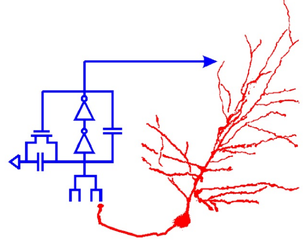


Computational
Sensorimotor
Systems Lab

Continuous speech and its neural representations, through auditory cortex and beyond

Jonathan Z. Simon

University of Maryland, College Park

Acknowledgements

Current Lab Members & Affiliates

Shohini Bhattasali
Regina Calloway
Jason Dunlap
Theo Dutcher
Sydney Hancock
Kevin Hu
Neha Joshi
Dushyanthi Karunathilake
Joshua Kulasingham
Janani Perara
Mohsen Rezaeizadeh
Behrad Soleimani

Current & Recent Collaborators

Samira Anderson
Behtash Babadi
Tom Francart
L. Elliot Hong
Stefanie Kuchinsky
Ellen Lau
Elisabeth Marsh
Philip Resnik

Recent Lab Members & Affiliates

Sahar Akram
Ross Baehr
Anurupa Bhonsale
Christian Brodbeck
Francisco Cervantes Constantino
Aura Cruz Heredia
Proloy Das
Lien Decruy
Marisel Villafane Delgado
Nai Ding
Marlies Gilles
Alex Jiao
Sina Miran
David Nahmias
Alex Presacco
Krishna Puvvada
Jonas Vanthornhout
James Williams
Peng Zan

Funding & Support

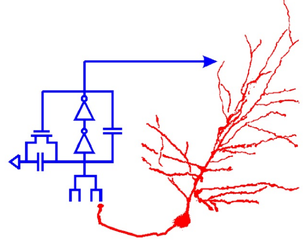
NIDCD

NIH
NIA

NSF

DARPA

UNIVERSITY OF MARYLAND 18 56



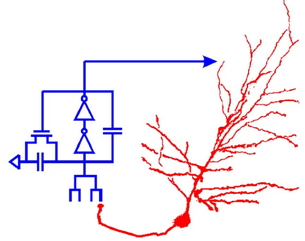
Computational
Sensorimotor
Systems Lab

Continuous speech and its neural representations, through auditory cortex and beyond

SCAN, 21 May 2021

To Intelligibility, and Beyond!

SCAN, 21 May 2021



Computational
Sensorimotor
Systems Lab

Continuous speech and its neural representations, through auditory cortex and beyond

SCAN, 21 May 2021

Cortical Representations of Continuous Speech

Continuous speech

- naturalistic
- redundant
- employs auditory cognition
- acoustically diverse
- drives most auditory areas
- …
- but also complicated

If you happened to find yourself on the banks of the Ohio River on a particular afternoon in the spring of 1806—somewhere just to the north of Wheeling, West Virginia, say …

The Botany of Desire — Michael Pollan

Alfred the Great was a young man, three-and-twenty years of age, when he became king. Twice in his childhood, he had been taken to Rome, where the Saxon nobles were in the habit of going on journeys which they supposed to be religious; …

A Child's History of England — Charles Dickens

In the bosom of one of those spacious coves which indent the eastern shore of the Hudson, at that broad expansion of the river denominated by the ancient Dutch navigators …

The Legend of Sleepy Hollow — Washington Irving

He was an old man who fished alone in a skiff in the Gulf Stream and he had gone eighty-four days now without taking a fish. In the first forty days a boy had been with him. But after forty days without a fish …

The Old Man and the Sea — Ernest Hemingway

Cortical Representations of Continuous Speech

Temporal neural patterns ⇆ temporal patterns in speech

- Need high temporal precision, for fast temporal speech features

  - EEG (electroencephalography): *whole brain*
  - MEG (magnetoencephalography): *whole brain but with strong cortical bias*
  - ECoG (electrocorticography): *placed cortical surface electrodes*
  - single- and multi-unit recording methods: *placed depth electrodes*

Cortical Representations of Continuous Speech

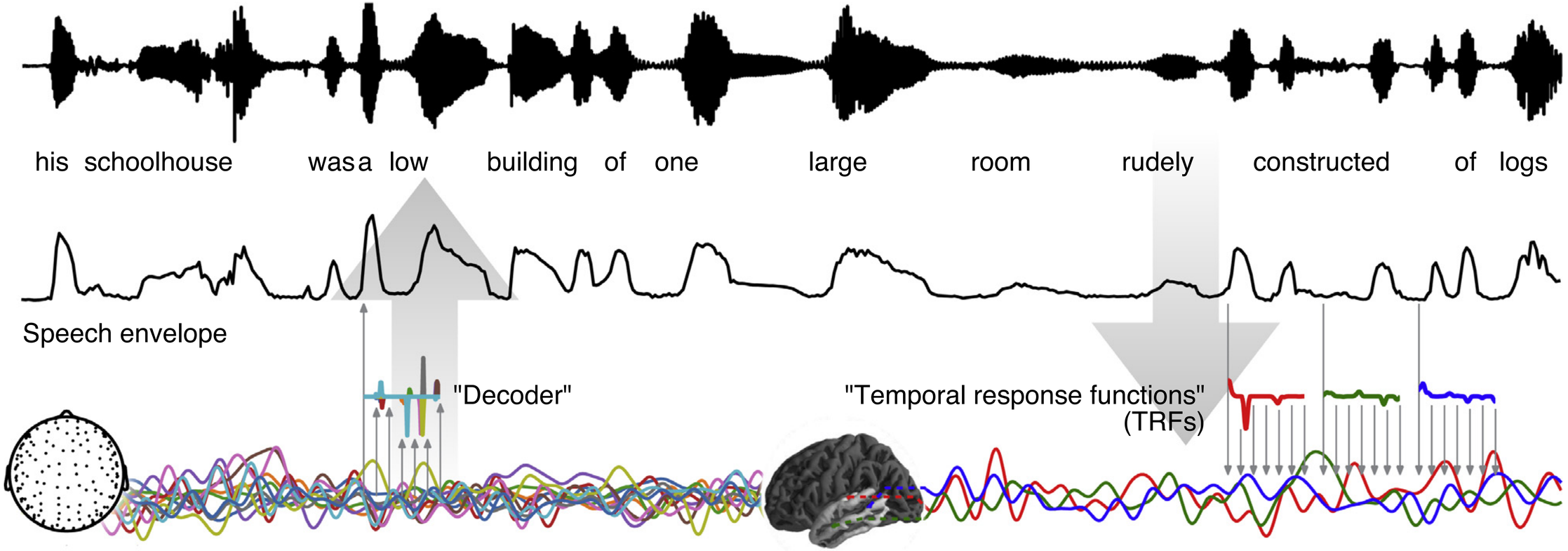
Neural Representations of Speech

- driven oscillations at pitch frequencies (mostly subcortical)
  - acoustic onset tracking
    - speech envelope rhythmic following
      - phoneme-based responses
        - phoneme-context-based responses
          - sentence-structure rhythm following
            - semantic structure tracking
- plus connections to **intelligibility/perception/behavior**

Brodbeck & Simon (2020) Continuous Speech Processing, Curr Op Physiol

Cortical Representations of Speech

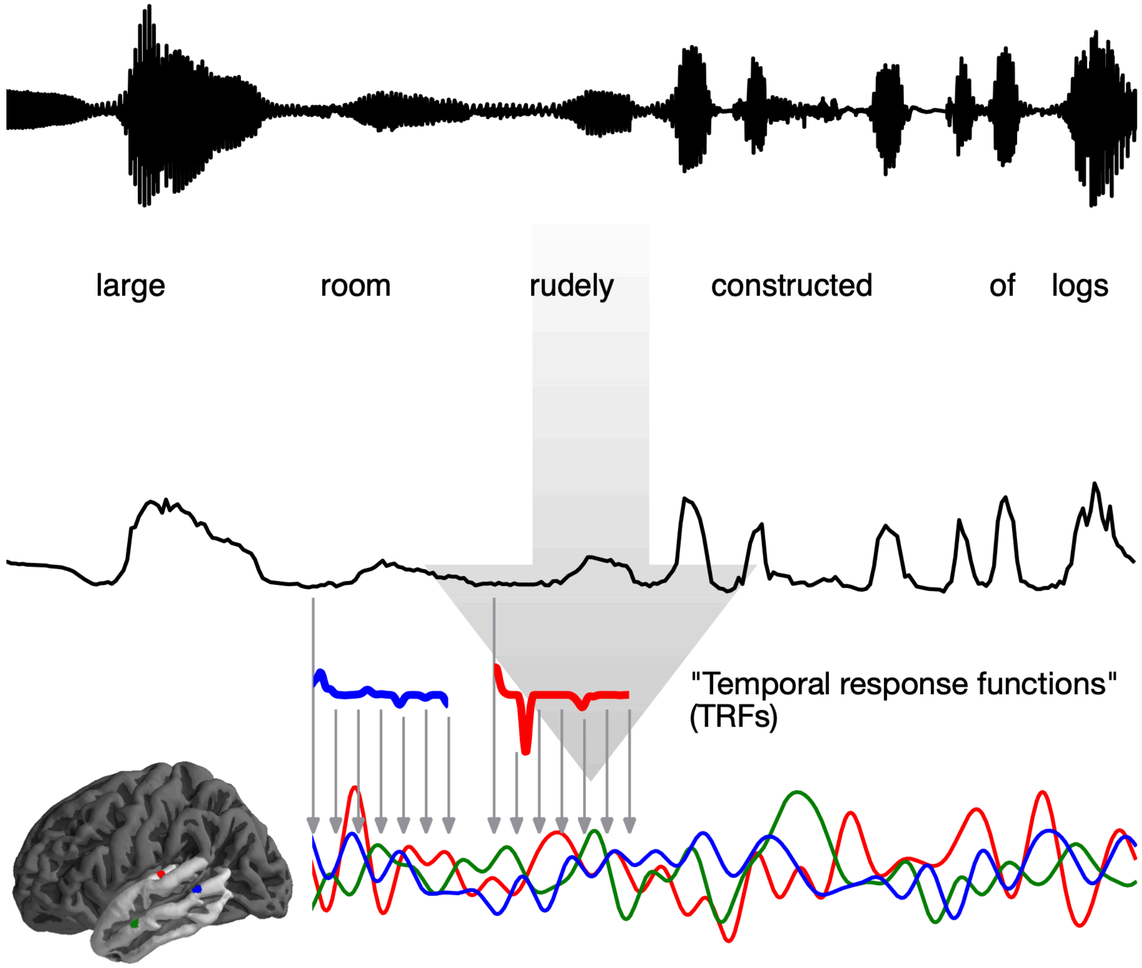
- Measure *time-locked* responses to temporal pattern of speech features (in humans)
- Any speech feature of interest: acoustic envelope, lexical, pitch, semantic, etc.
- Infer spatio-temporal neural origins of neural responses



Brodbeck & Simon (2020) Continuous Speech Processing, Curr Op Physiol

Cortical Representations: Encoding

- Predicting future neural responses from present stimulus features,
  - wide variety of stimulus features
  - via Temporal Response Function (TRF)
- Why look at encoding? It *often* tells us more about the brain
  - TRF analogous to evoked response
  - peak amplitude ≈ processing intensity
  - peak latency ≈ source location
  - multiple TRFs simultaneously

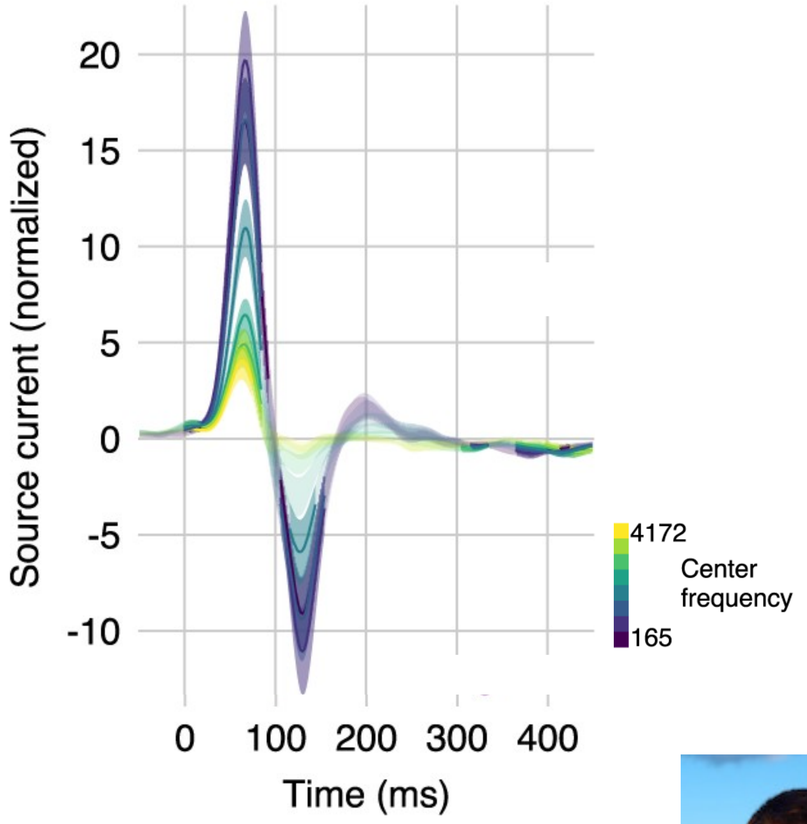


Example: MEG Prediction of Voxel Responses

Cortical Representations: Speech Envelope

- TRF interpretable a la evoked response
  - Has M50 (“P1”) & M100 (“N1”) peaks, but from instantaneous speech envelope
  - early peak localizes to primary auditory areas (HG)
  - later peak localizes to associative areas (PT)
  - caveat: actually from envelope *onset*
- This is from a single talker, clean speech
  - simple but limiting
  - what about noise? other speakers? attention?
  - can the speech representation be cleaned?

Temporal Response Fields

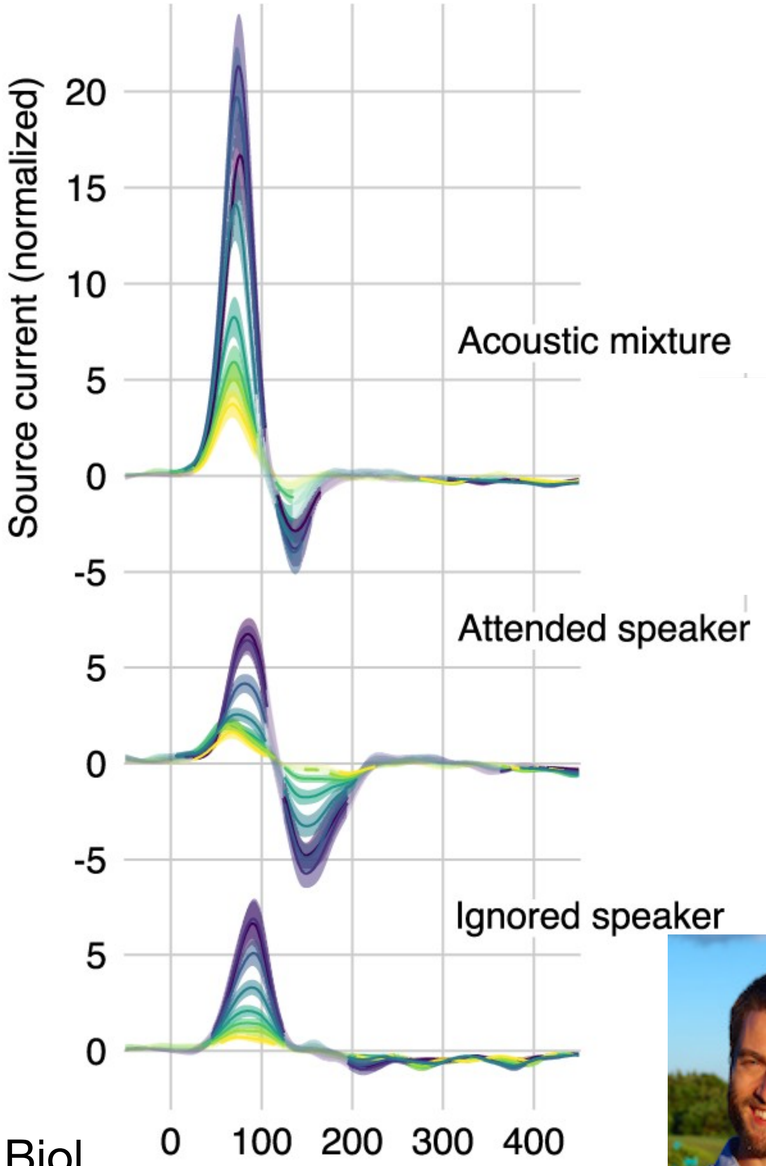


Brodbeck et al. (2020) *Neural Speech Restoration at the Cocktail Party …*, PLoS Biol

Cortical Representations: Attention

Two competing speakers, selectively attend to one

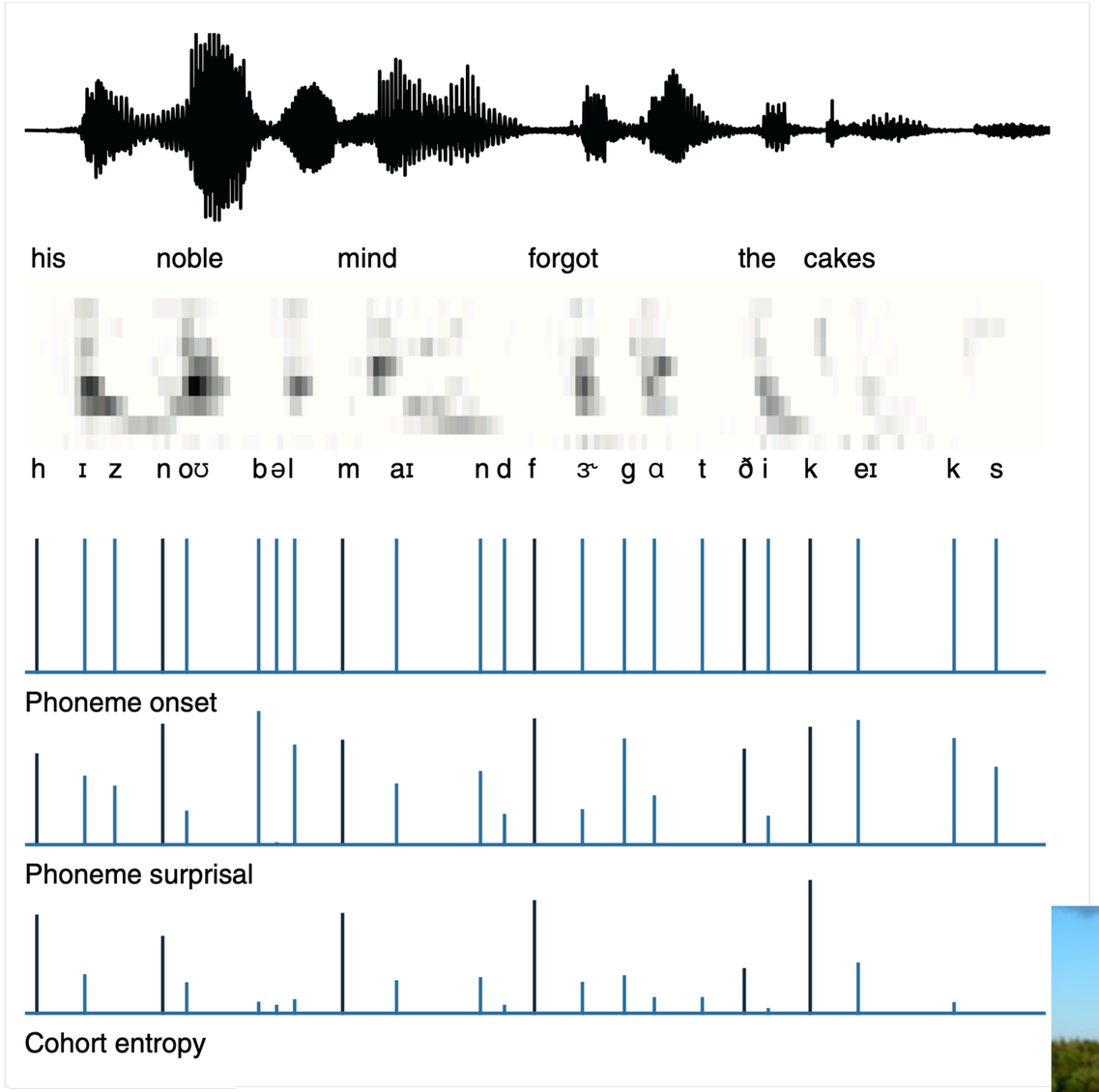
- more illuminating since more complex auditory scene
- need more care re: “stimulus” responsible for responses
  - acoustic mixture entering ears
  - foreground speech
  - background speech
- estimate all TRFs simultaneously
  - compete to explain variance



