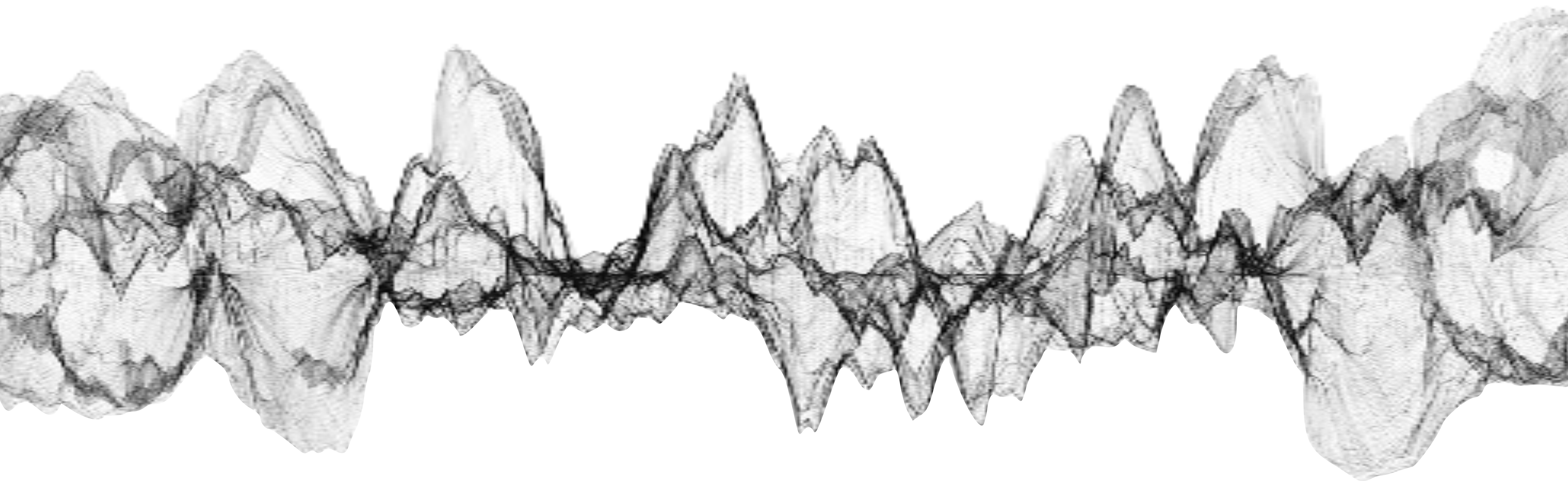
NEW YORK UNIVERSITY

In spoken word recognition, the future predicts the past

Laura Gwilliams

5th June 2017



Road Map

Completed Research:

- Sensitivity to **phonological ambiguity** is reflected in the very initial stages (~50 ms) of processing a speech sound
- Sub-phonemic information is **maintained** for long periods of time, and is **re-evoked** at subsequent phoneme positions in the spoken word
- The system **commits to phonological interpretations** on a shorter time-scale in parallel to phonetic maintenance

Future Directions:

