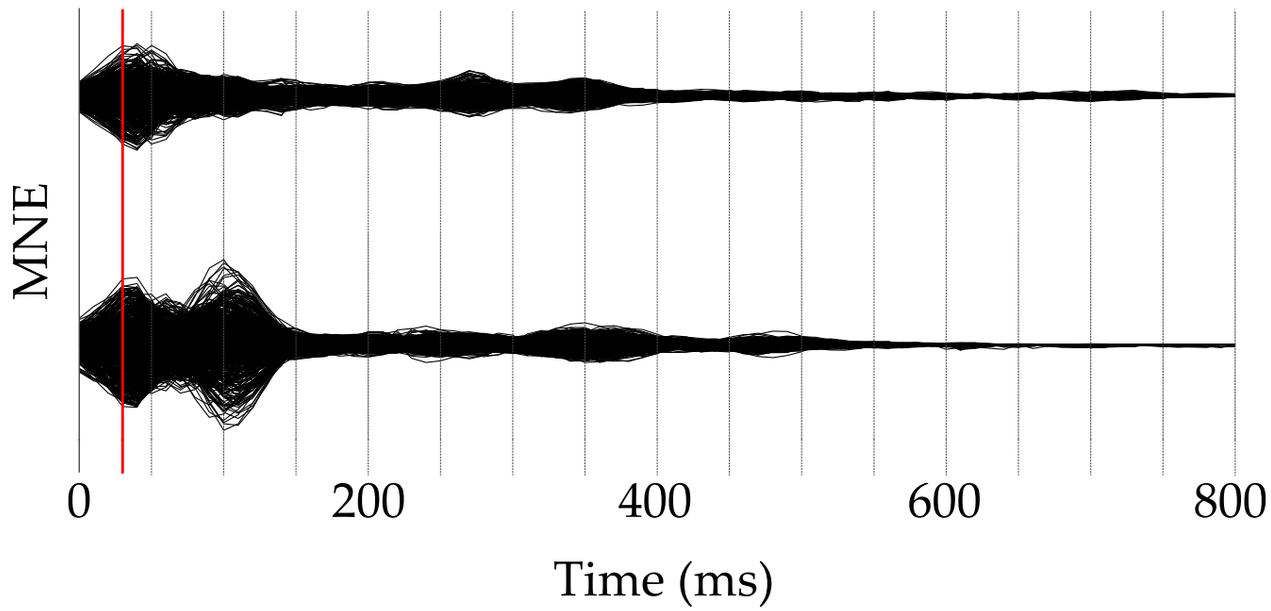
- Cortical Representations of Speech via MEG
- Missing speech / Familiar Speech
- Spatial Localization of Cortical Representations
  - Multiple Levels of Cortical Representations
- Representation Transition from Auditory/Phonetic → Lexical/Word-based
- Recent advances in Spatial Localization

Work in Progress

- Spatial localization's broad *spread* artifacts could be tamed by removing independent source analysis
  - Allow TRF sources to compete with each other
  - Keep competition for stimulus representations
- Difficult problem (non-convex)!
- But use of State-Space estimators and Expectation Maximization (EM) allows Direct TRF Localization