Brodbeck et al. (2020) *Neural Speech Restoration at the Cocktail Party …*, PLoS Biol

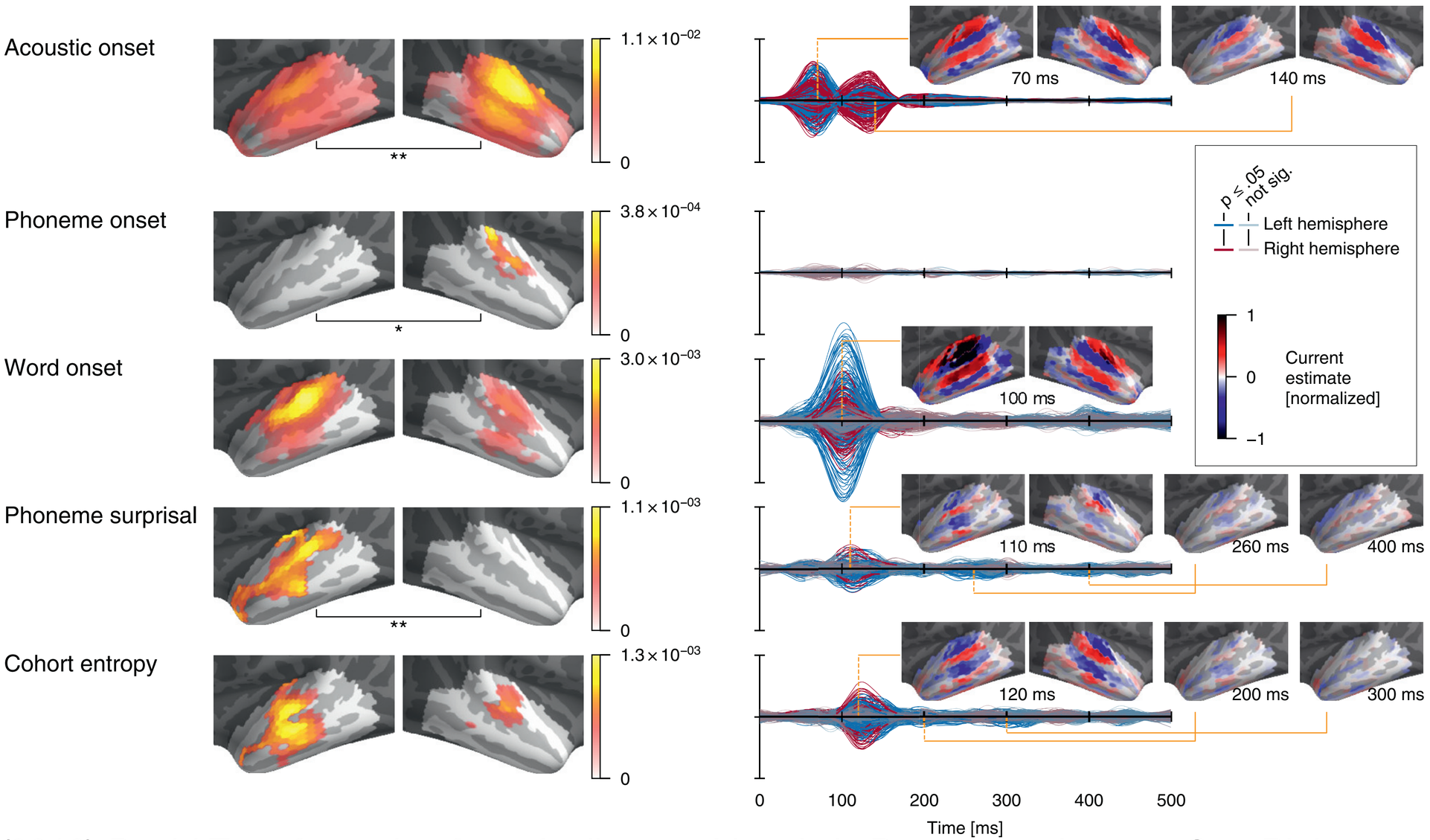
Cortical Representations: Language Features

- Language-based speech features
  - phonemes
  - words & word boundaries
  - phoneme context
- All TRFs estimated simultaneously
  - compete to explain variance



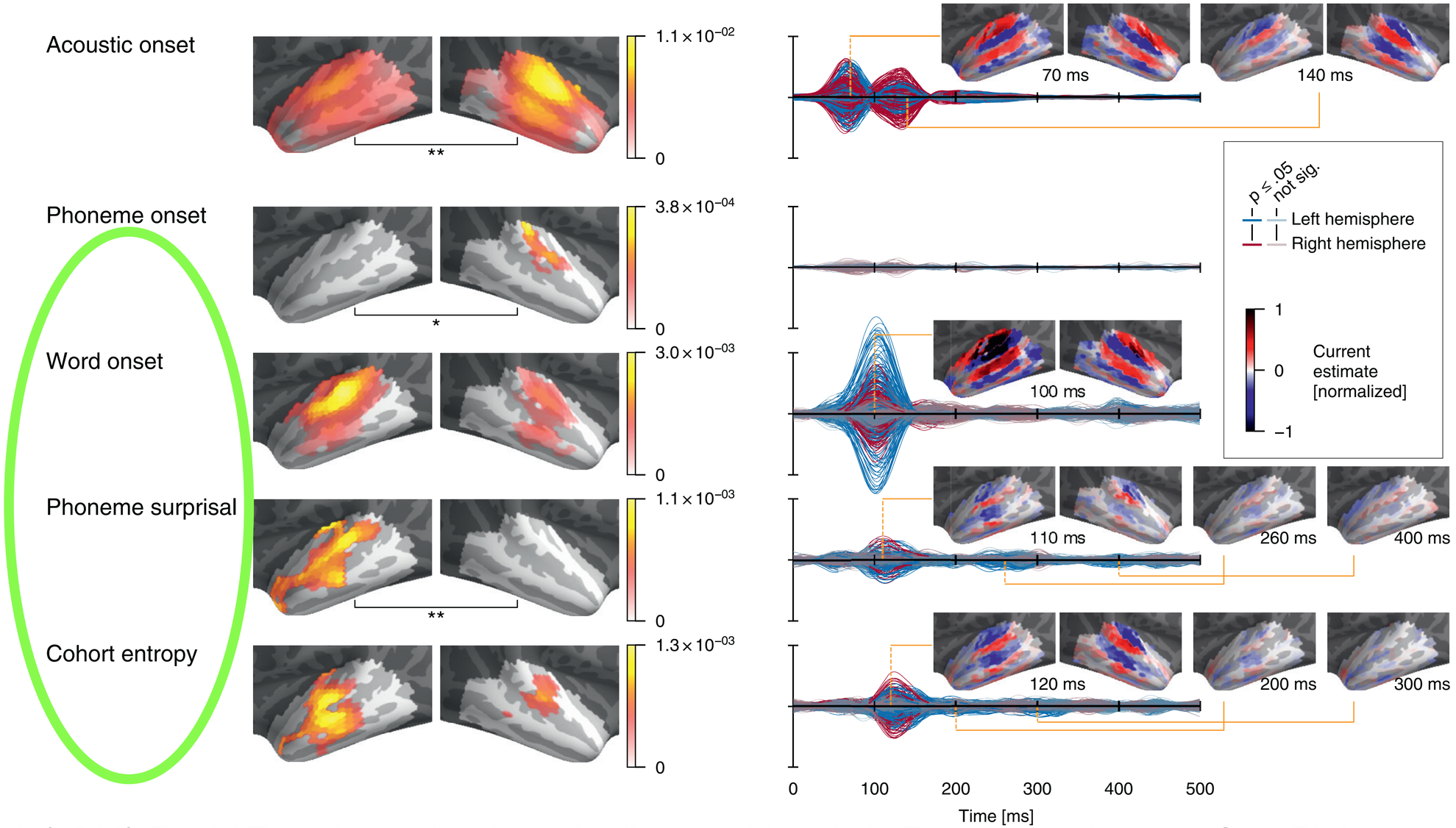
Brodbeck et al. (2018) *Rapid Transformation from Auditory to Linguistic Representations …*, Curr Biol

Language-feature based TRFs



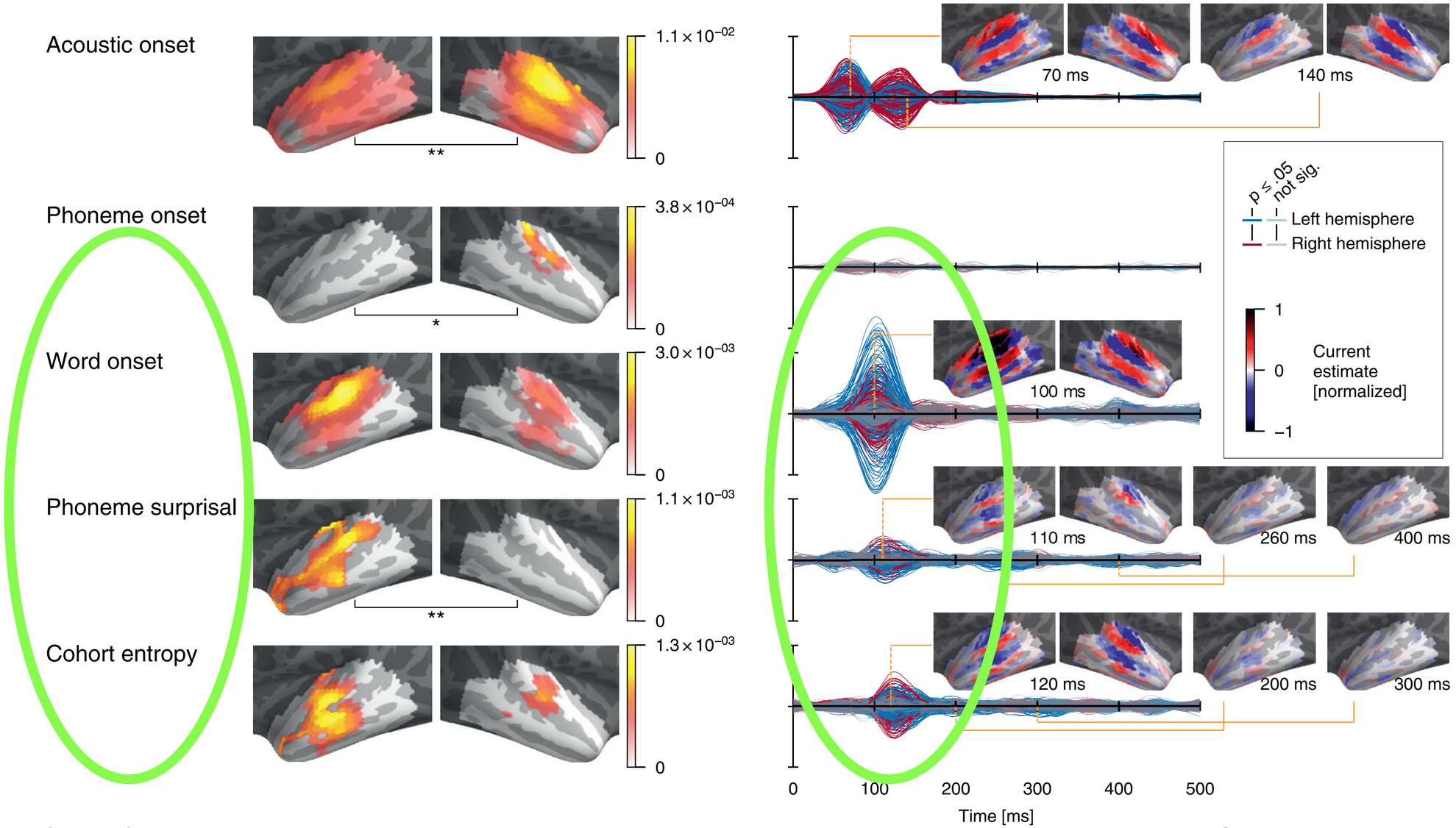
Brodbeck et al. (2018) *Rapid Transformation from Auditory to Linguistic Representations …*, Curr Biol

Language-feature based TRFs



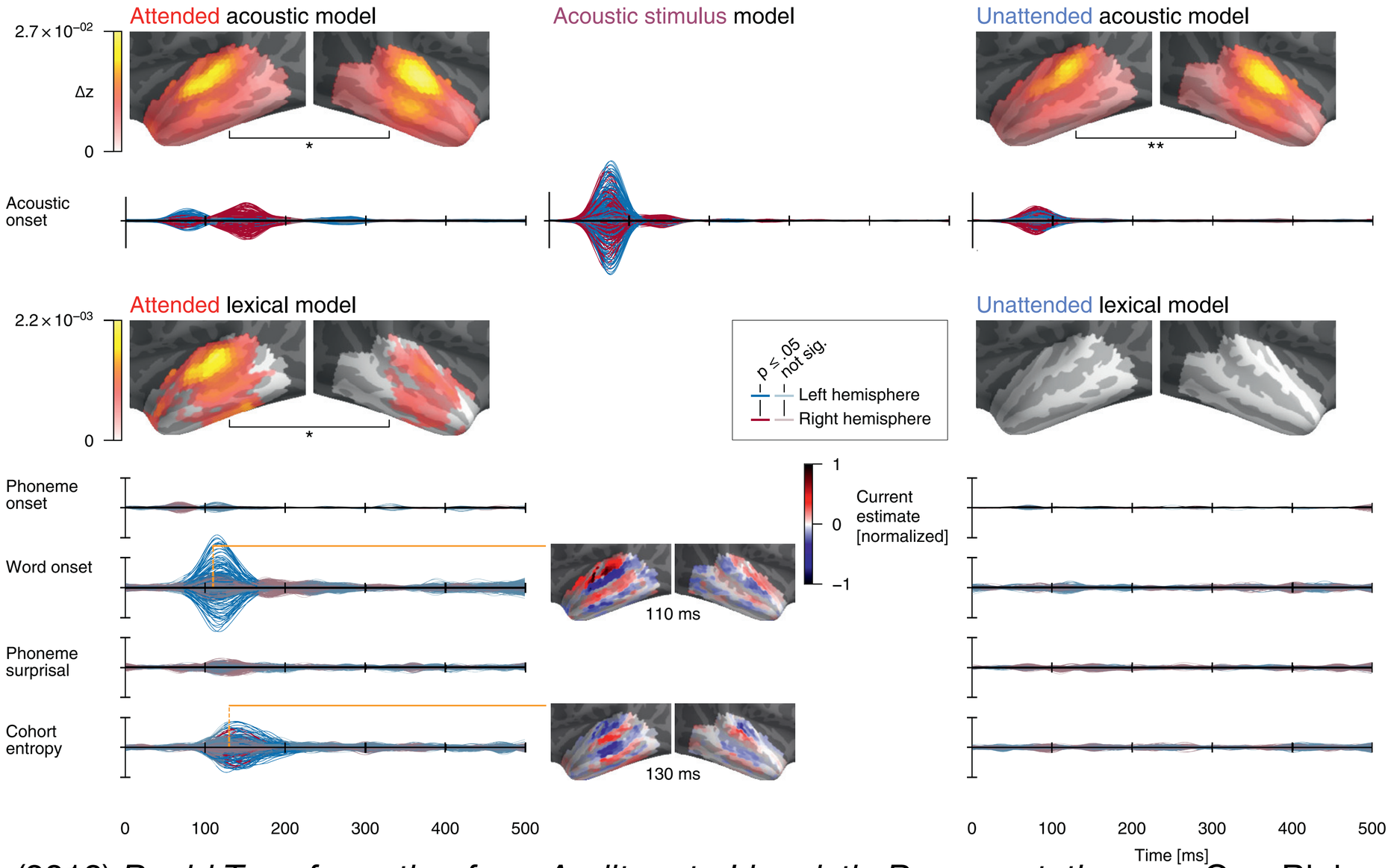
Brodbeck et al. (2018) *Rapid Transformation from Auditory to Linguistic Representations …*, Curr Biol

Language-feature based TRFs



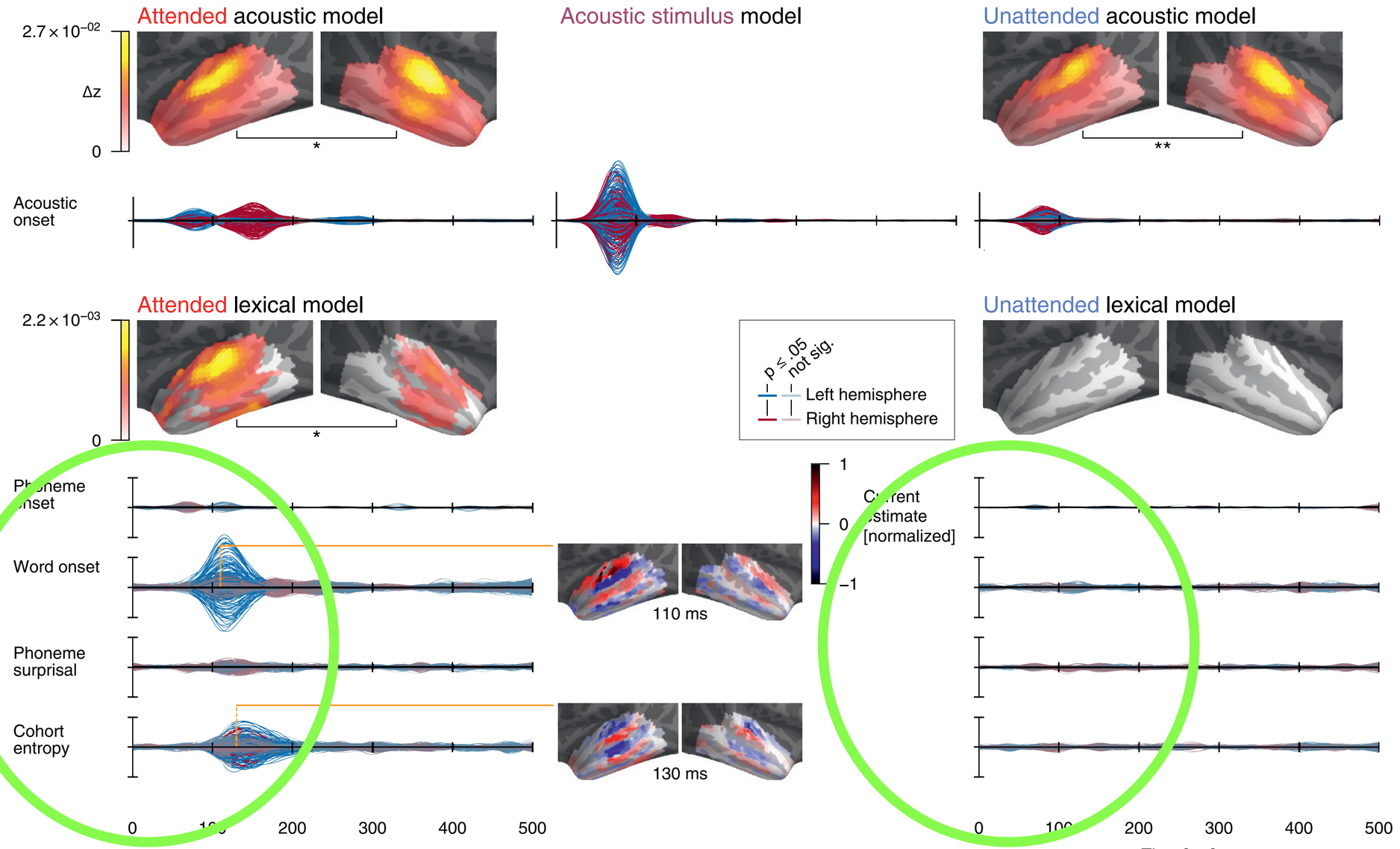
Brodbeck et al. (2018) *Rapid Transformation from Auditory to Linguistic Representations …*, Curr Biol