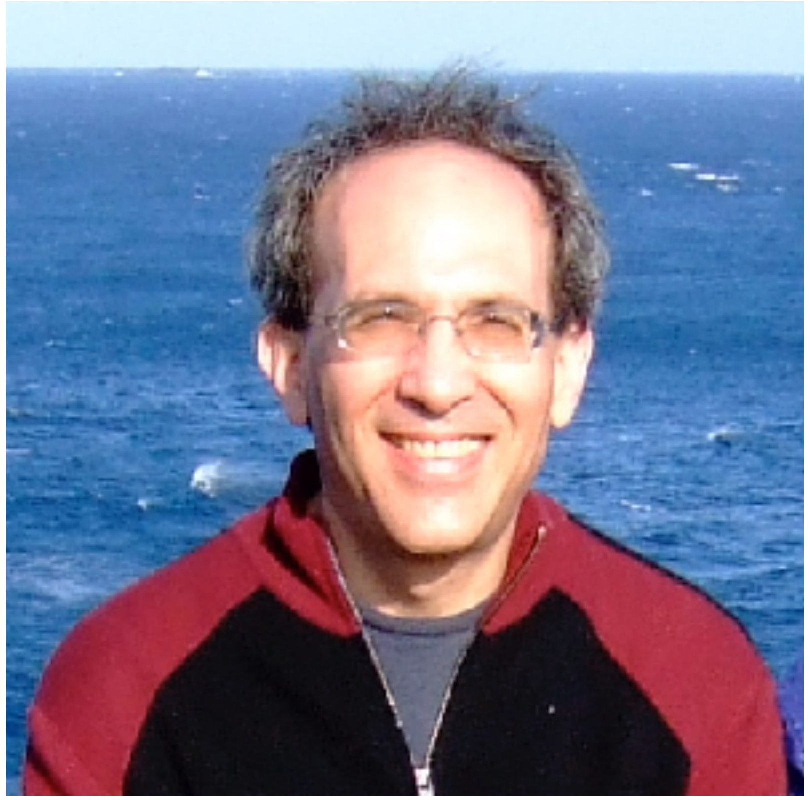
- Can we apply **machine-learning** analysis techniques to MEG data to unveil the dynamics with which sub-phonemic information is processed?

Recognition & Resolution of Phoneme Ambiguity in Spoken Words

Collaborators

Tal Linzen

David Poeppel



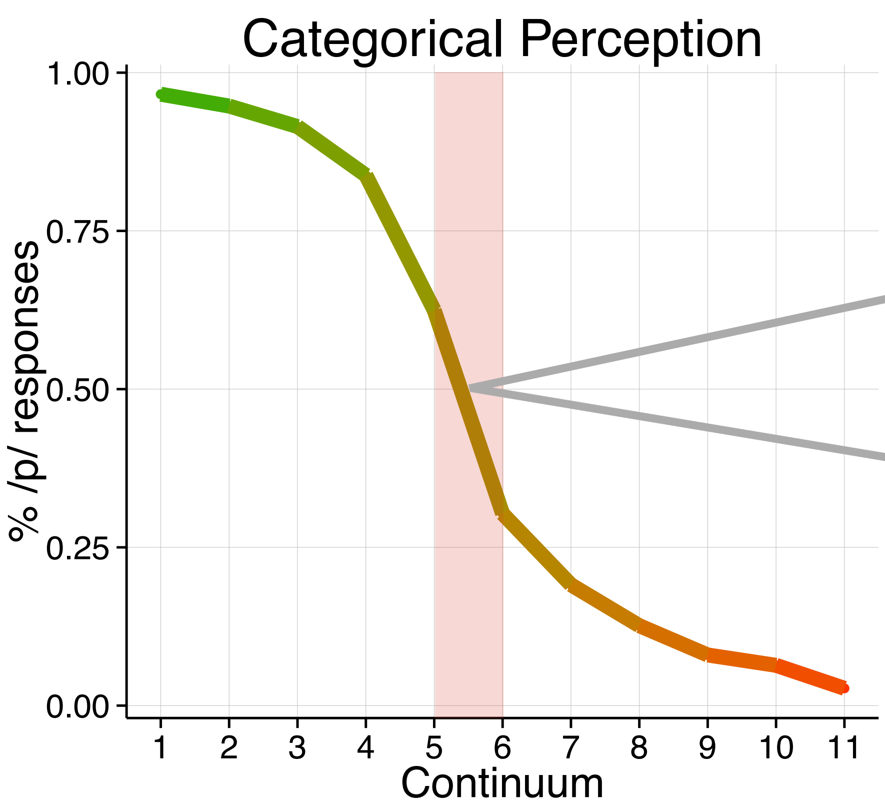
Alec Marantz

Gwilliams, Linzen, Poeppel & Marantz (Submitted)

Future Influences on Perception

- Speech is an inherently **noisy and ambiguous** signal
- To fluently derive meaning, listeners must **integrate top-down** contextual information to guide their interpretation
- Top-down input occurring *after* an acoustic signal can be integrated to **affect the perception of earlier sounds** (Bicknell et al., submitted; Connine et al., 1991; Samuel, 1981; Szostak & Pitt, 2013; Warren & Sherman, 1974)