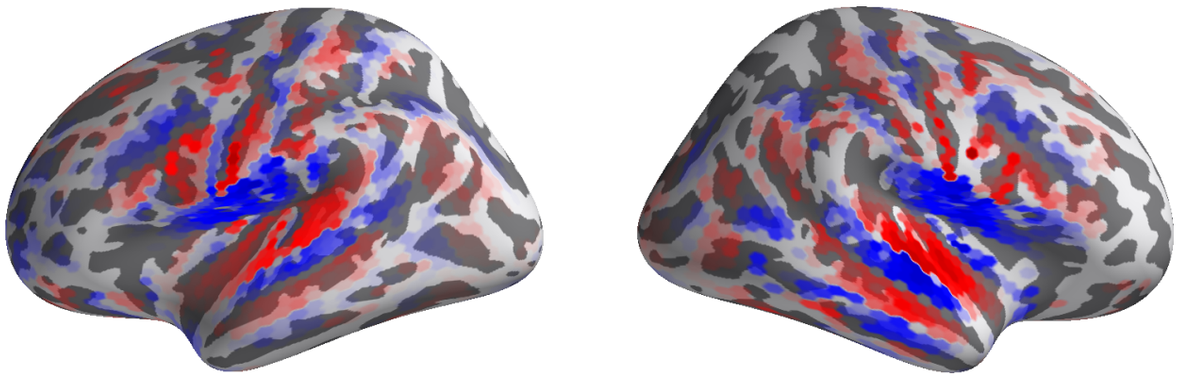
Direct TRF Localization

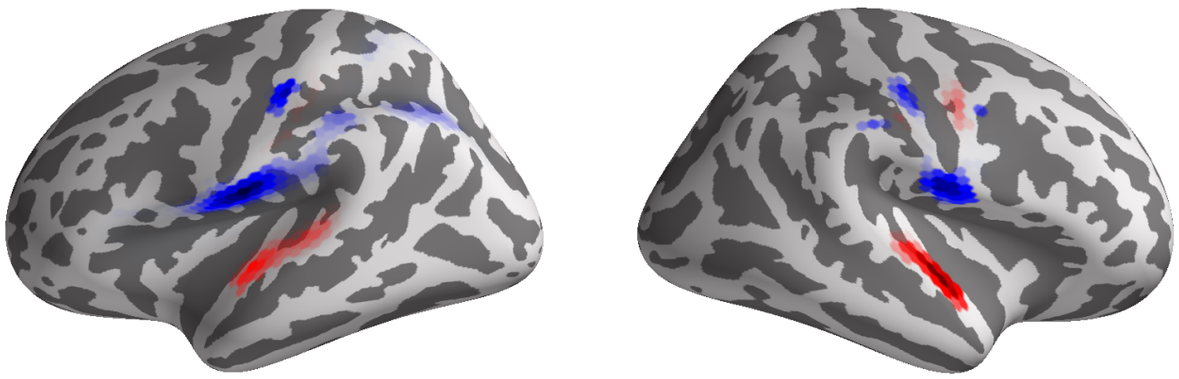


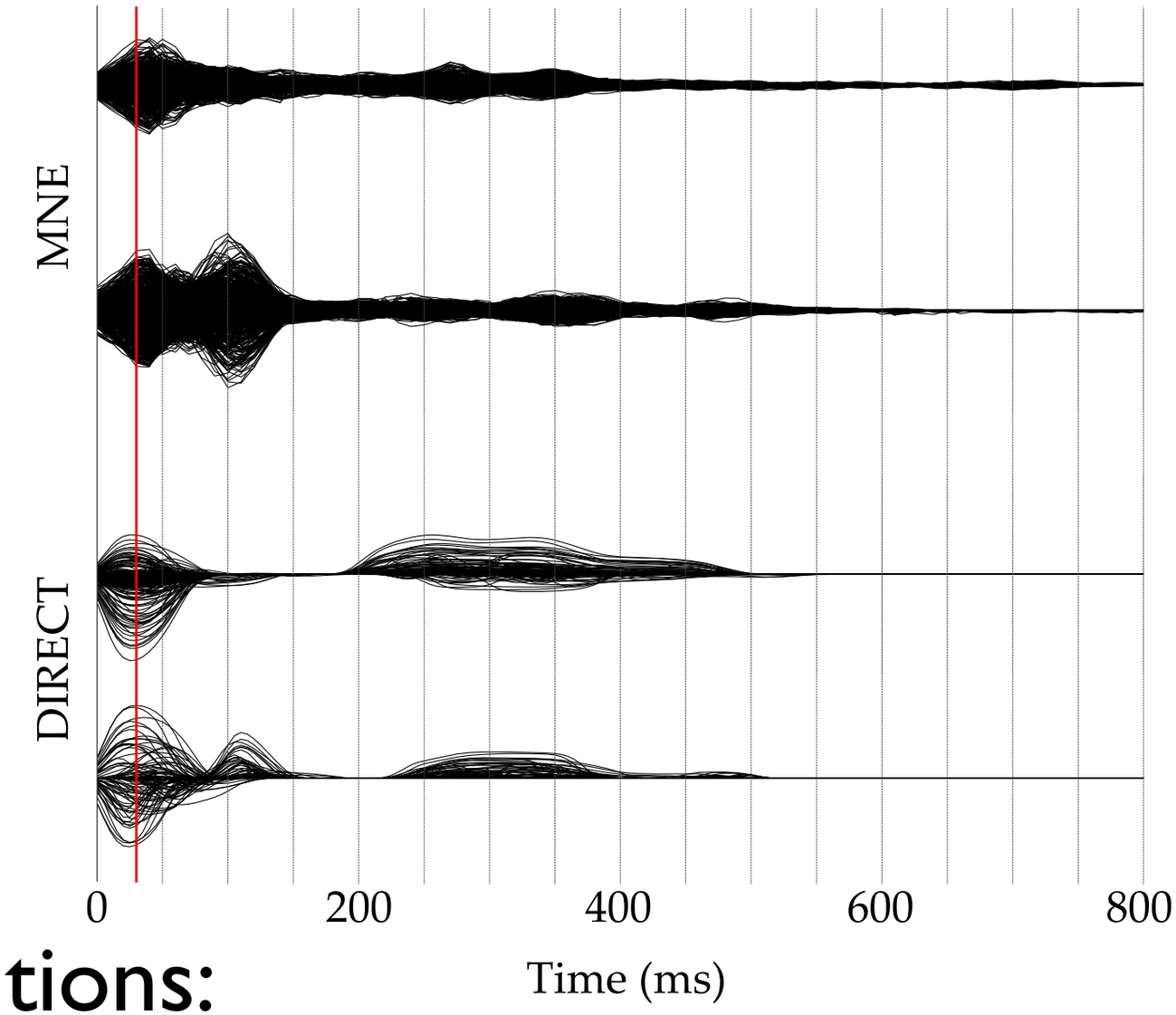