Attention + Language-feature based TRFs



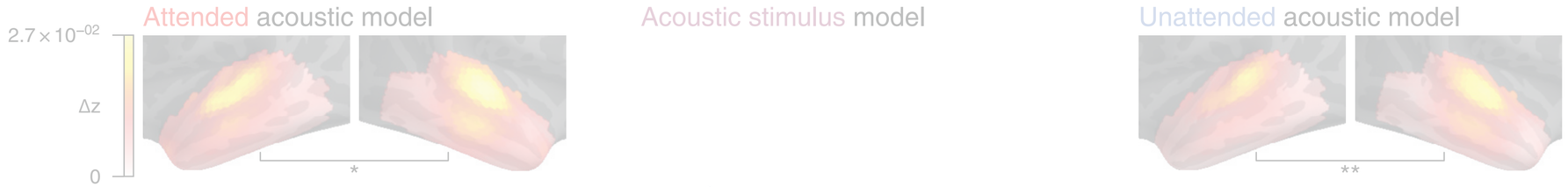
Brodbeck et al. (2018) *Rapid Transformation from Auditory to Linguistic Representations …*, Curr Biol

Attention + Language-feature based TRFs



Brodbeck et al. (2018) *Rapid Transformation from Auditory to Linguistic Representations …*, Curr Biol

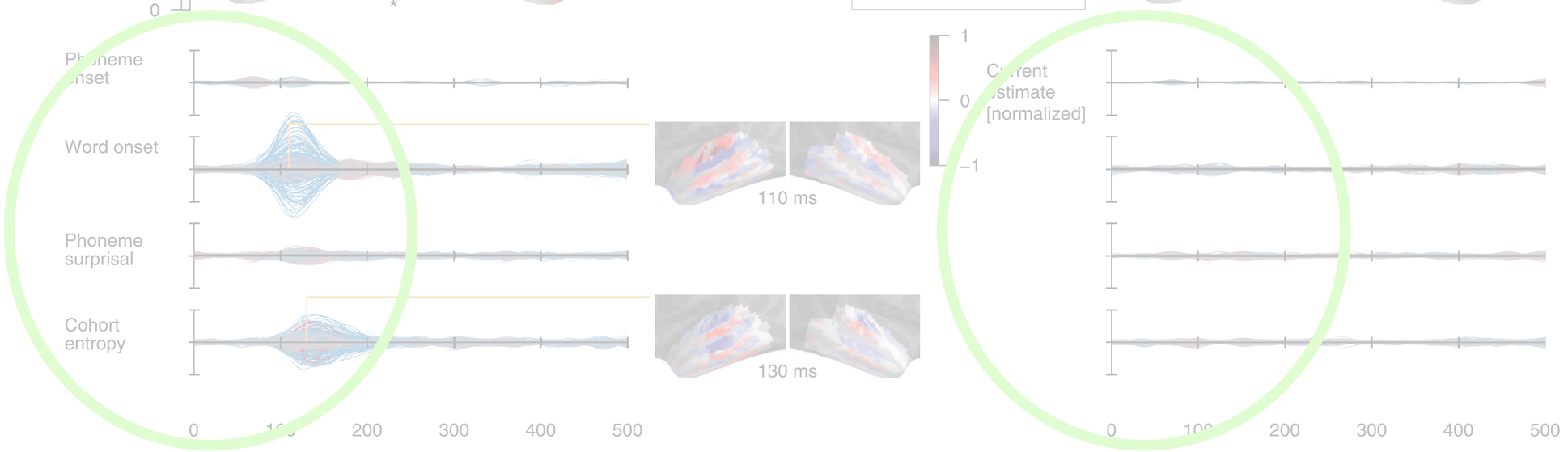
Attention + Language-feature based TRFs



See also:

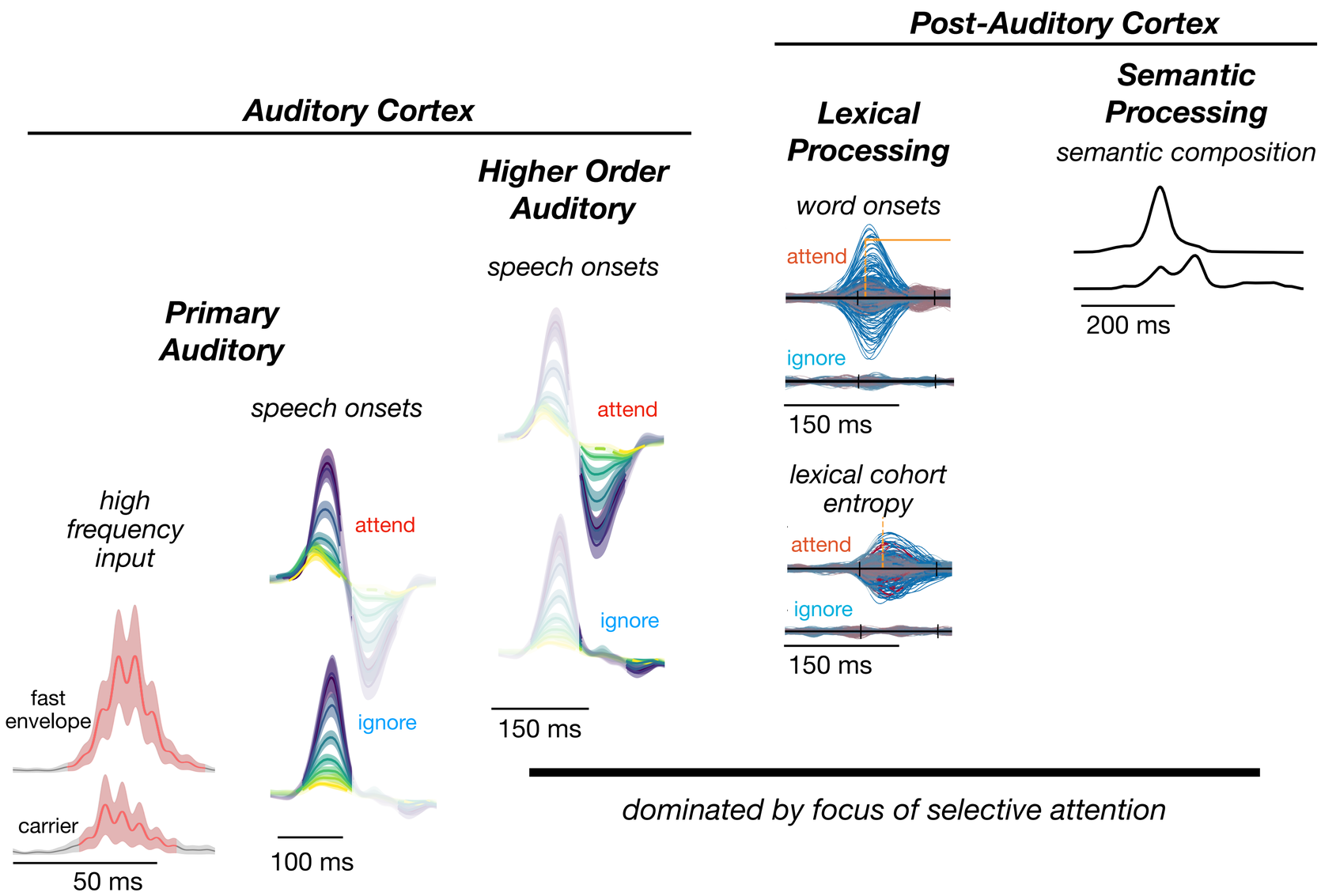
Gillis et al., (2021) bioRxiv

Neural Markers of Speech Comprehension: Measuring EEG Tracking of Linguistic Speech Representations, Controlling the Speech Acoustics

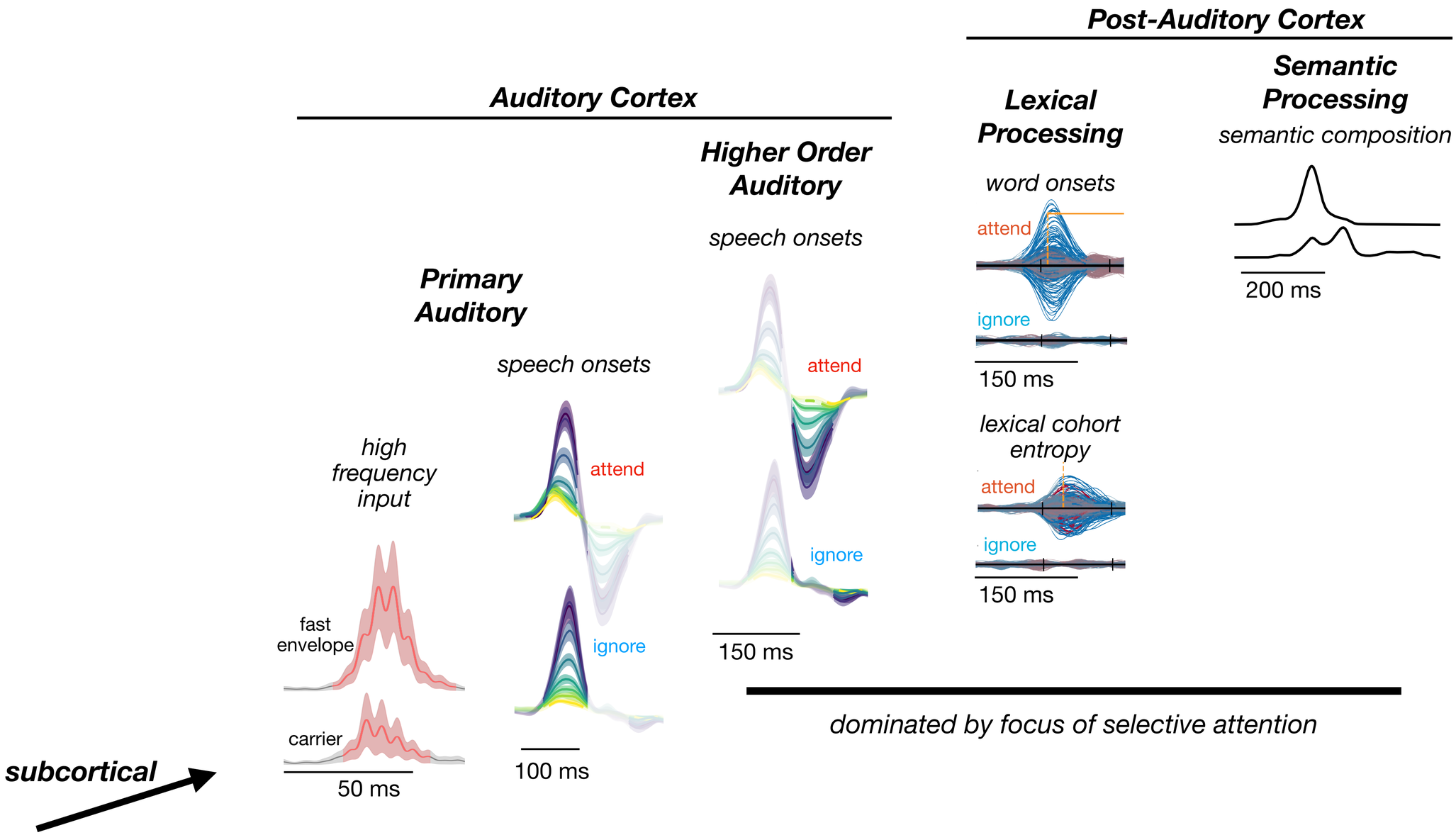


Brodbeck et al. (2018) *Rapid Transformation from Auditory to Linguistic Representations …*, Curr Biol