Future Influences on Perception



p b b b b

(this is a parakeet)

p ee t

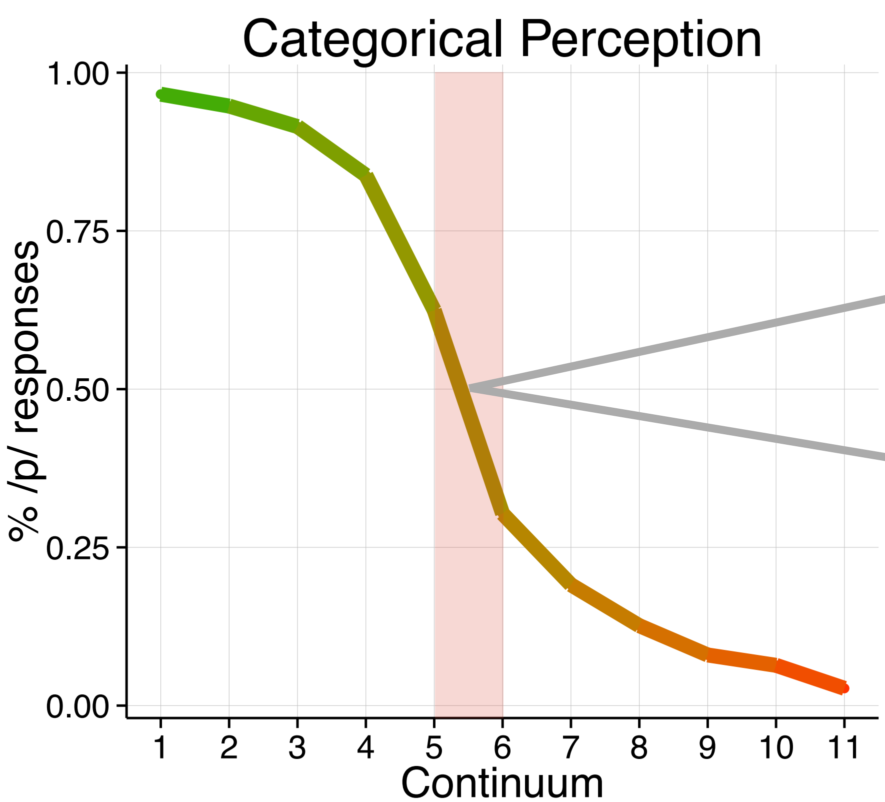
b ai d

(this is a barricade)

“**P**oint **o**f **D**isambiguation” (POD)

Future Influences on Perception

(this is a parakeet)



p b b b b

p a r a k **ee** t

b a r a k **ai** d

(this is a barricade)

“**P**oint **of** **D**isambiguation” (POD)