Work by Proloy Das, Behtash Babadi, and Christian Brodbeck

Direct TRF Localization





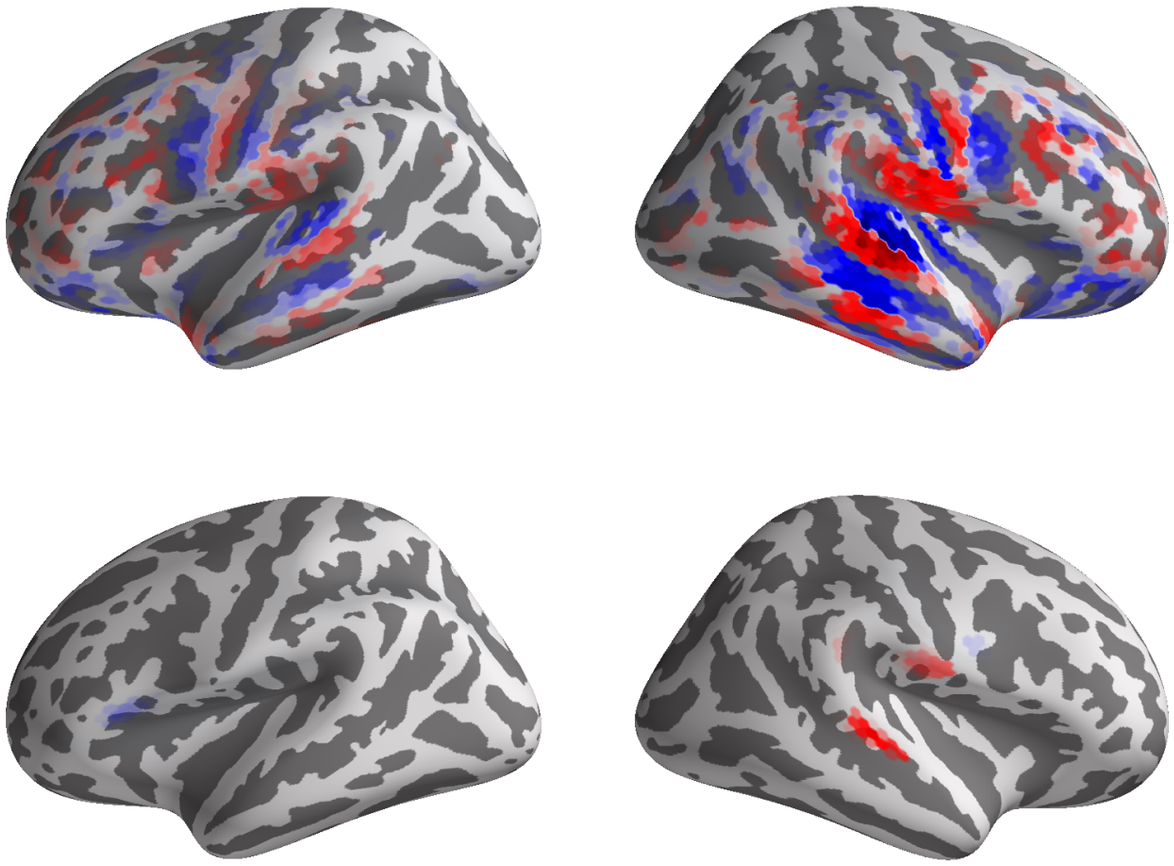