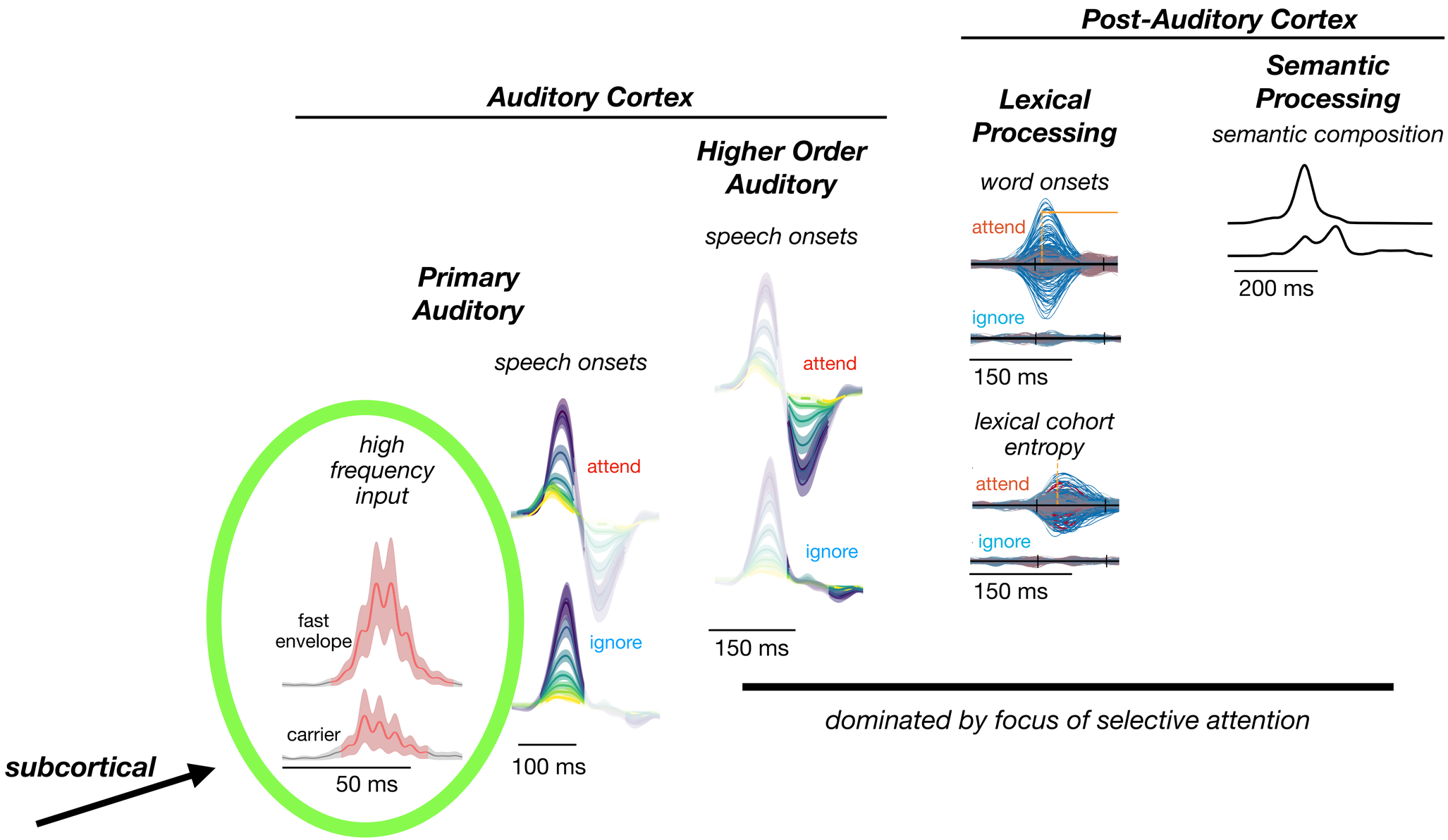
Cortical Representations Across Cortex



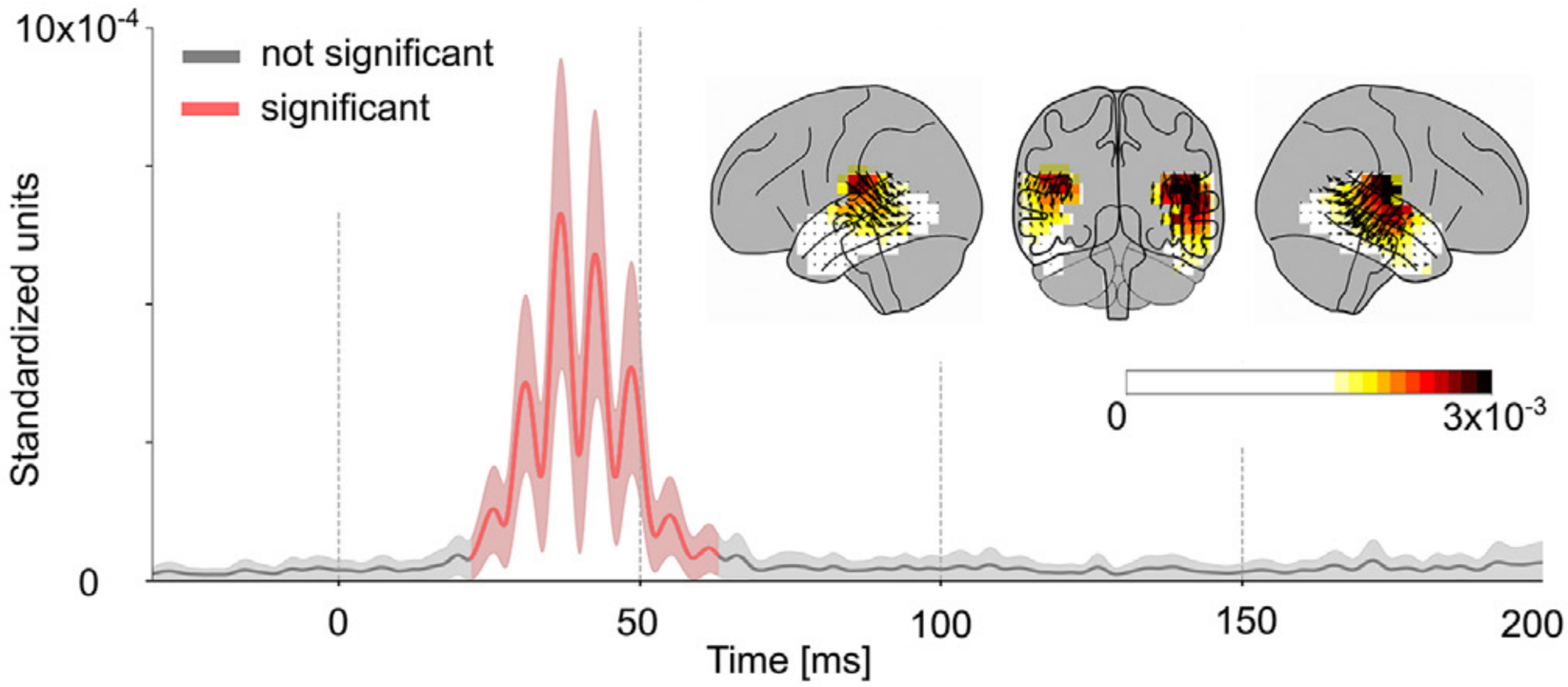
Cortical Representations Across Cortex



Cortical Representations Across Cortex

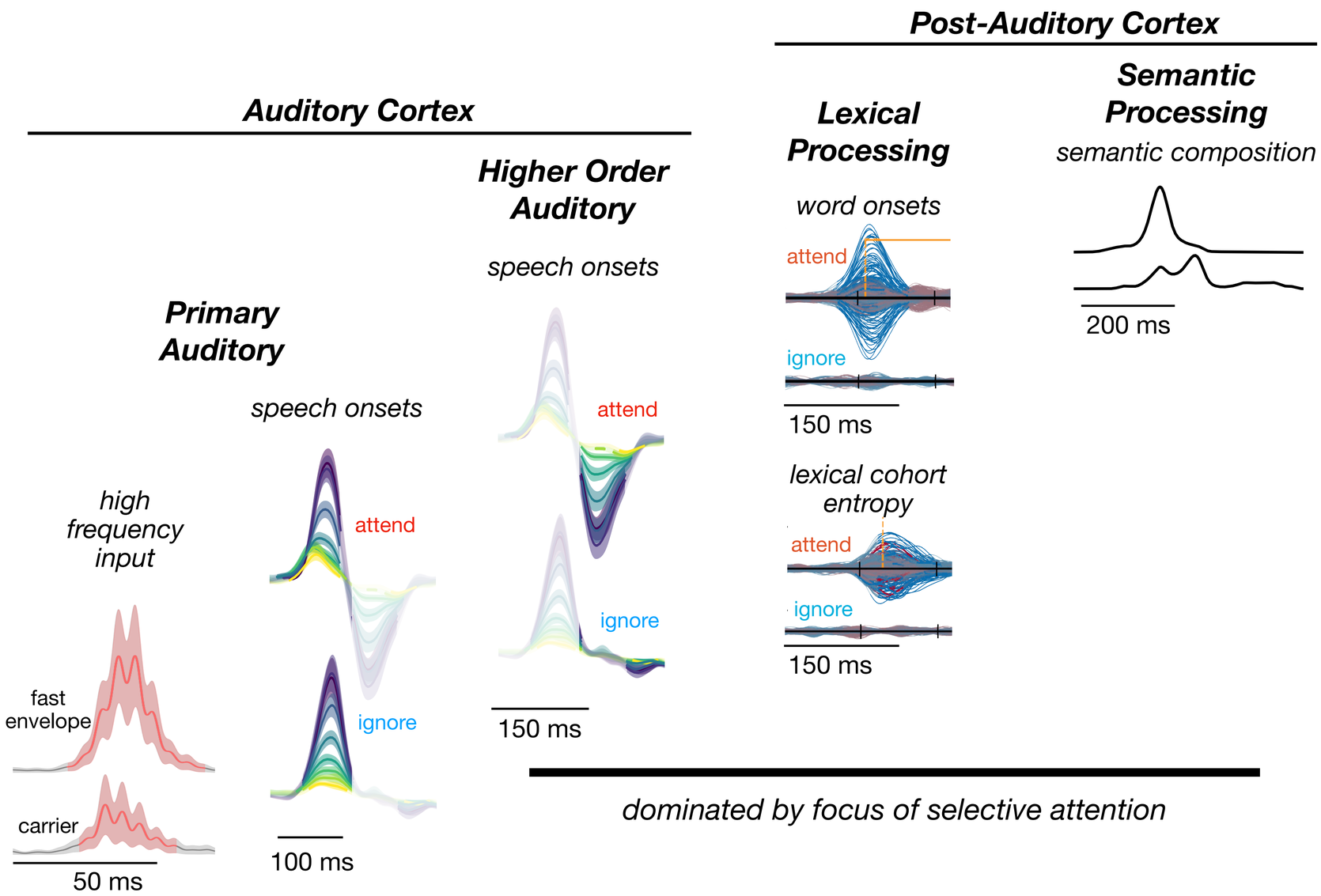


Fast & Early Cortical Representations

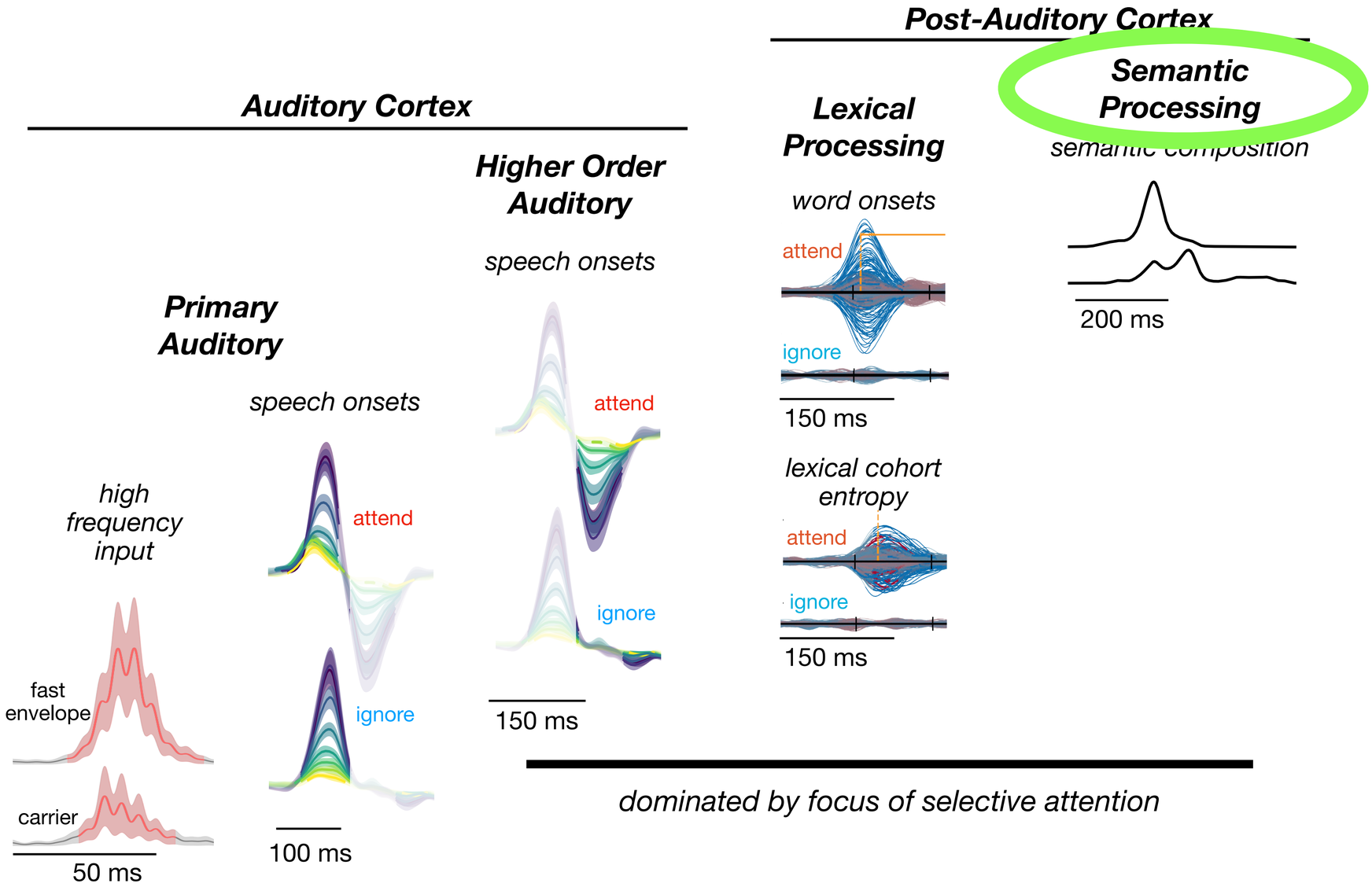


Kulasingham et al. (2020) *High Gamma Cortical Processing of Continuous Speech …*, NeuroImage

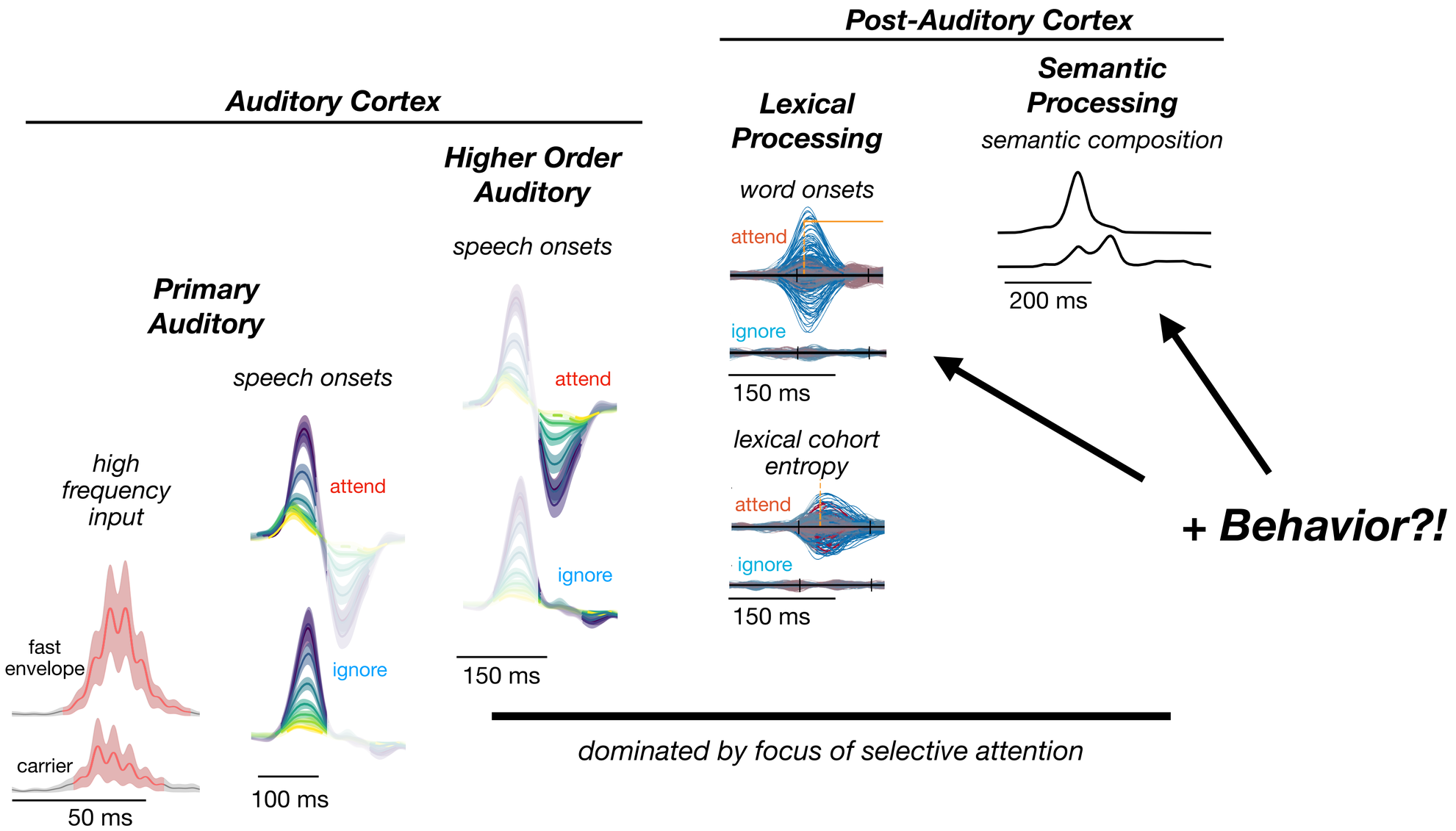
Cortical Representations Across Cortex



Cortical Representations Across Cortex



Cortical Representations Across Cortex



To Intelligibility, and Beyond?

Cortical Representations of Speech Understanding

- Behavioral correlates of speech understanding
  - implies language comprehension
  - higher order comprehension (?)
    - sentence structure
    - other structures, e.g. poetic, logical

- Neural correlates of speech understanding
  - rhythms of higher order structures, even if *totally absent in the acoustics*
    - sentence structures

Ding et al., Nat Neurosci 2016

Cortical Representations of Speech Understanding

- Behavioral correlates of speech understanding
  - implies language comprehension
  - higher order comprehension (?)
    - sentence structure
    - other structures, e.g. poetic, logical

- Neural correlates of speech understanding
  - rhythms of higher order structures, even if *totally absent in the acoustics*
    - sentence structures
    - poetic structures

Ding et al., Nat Neurosci 2016
Teng et al., Curr Biol 2020

Cortical Representations of Speech Understanding

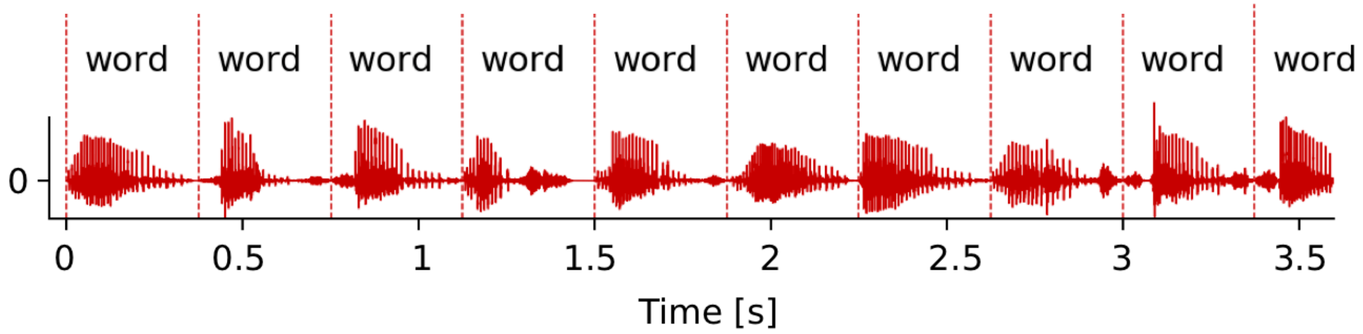
- Behavioral correlates of speech understanding
  - implies language comprehension
  - higher order comprehension (?)
    - sentence structure
    - other structures, e.g. poetic, logical

- Neural correlates of speech understanding
  - rhythms of higher order structures, even if *totally absent in the acoustics*
    - sentence structures
    - poetic structures
    - mathematical structures

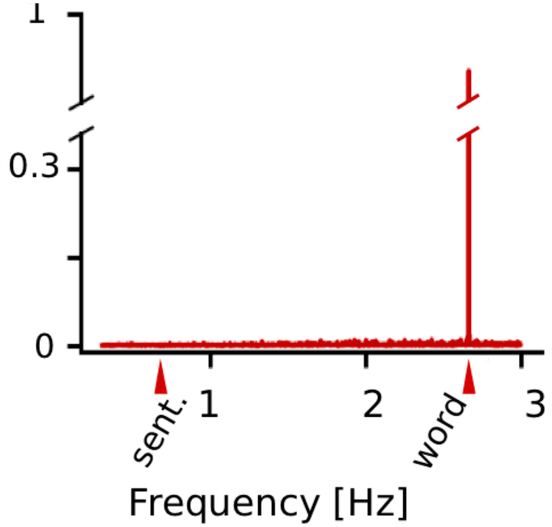
Ding et al., Nat Neurosci 2016
Teng et al., Curr Biol 2020

Isochronous Speech

Acoustics

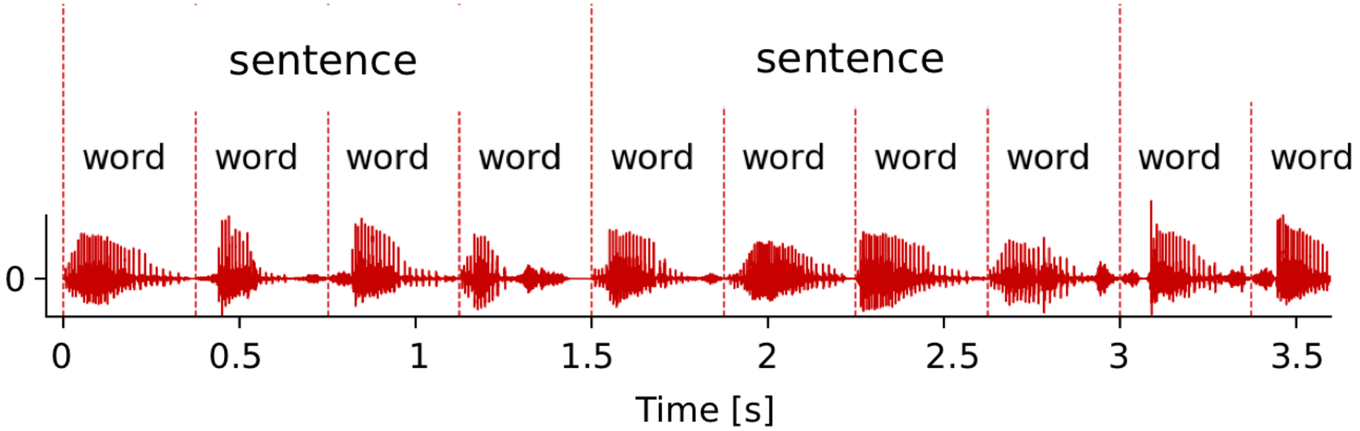


Acoustical Spectrum (envelope)

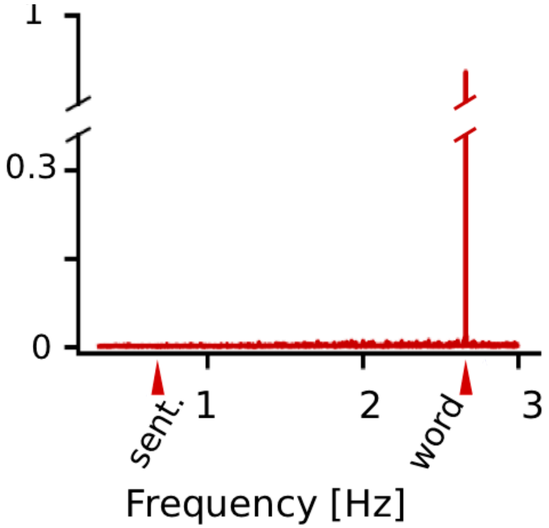


Isochronous Speech

Acoustics

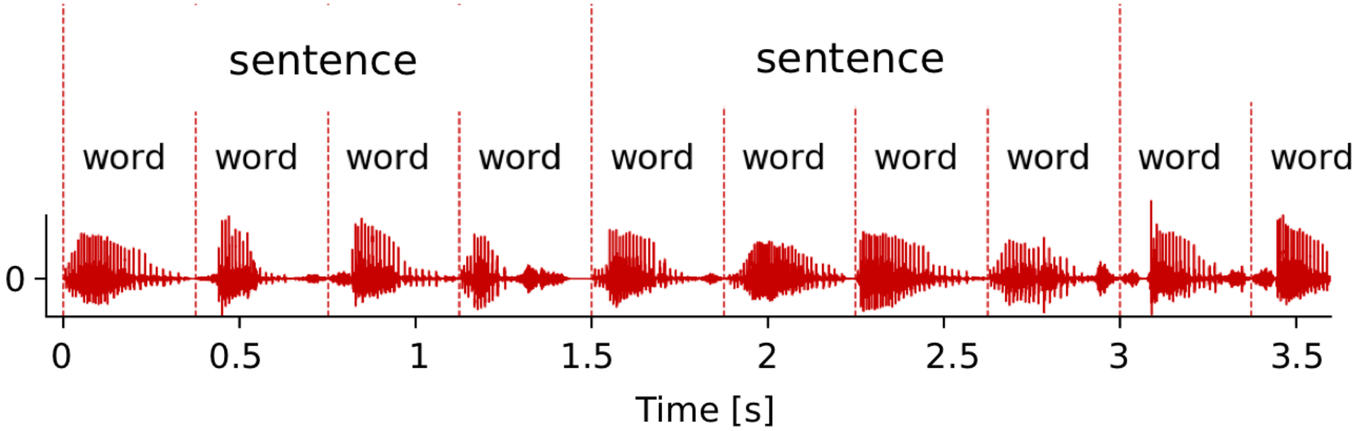


Acoustical Spectrum (envelope)