Today's Questions

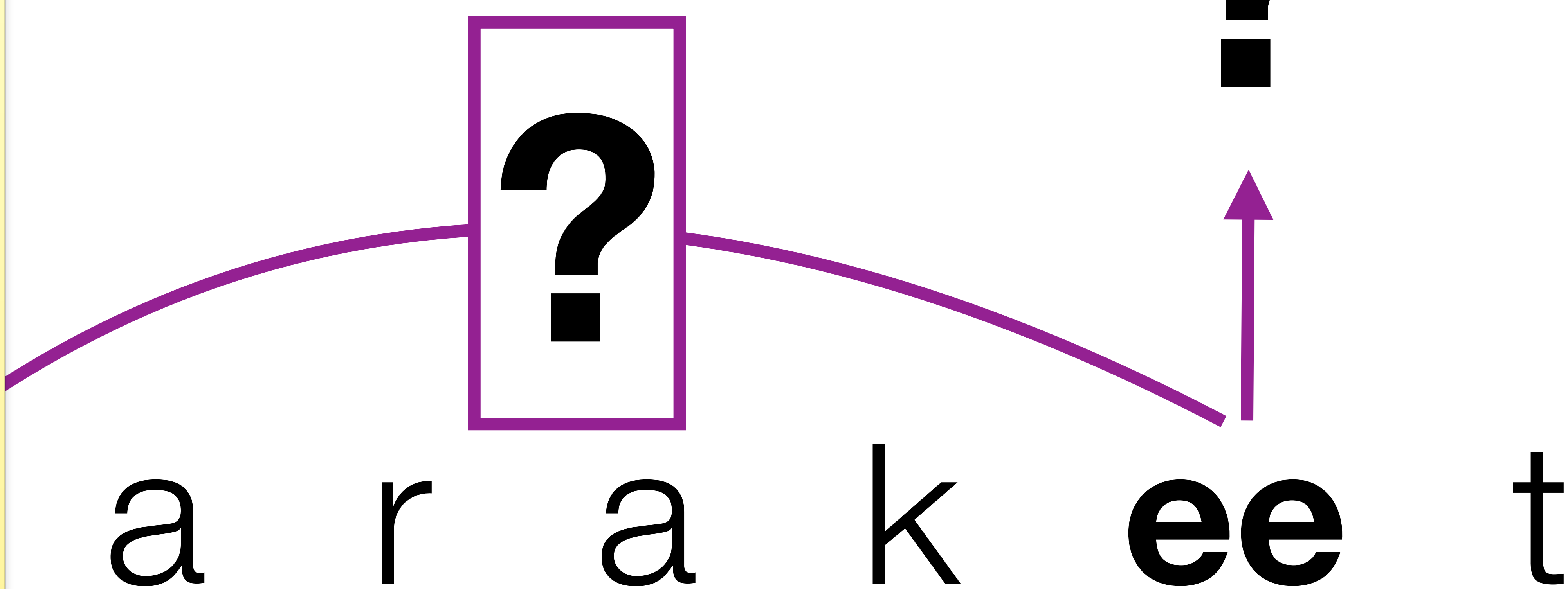
How does the auditory cortex **respond** to phonological ambiguity?

?

What is the **time-limit** on how late subsequent context can be received?

What are the neural signatures of ambiguity **resolution**?

?