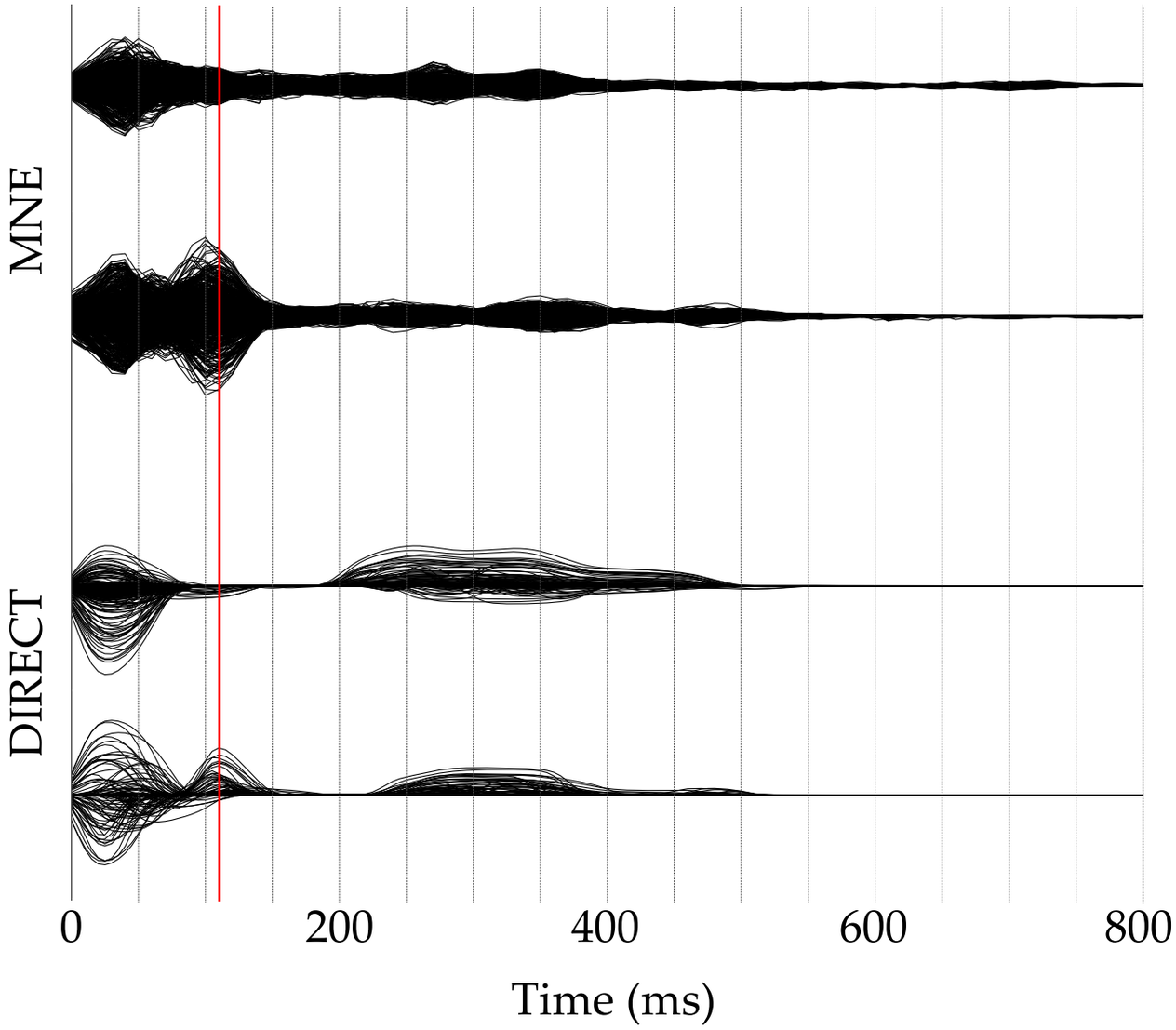
- Direct TRF localized solutions:
  - Naturally Sparse
  - Dominantly Temporal Lobe