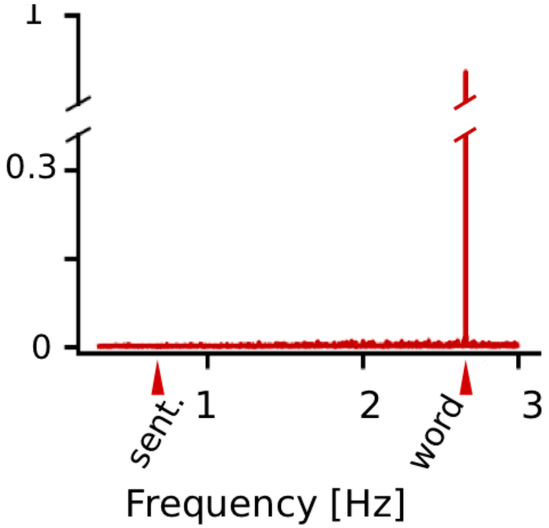


Isochronous Speech

Acoustics

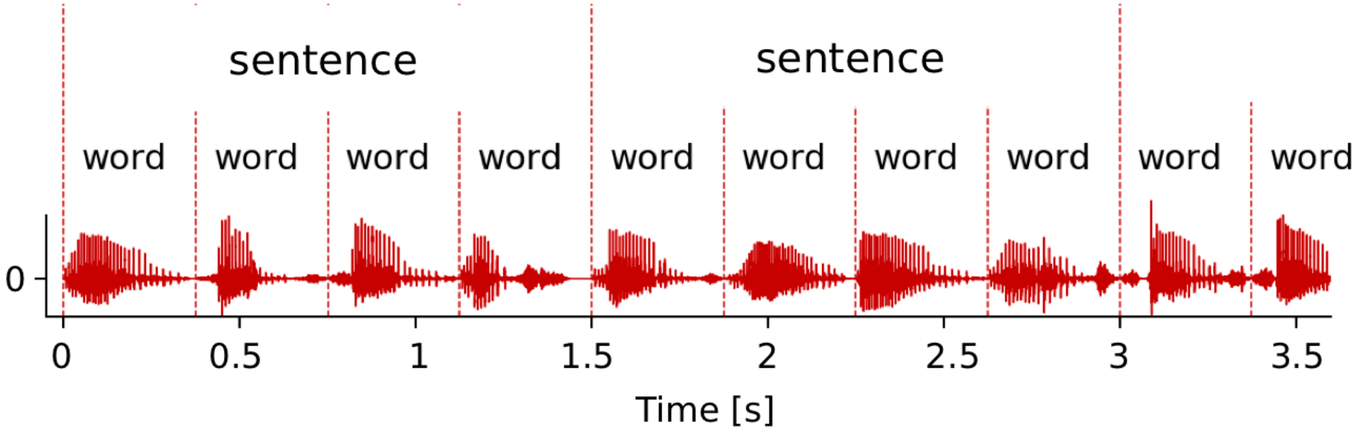


Acoustical Spectrum (envelope)

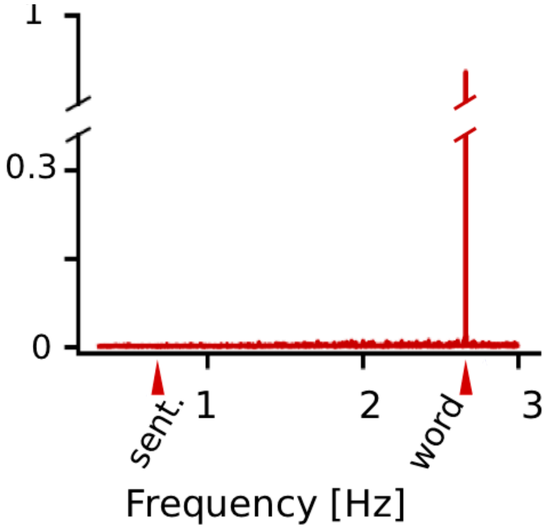


Isochronous Speech

Acoustics

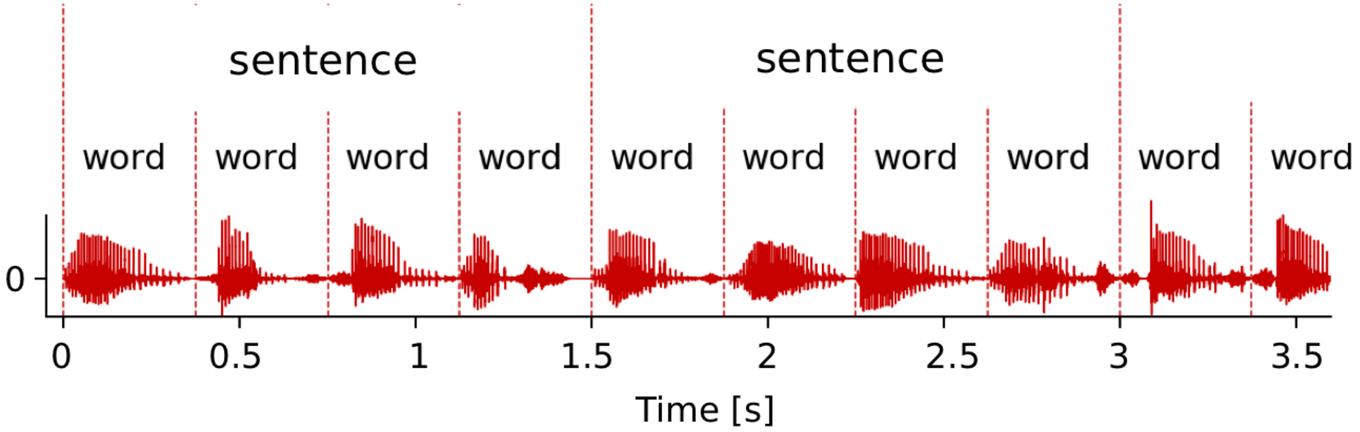


Acoustical Spectrum (envelope)

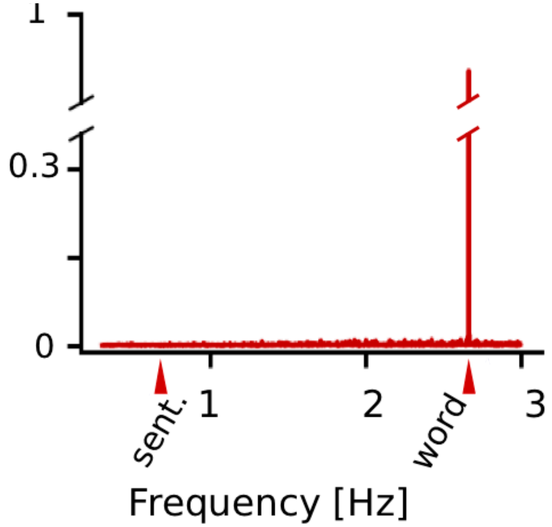


Isochronous Speech

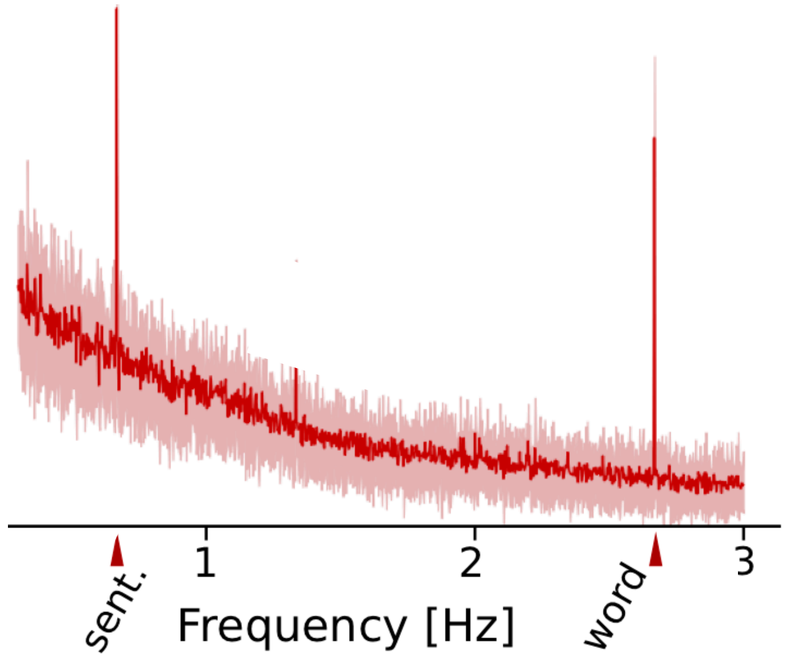
Acoustics



Acoustical Spectrum (envelope)

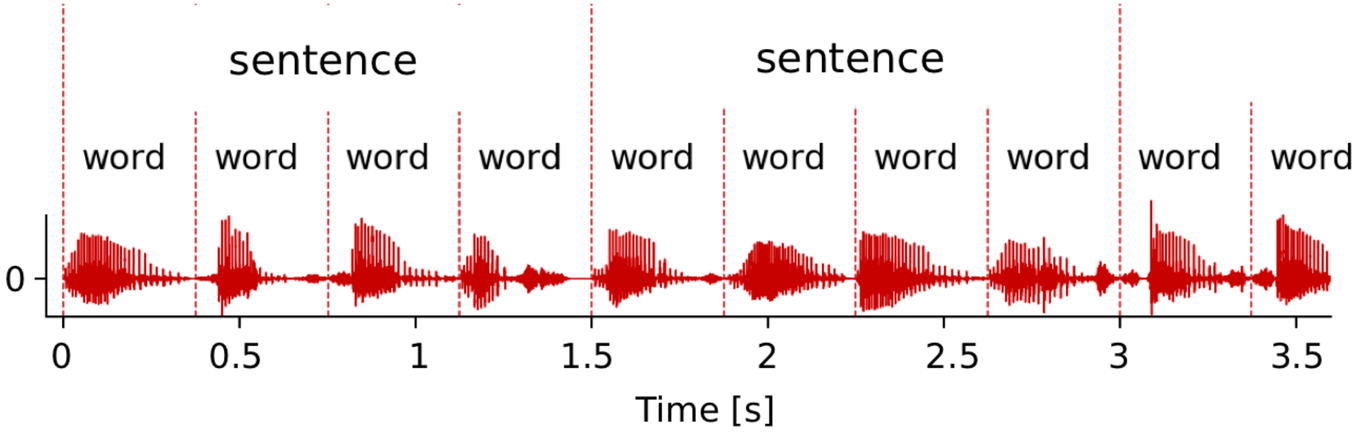


Perception?

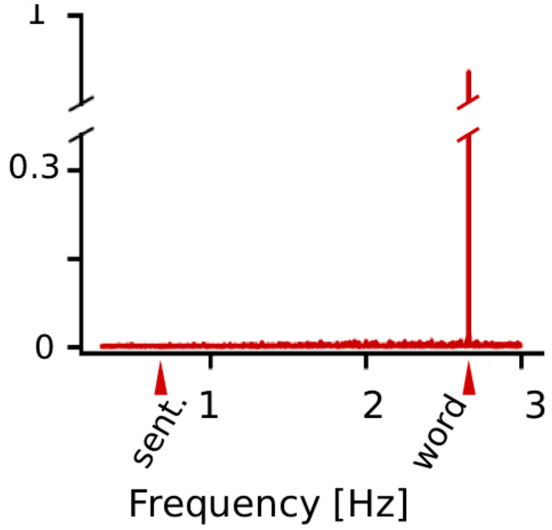


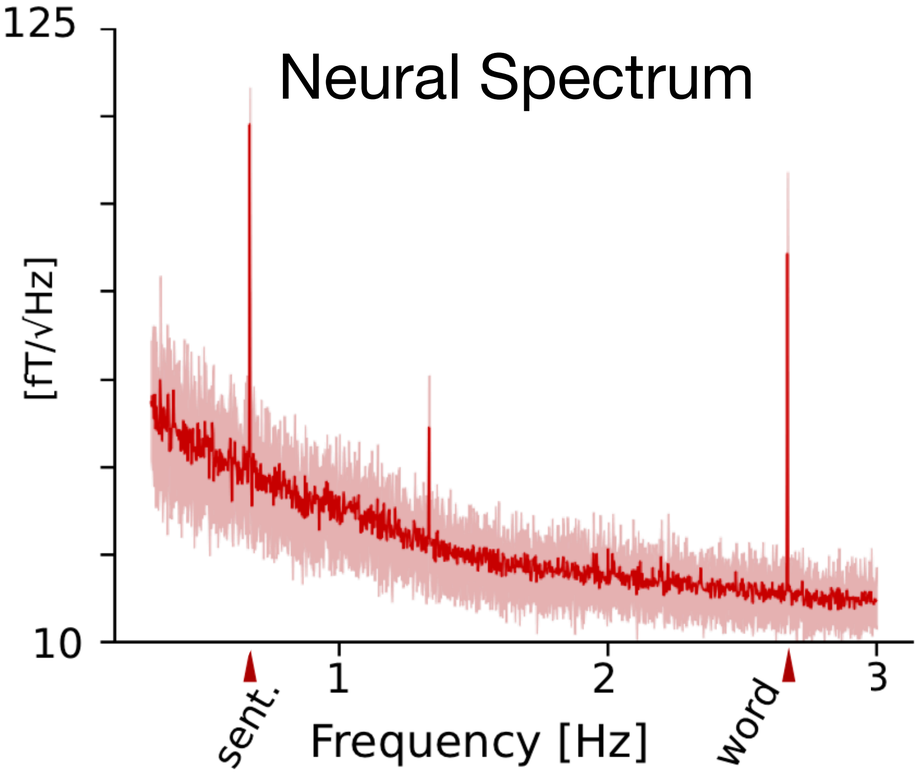
Isochronous Speech

Acoustics

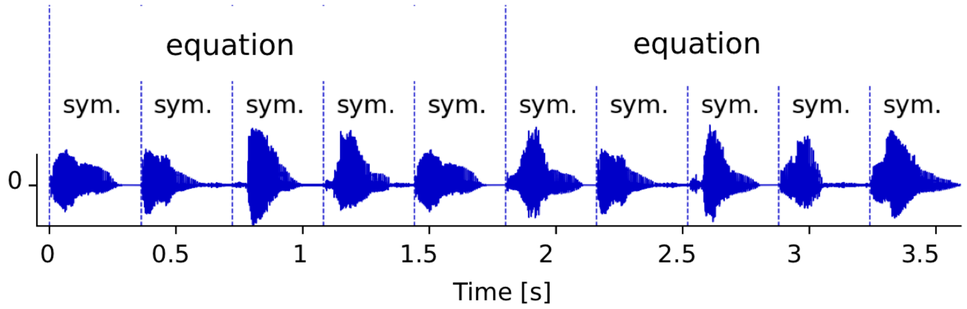


Acoustical Spectrum (envelope)

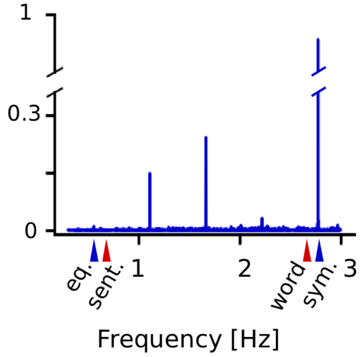




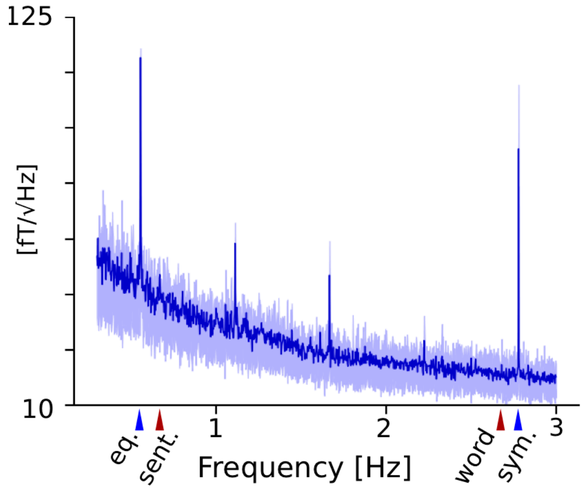
Isochronous Arithmetic



Acoustics



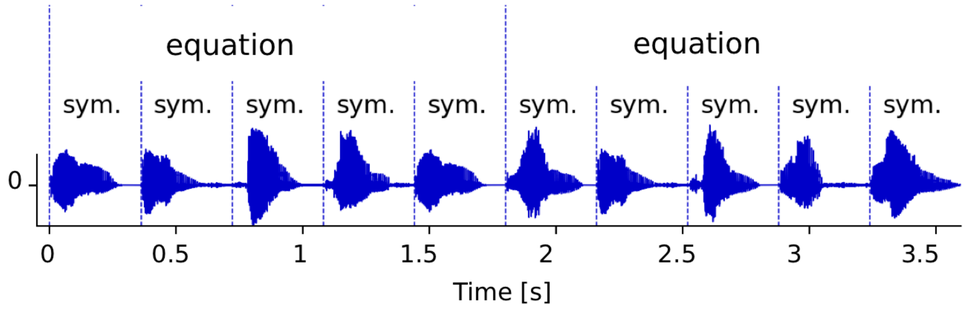
Acoustical Spectrum



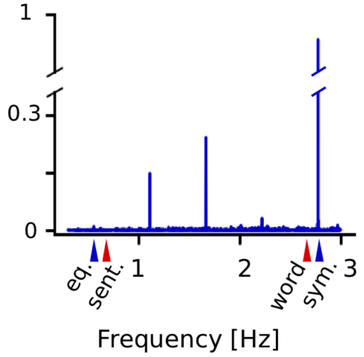
Neural Spectrum

Kulasingham et al. (2021) *Cortical Processing of Arithmetic and Simple Sentences …*, bioRxiv

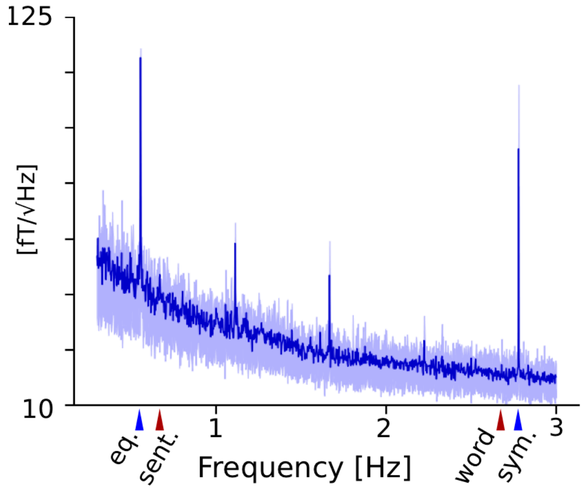
Isochronous Arithmetic



Acoustics



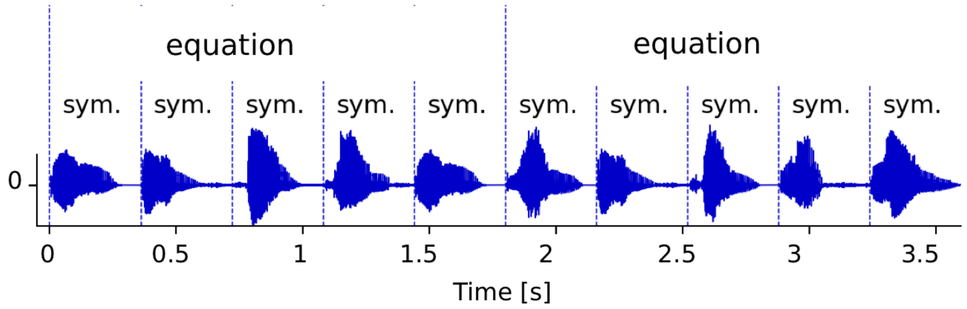
Acoustical Spectrum



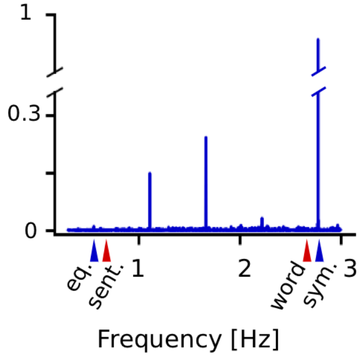
Neural Spectrum

Kulasingham et al. (2021) *Cortical Processing of Arithmetic and Simple Sentences …*, bioRxiv