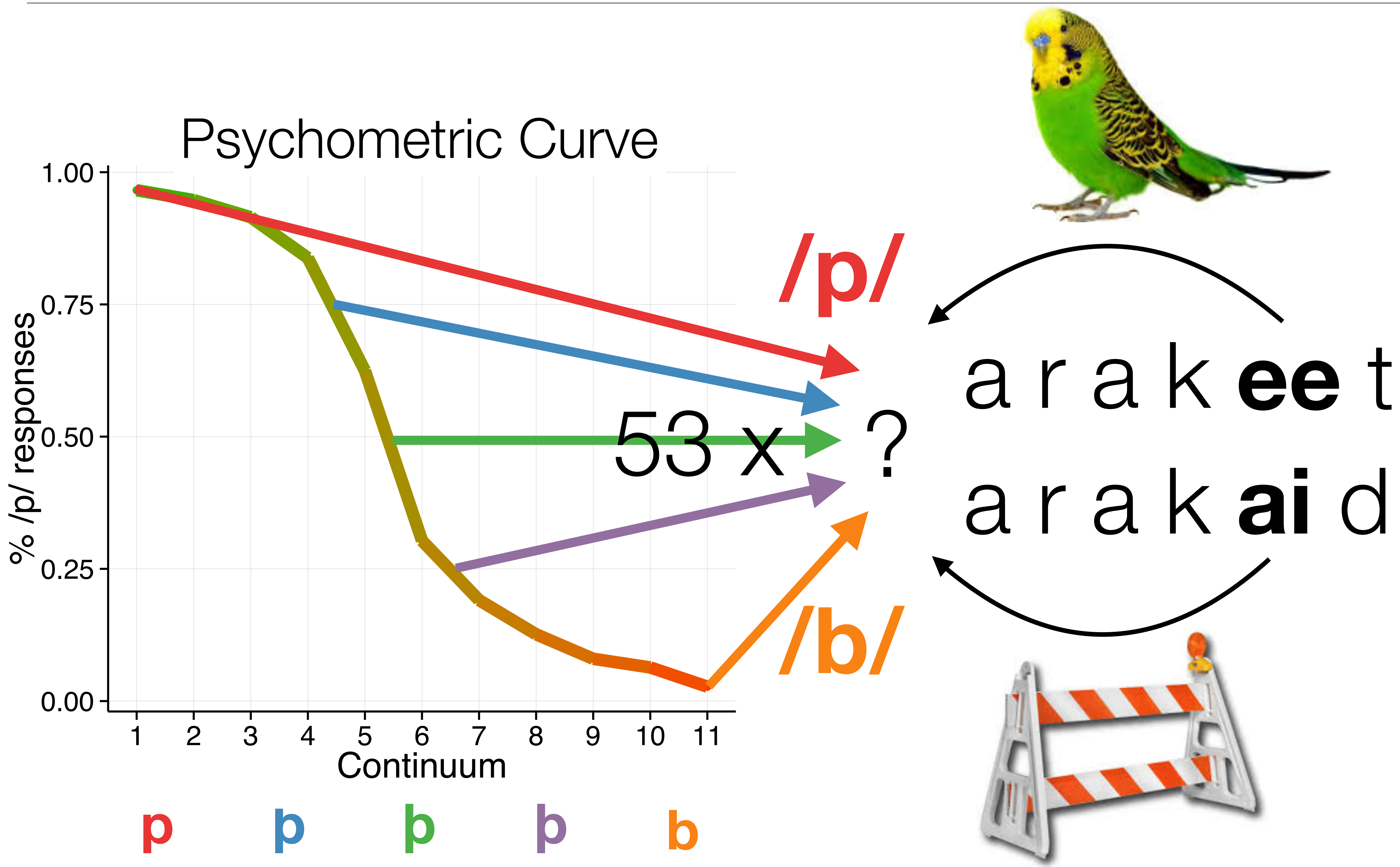


uity surface? most studies looking at
nri, but here we can associate it with

n be an open question.

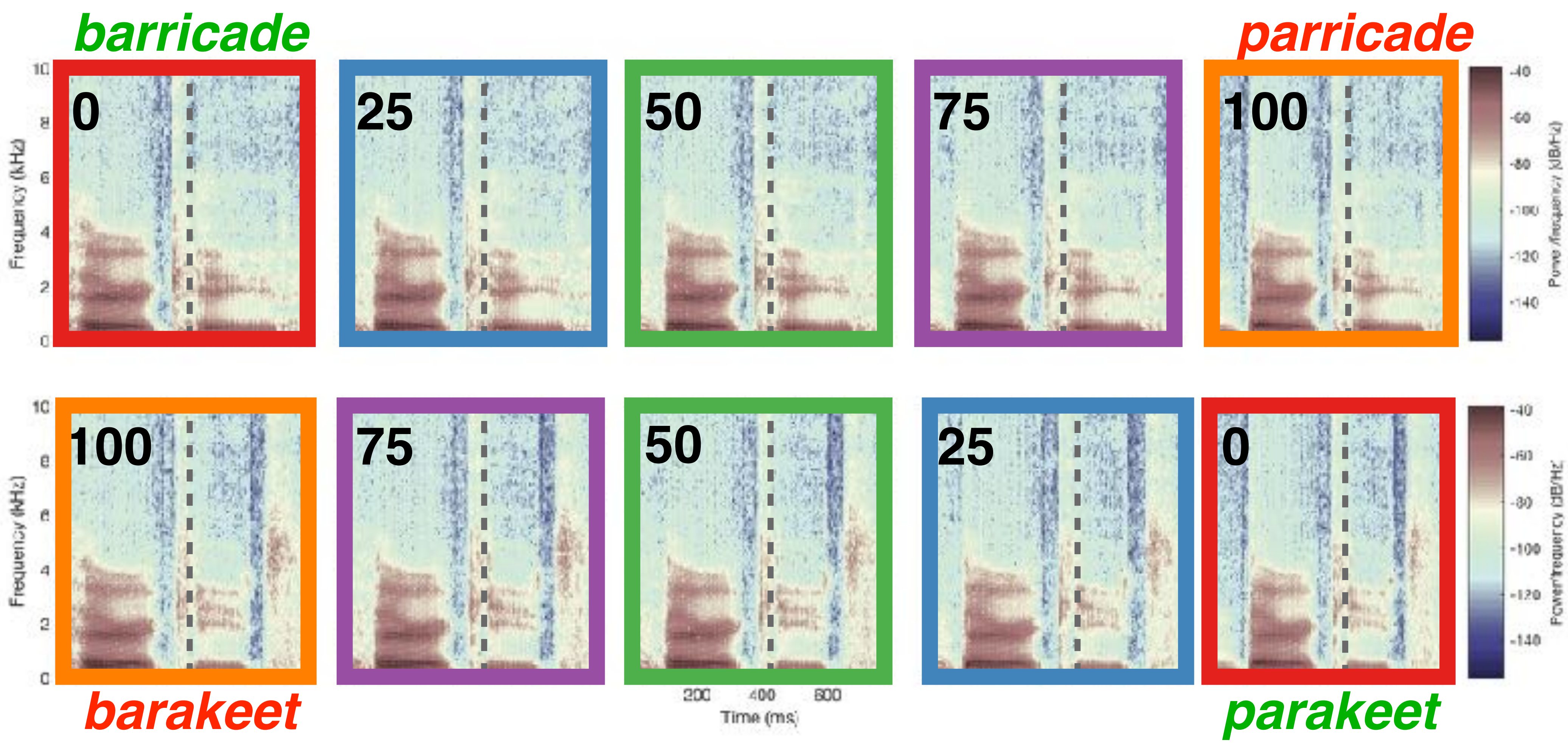
pute as to when it is. resolve this by
onetic maintenance. here i will argue
d are not mutually exclusive as has
e but important point, which i will get