Work by Proloy Das, Behtash Babadi, and Christian Brodbeck

Work in Progress

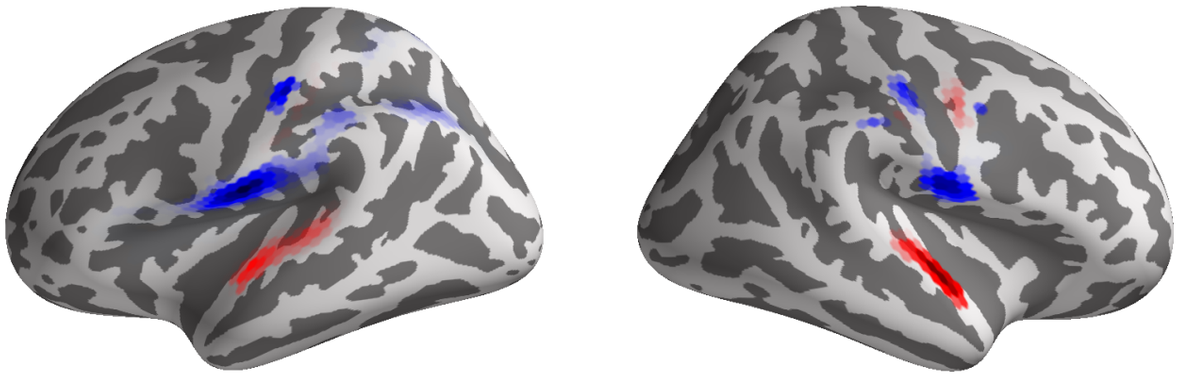
Auditory M100





Work by Proloy Das, Behtash Babadi, and Christian Brodbeck

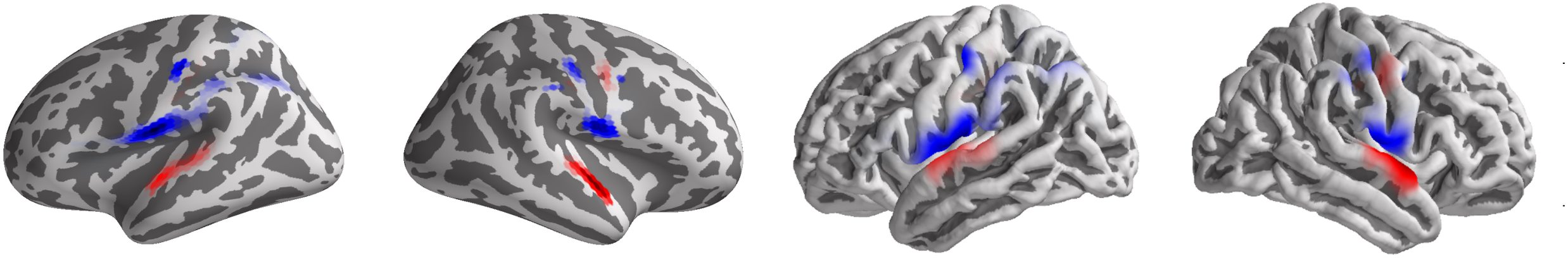
Work in Progress



- Cross-Sulcus artifacts explanation

Work by Proloy Das, Behtash Babadi, and Christian Brodbeck

Work in Progress



- Cross-Sulcus artifacts explanation

Work by Proloy Das, Behtash Babadi, and Christian Brodbeck

Summary

- Cortical representations of speech in MEG
  - temporal representation maintained even in the absence of speech, for familiar speech
  - facilitated for familiar speech
- Different temporal representations in different cortical areas
- Different cortical areas process different aspects of speech stimulus (e.g. acoustic, lexical, semantic)
- Transition from Acoustic/Phonetic to Lexical Processing is early (~110 ms), left-hemisphere dominant, and attention-dependent

Thank You