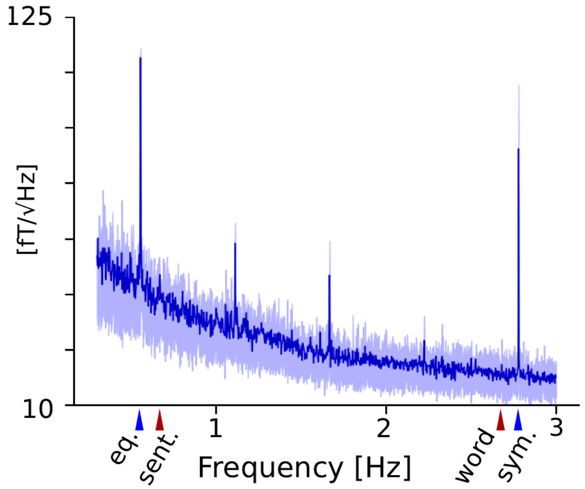
Isochronous Arithmetic



Acoustics



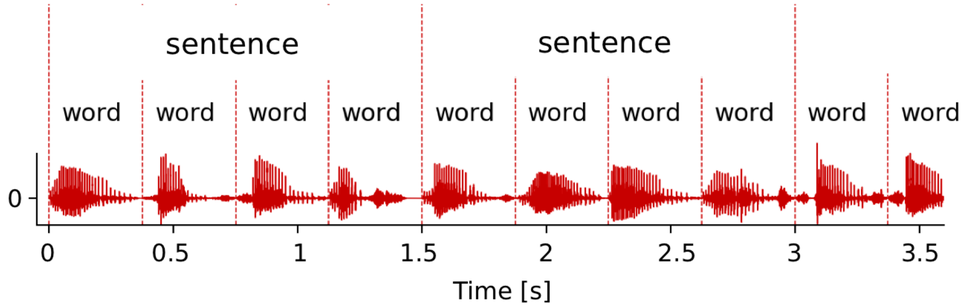
Acoustical Spectrum



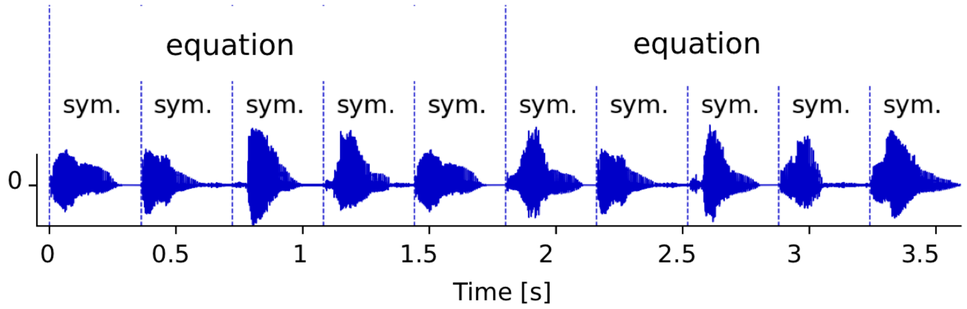
Neural Spectrum

Kulasingham et al. (2021) *Cortical Processing of Arithmetic and Simple Sentences …*, bioRxiv

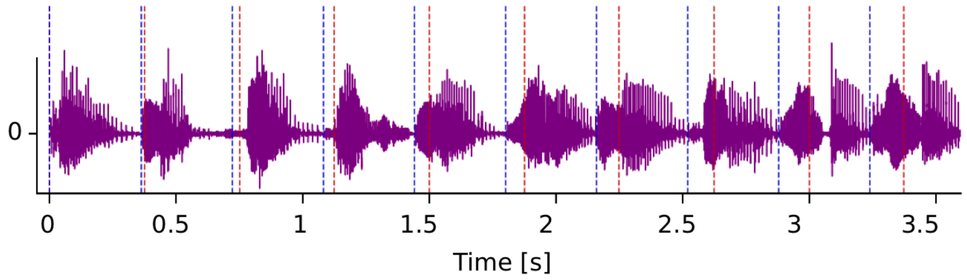
Isochronous Cocktail Party

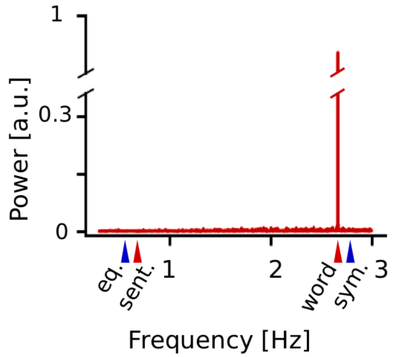


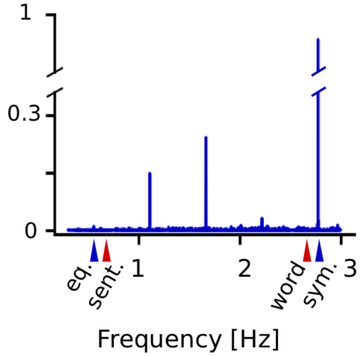
+

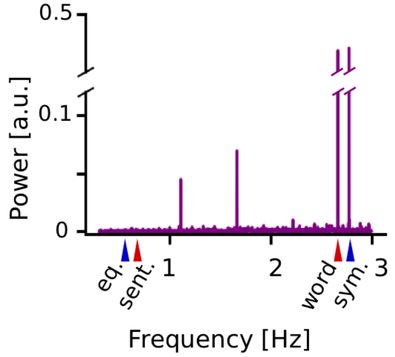


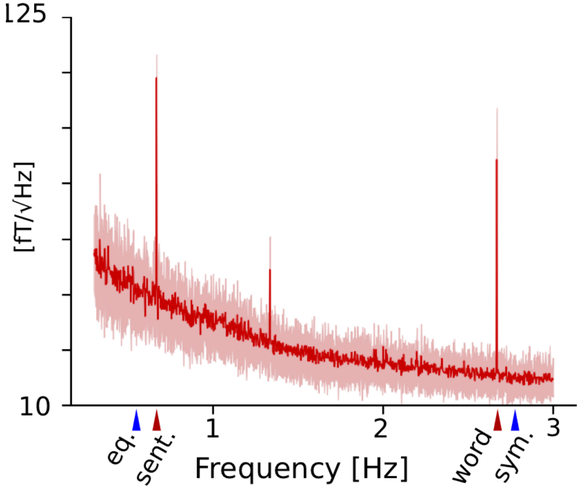
=

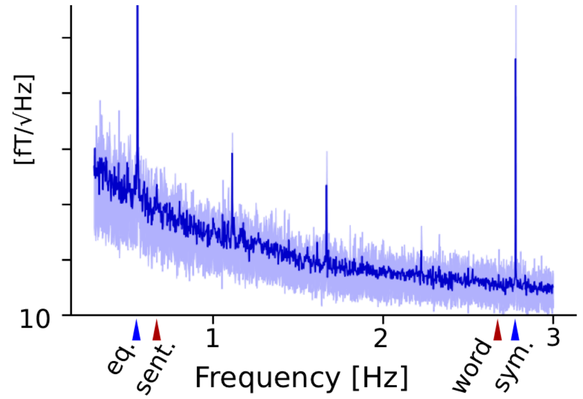






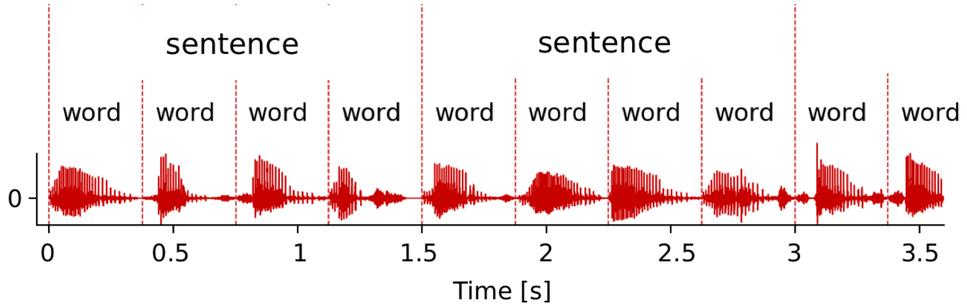




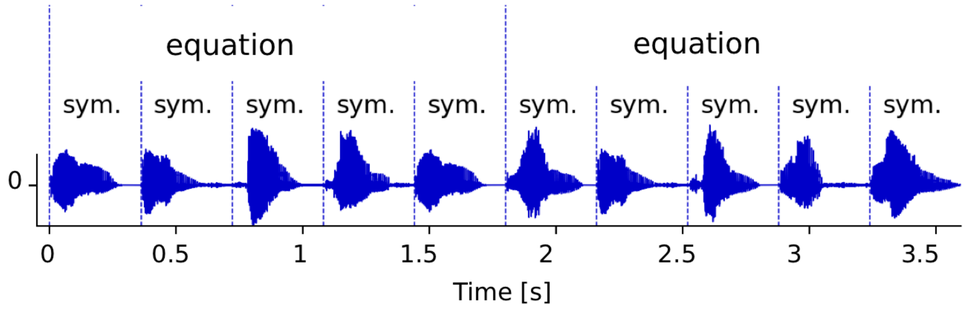


?

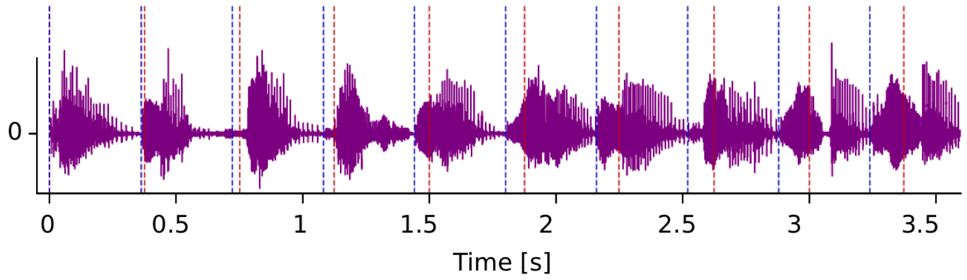
Isochronous Cocktail Party

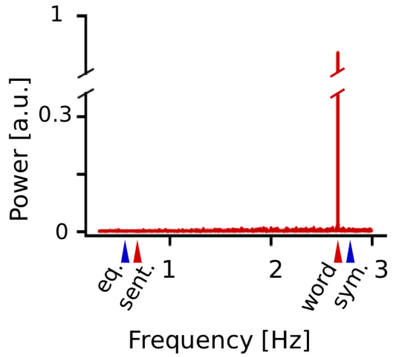


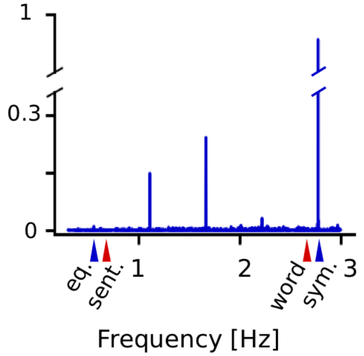
+

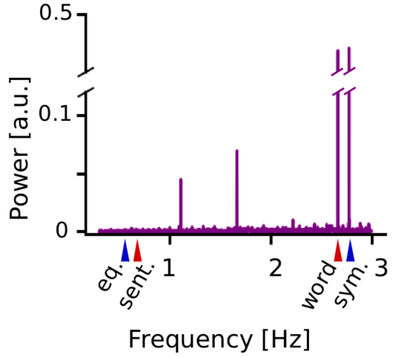


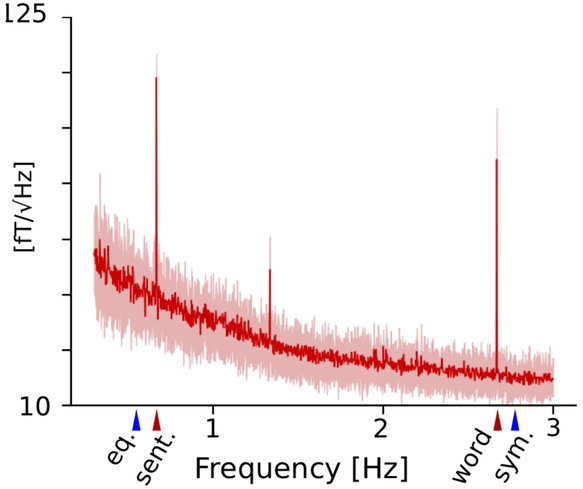
=

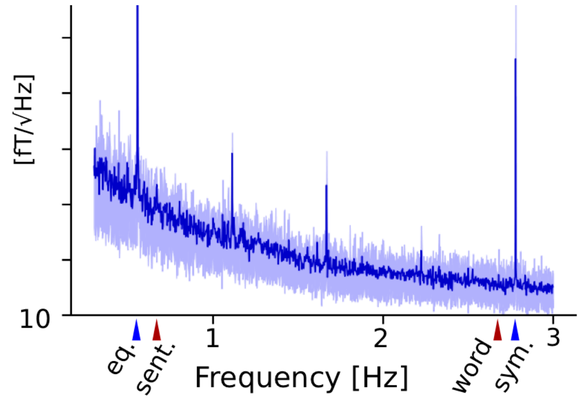






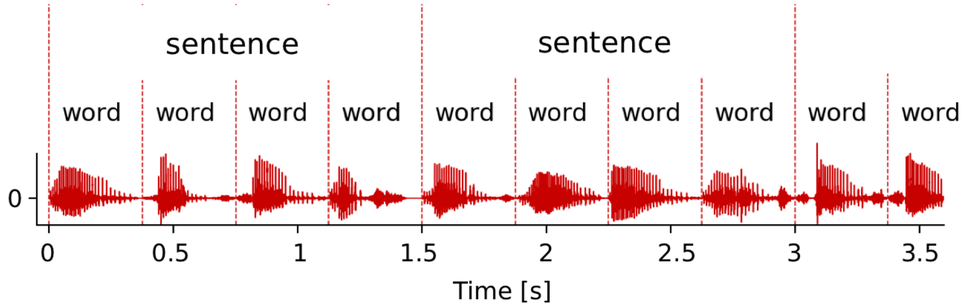




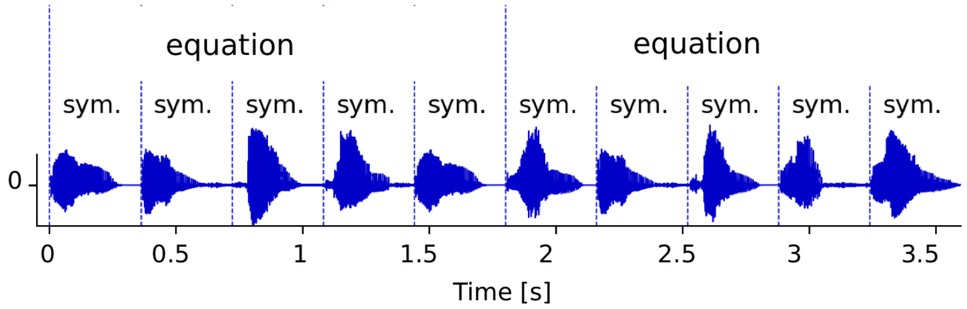


?

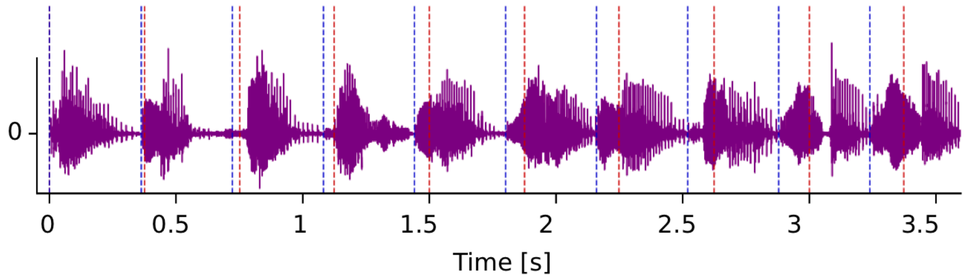
Isochronous Cocktail Party

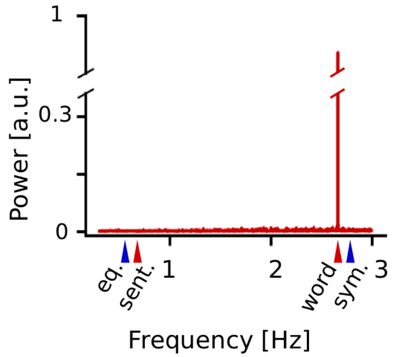


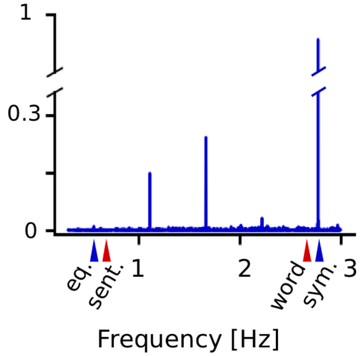
+

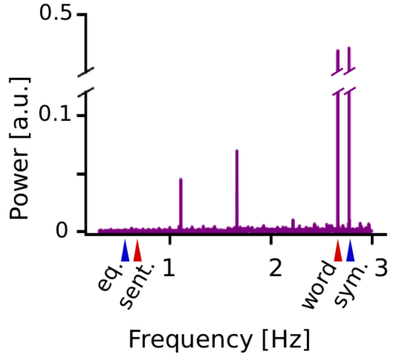


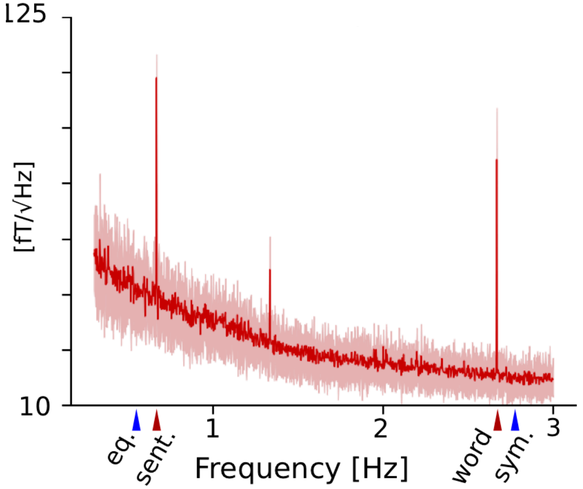
=

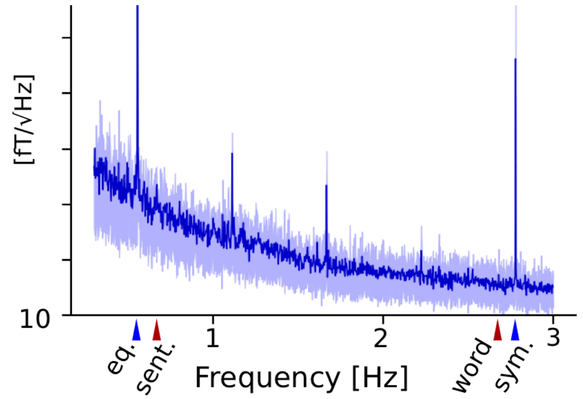










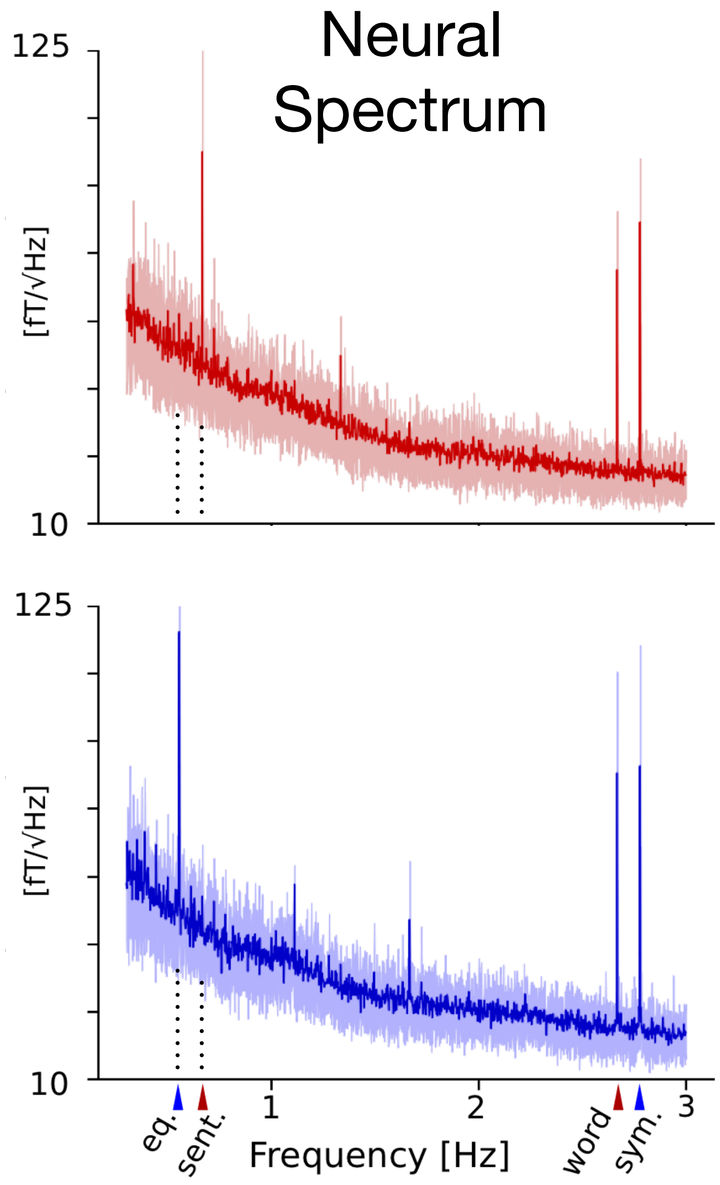


?