Design & Materials



(see also McMurray et al., 2009)

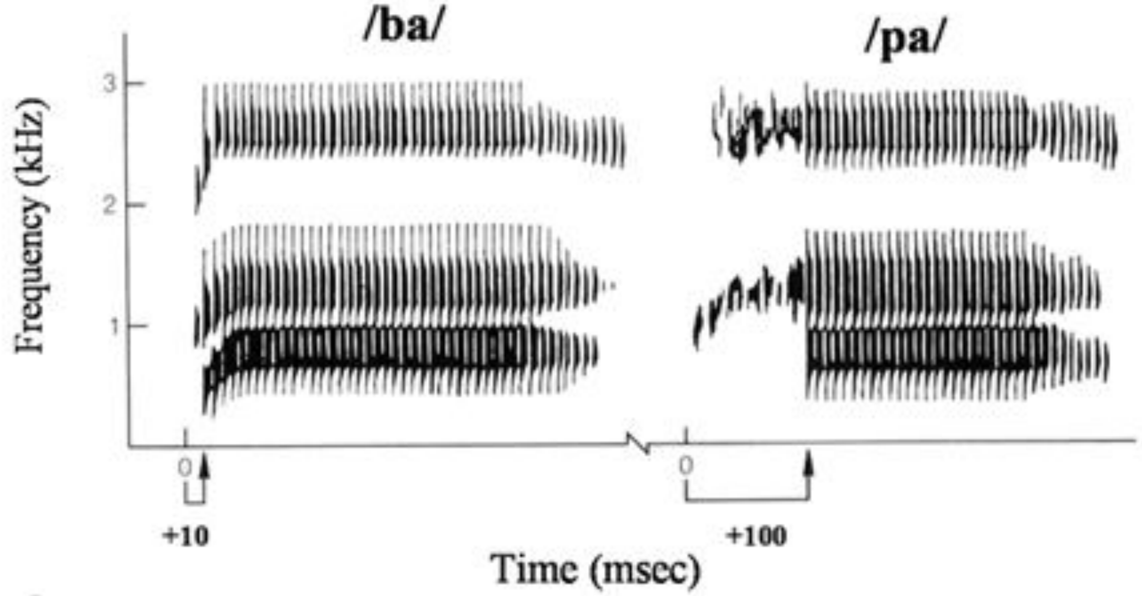
Design & Materials



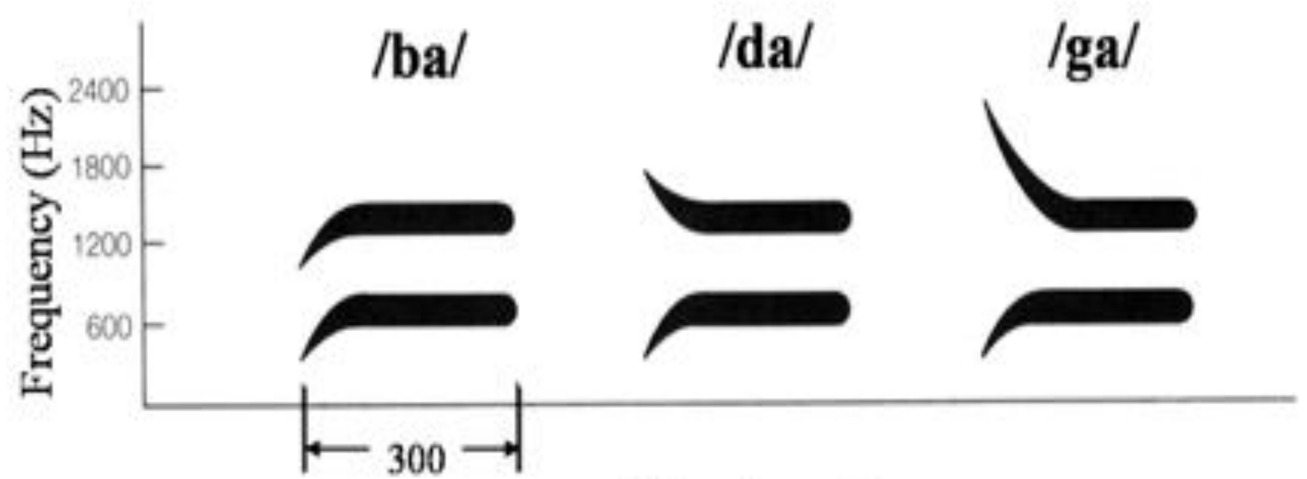
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms
- VOT (31 pairs) {p-b, t-d, k-g} and POA (22 pairs) {t-k, p-t}

Design & Materials

Contrast: Voice Onset Time (VOT)