Isochronous Cocktail Party

Attend to Sentences

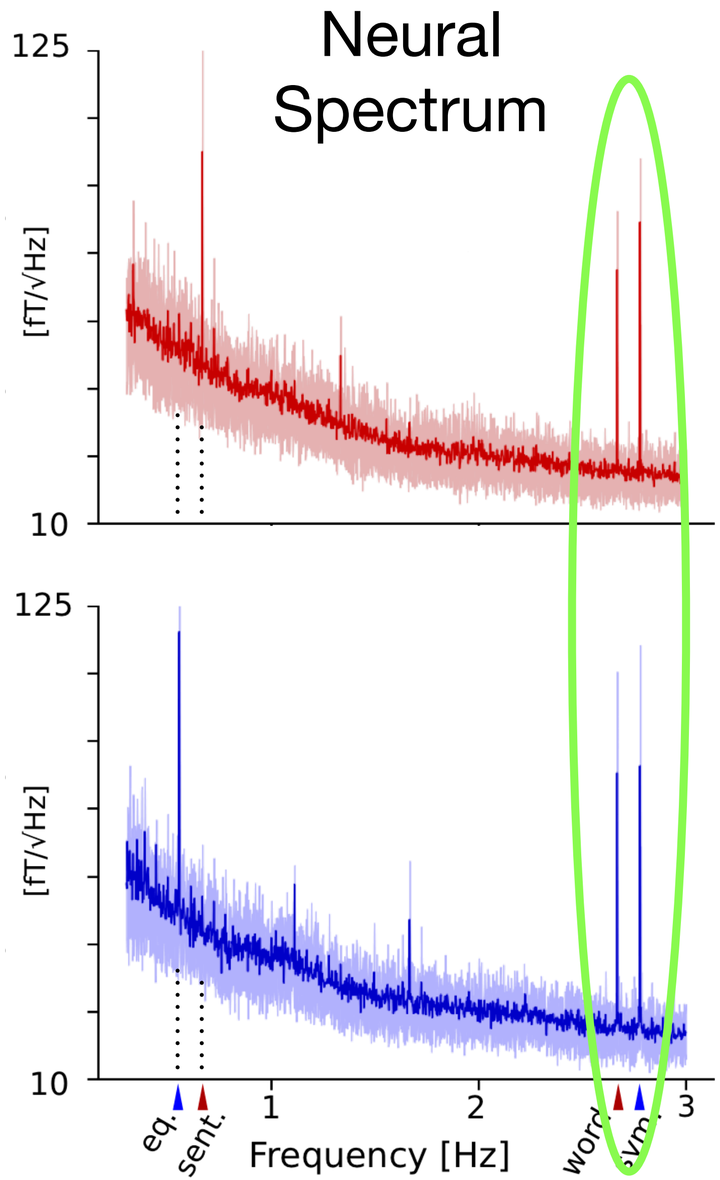
Attend to Equations



Isochronous Cocktail Party

Attend to Sentences

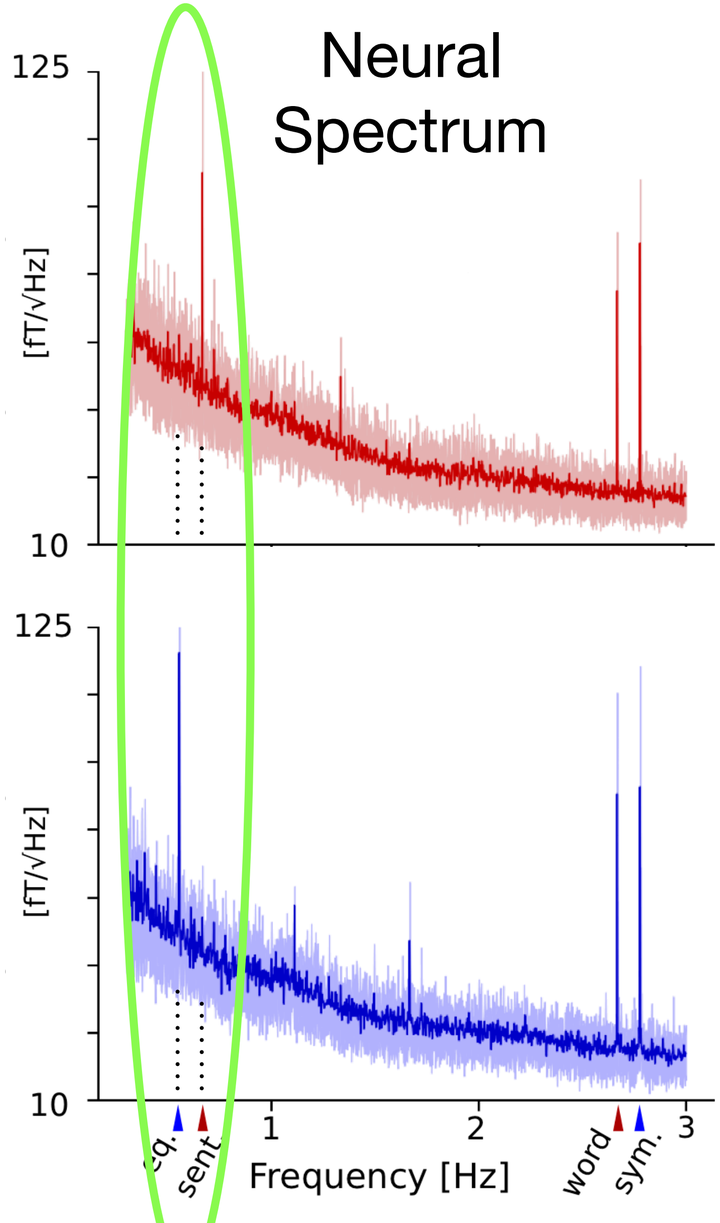
Attend to Equations



Isochronous Cocktail Party

Attend to Sentences

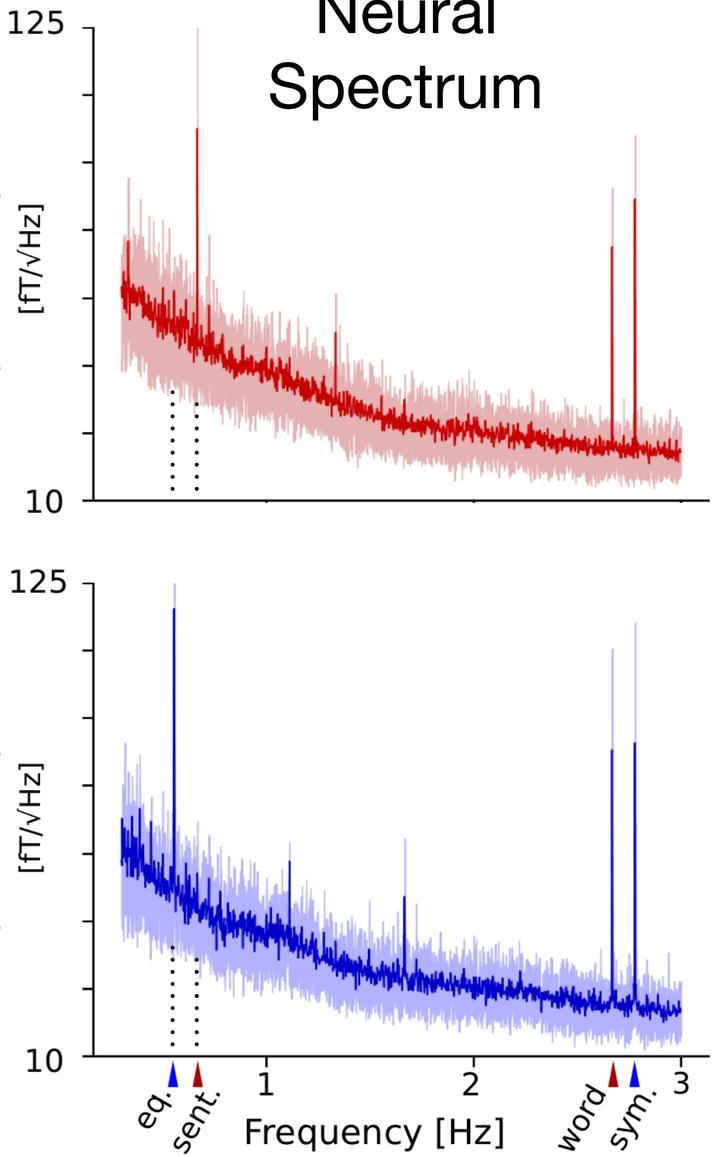
Attend to Equations

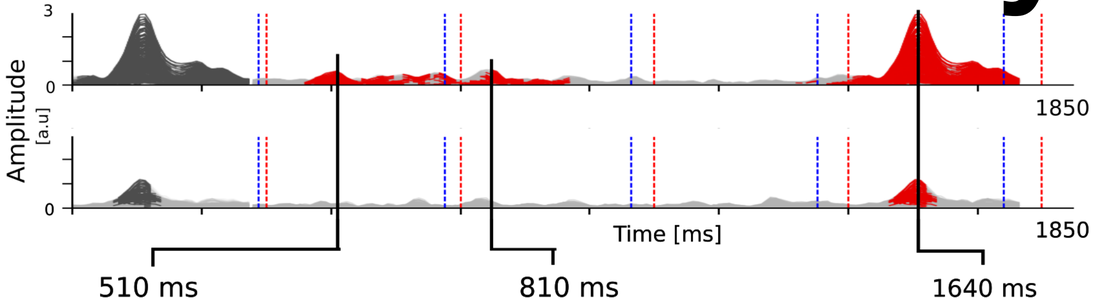


Isochronous Cocktail Party

Attend to Sentences

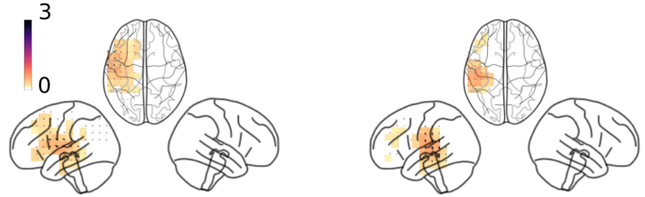
Attend to Equations

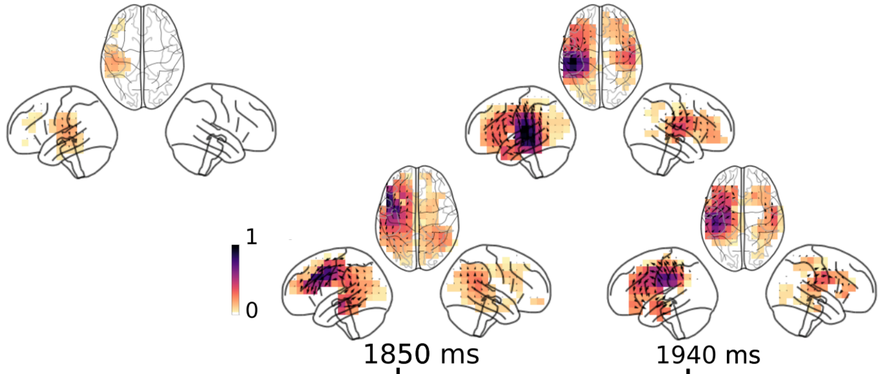




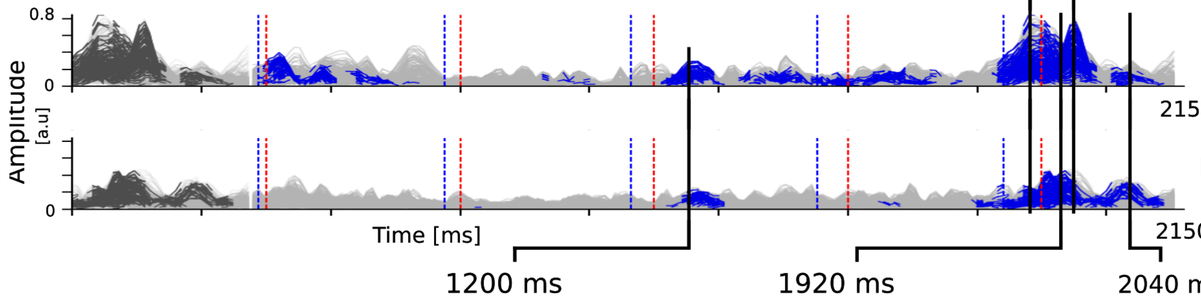
left hemisphere sources

right hemisphere sources



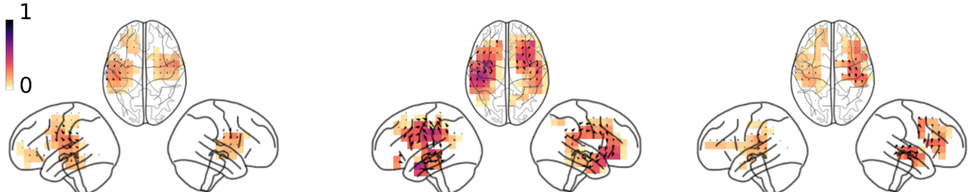


TRFs



left hemisphere sources

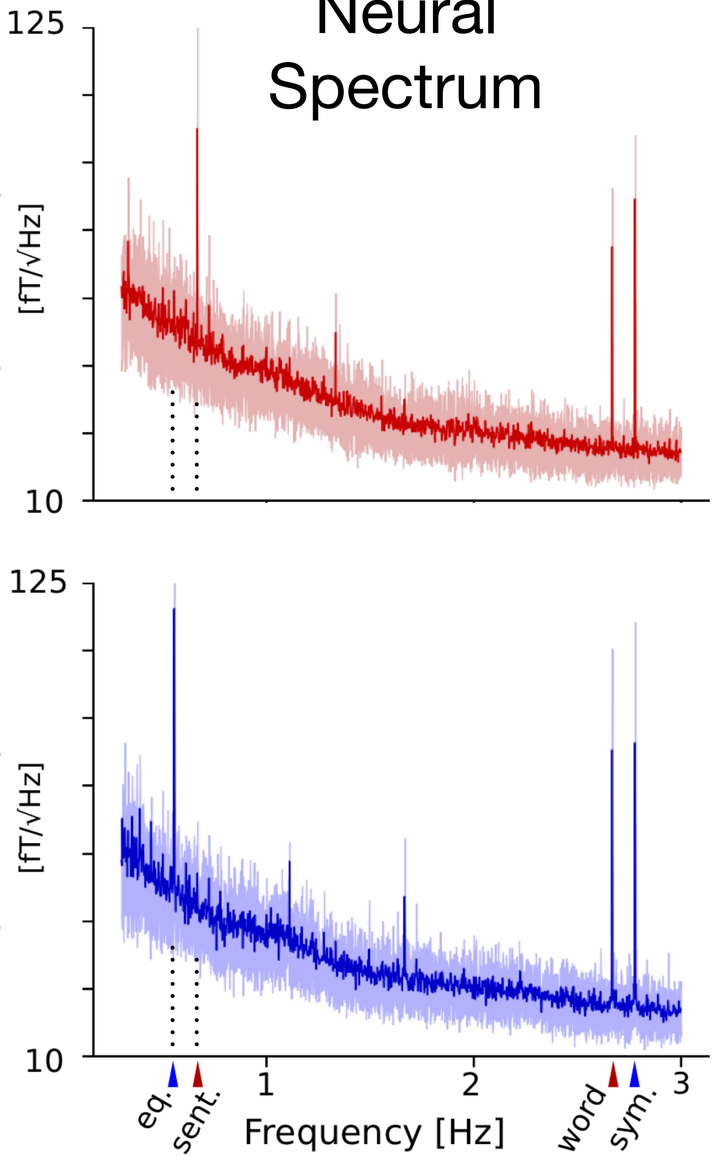
right hemisphere sources

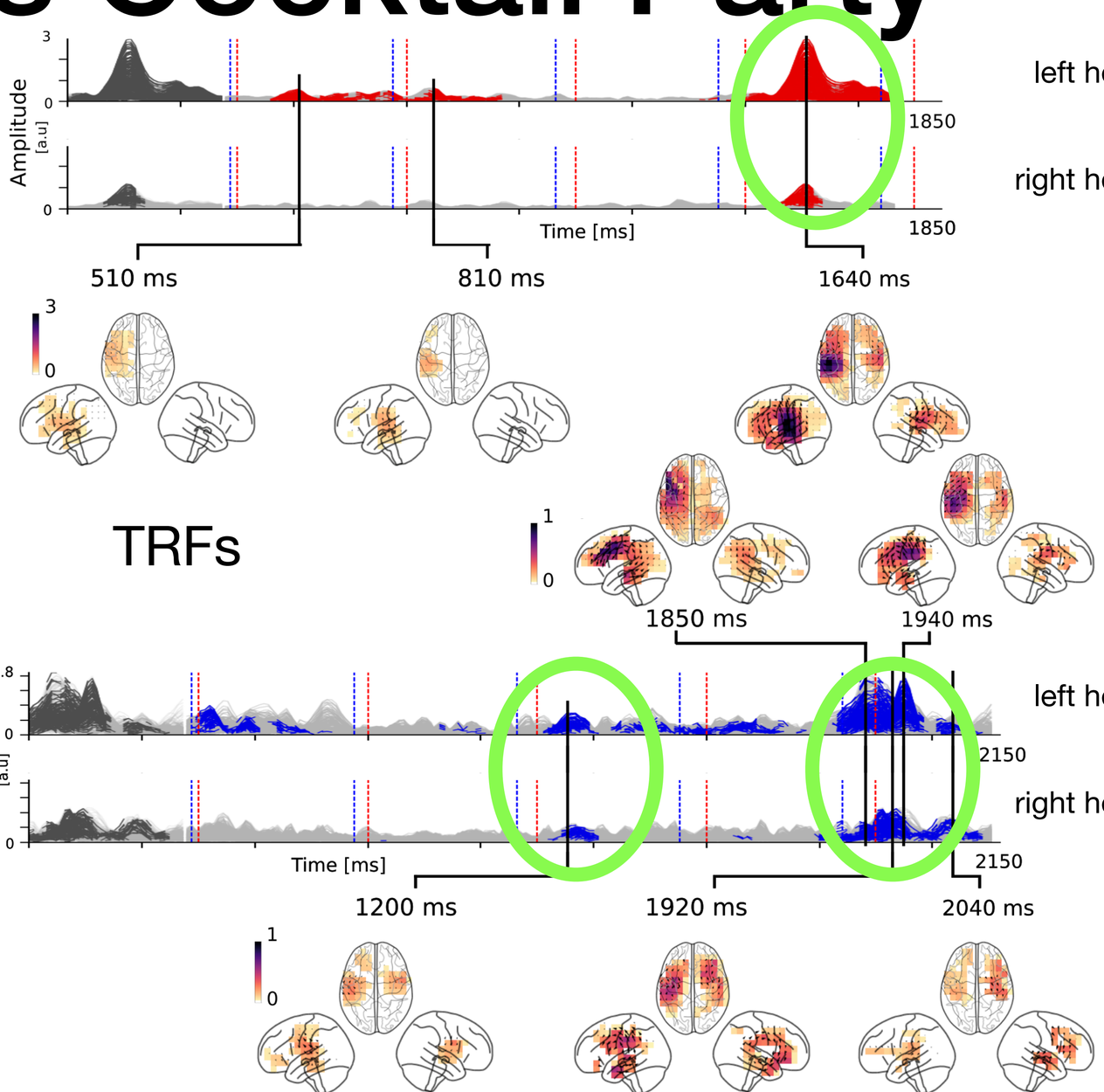


Isochronous Cocktail Party

Attend to Sentences

Attend to Equations

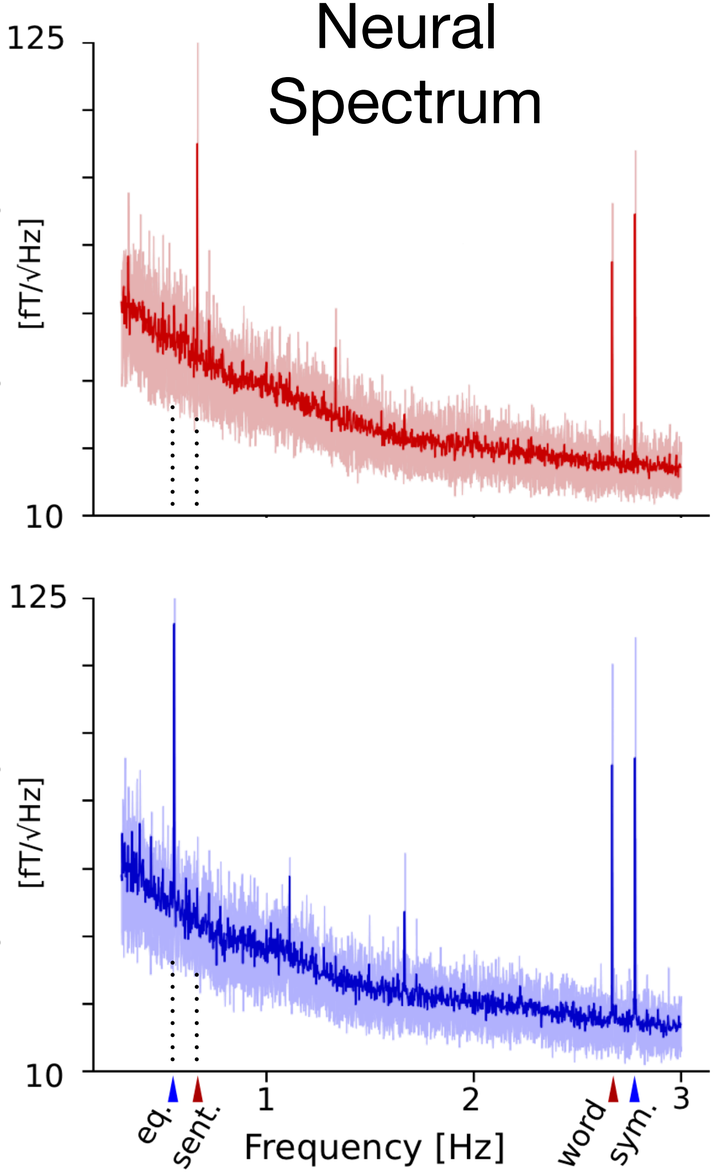


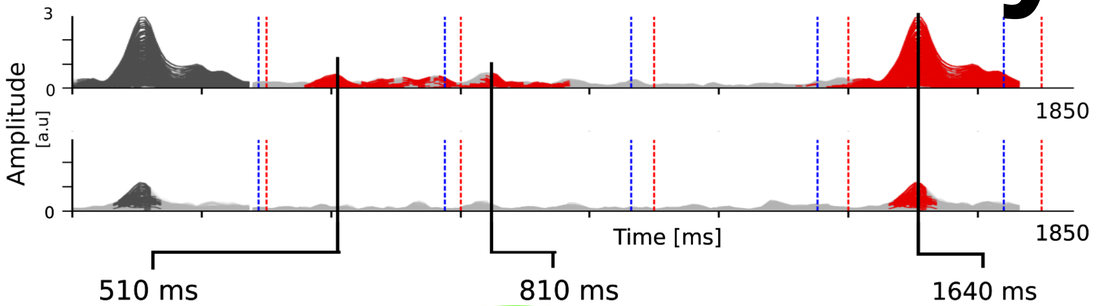


Isochronous Cocktail Party

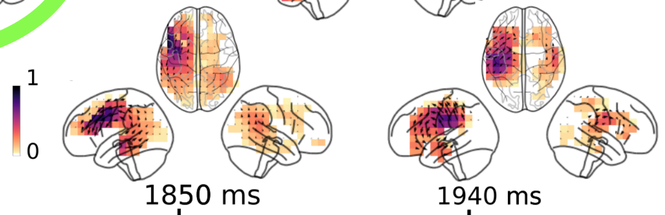
Attend to Sentences

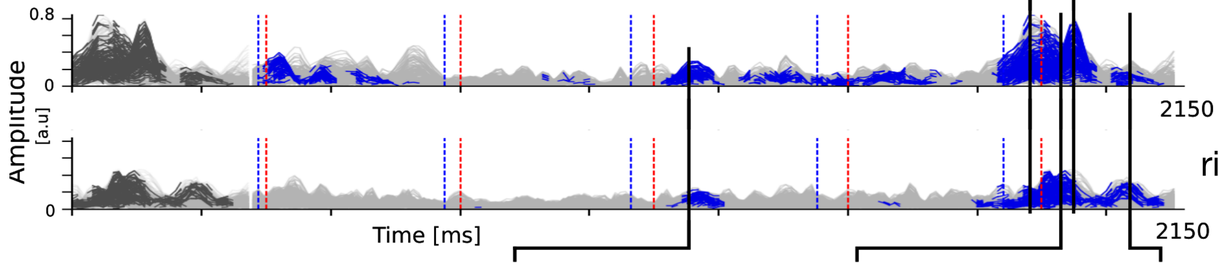
Attend to Equations

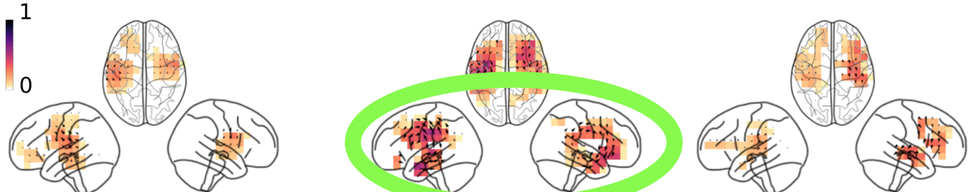




TRFs

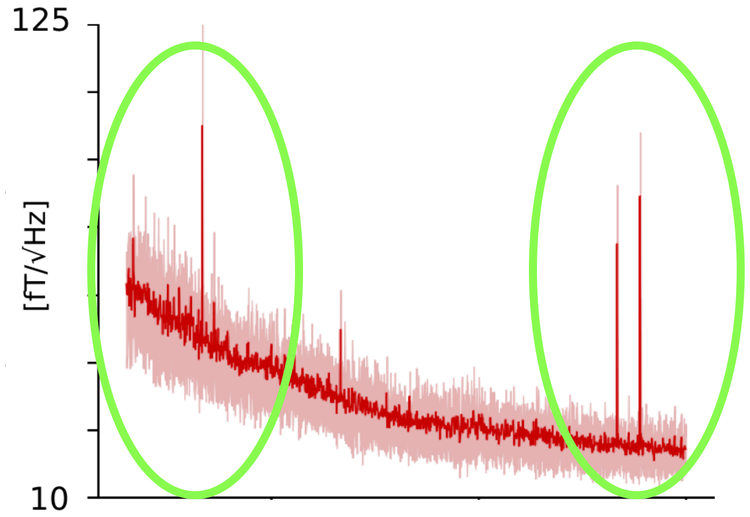




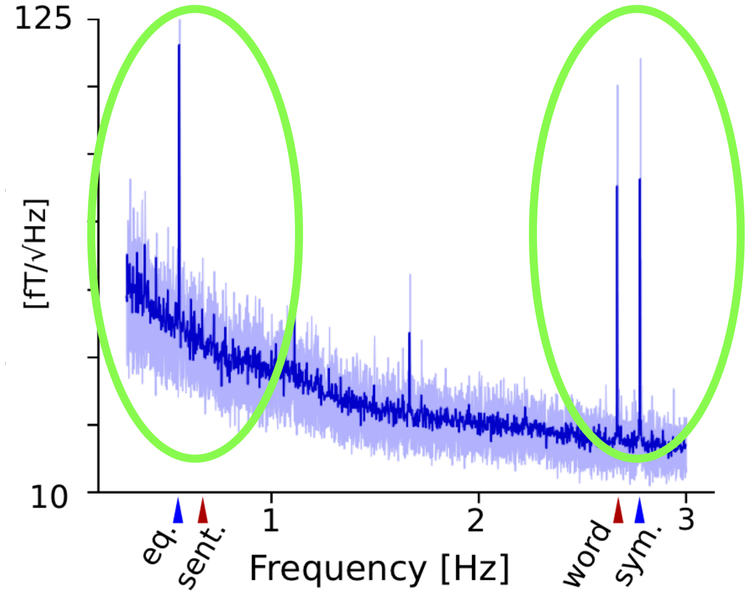


Representations of Understanding

Attend to Sentences

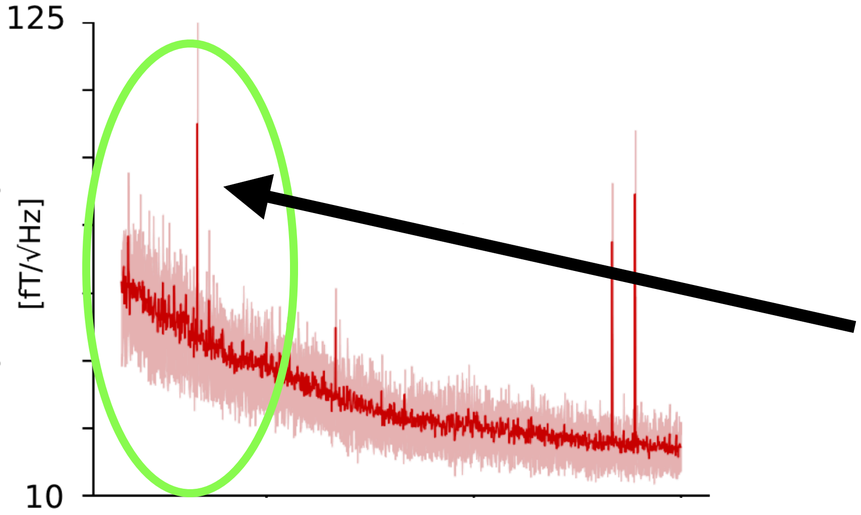


Attend to Equations

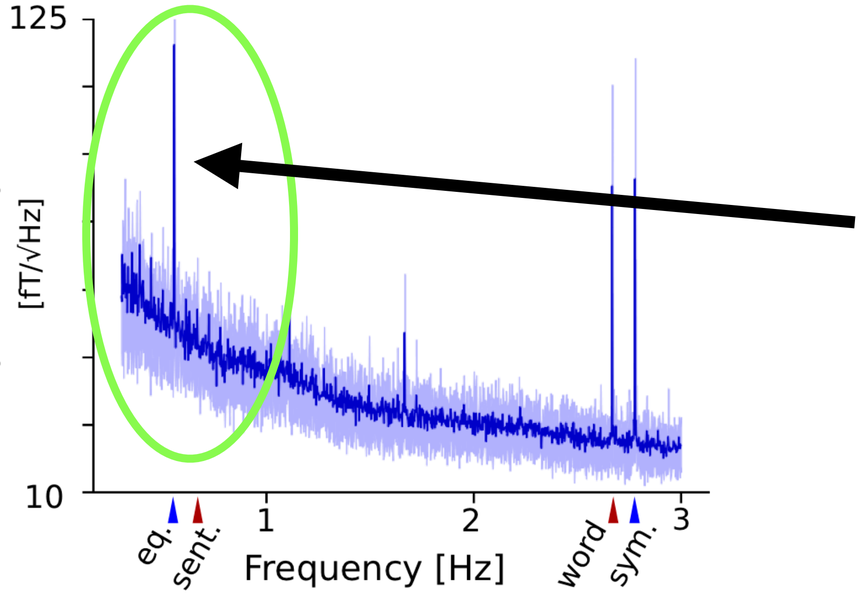


Representations of Understanding

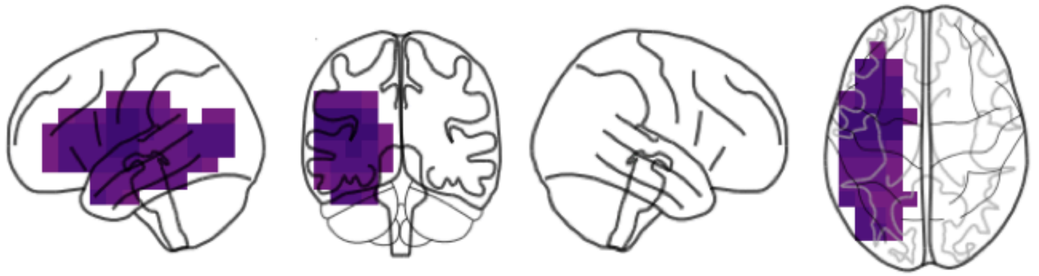
Attend to Sentences

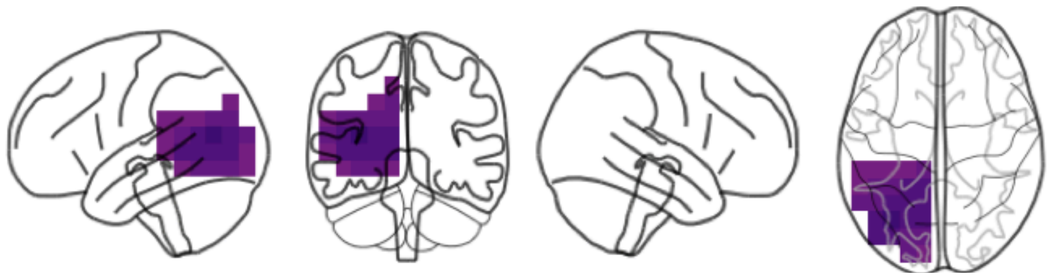


Attend to Equations



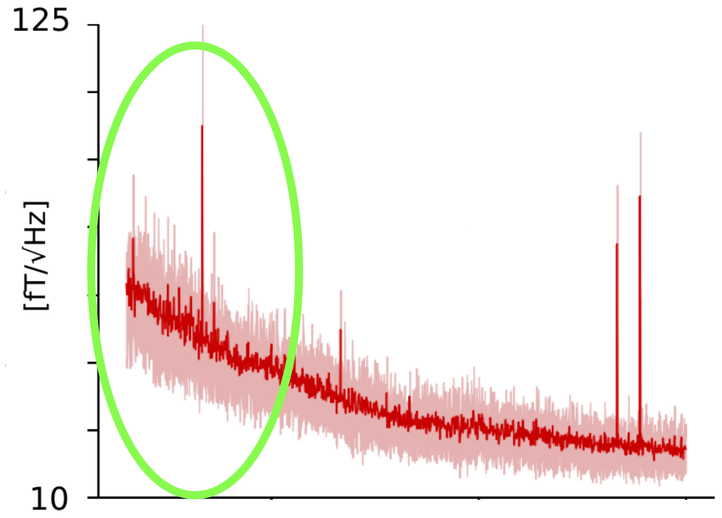
Neural Correlation with Behavior



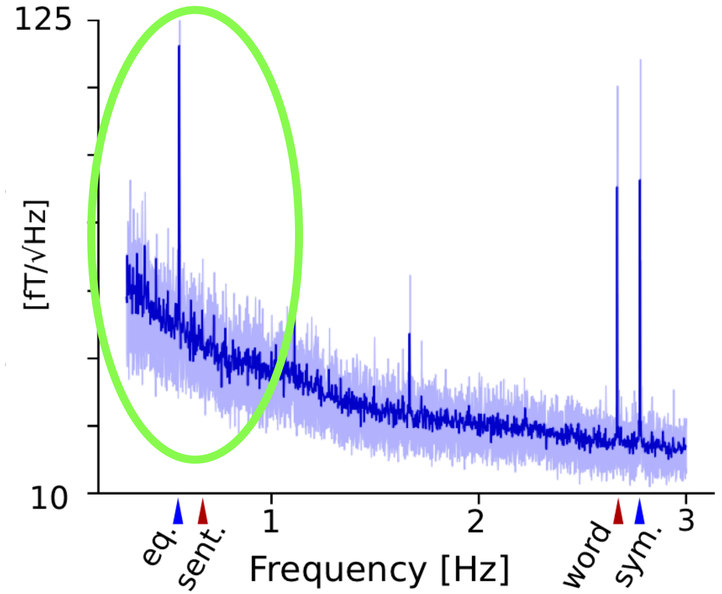


Representations of Understanding

Attend to Sentences



Attend to Equations



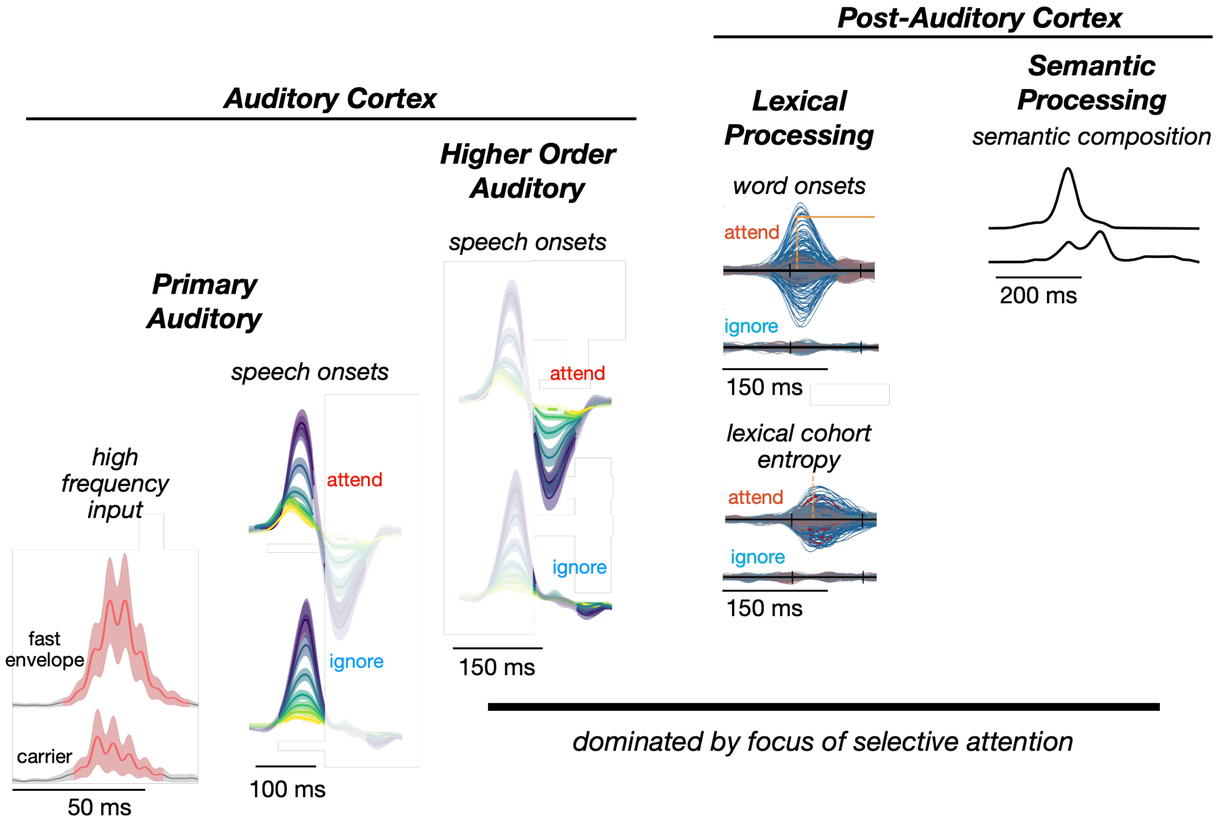
- Neural correlates of understanding
  - rhythms of higher order structures
    - sentence structures
    - poetic structures
    - mathematical structures
    - …

Summary

*temporal **neural** patterns* ⇆

*temporal patterns in **speech acoustics***
*temporal patterns in **speech perception***
*temporal patterns in **language perception***
*temporal patterns in **understanding***

- *Continuous speech* allows acquiring entire hierarchy from same stimulus
- Using *simultaneous TRFs* allows segregation of neural processes
- How is each process linked to intelligibility/understanding?
- Which links are predictive/causal?



thank you

http://www.isr.umd.edu/Labs/CSSL/simonlab