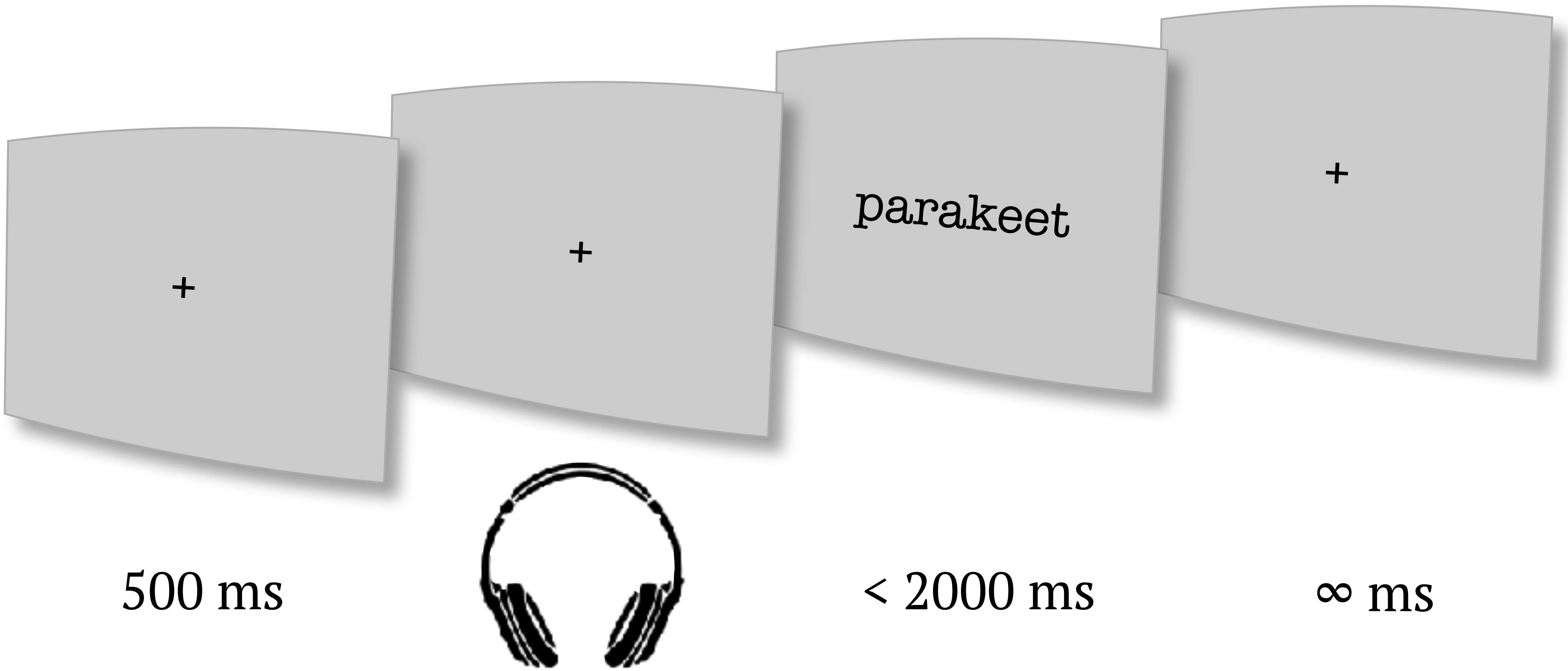


Contrast: Place of Articulation (POA)



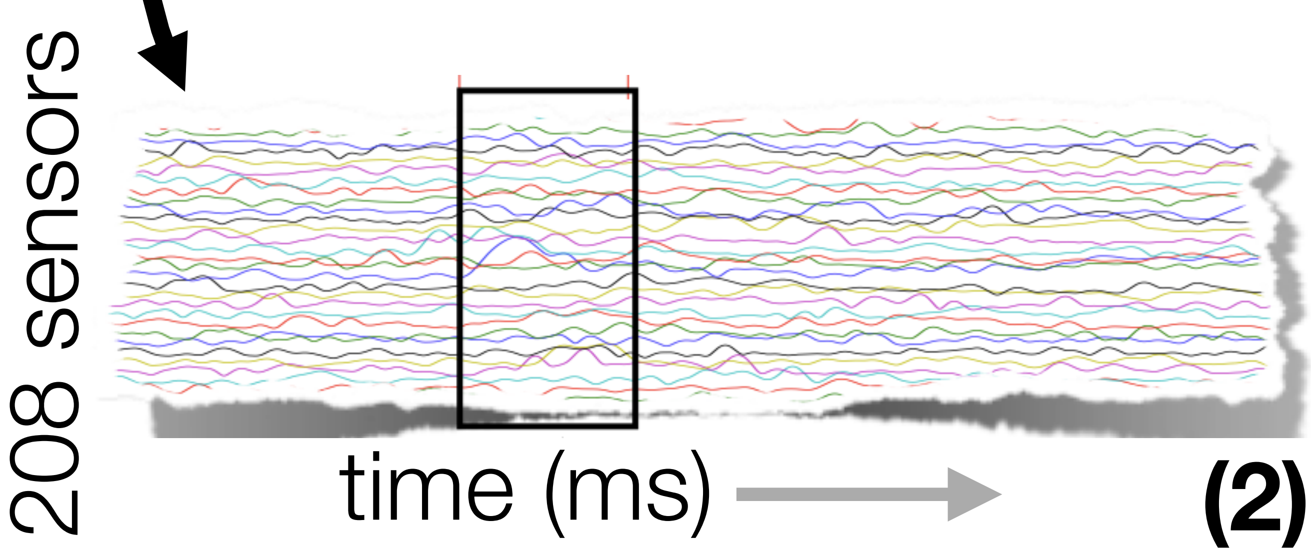
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms
- VOT (31 pairs) {p-b, t-d, k-g} and POA (22 pairs) {t-k, p-t}

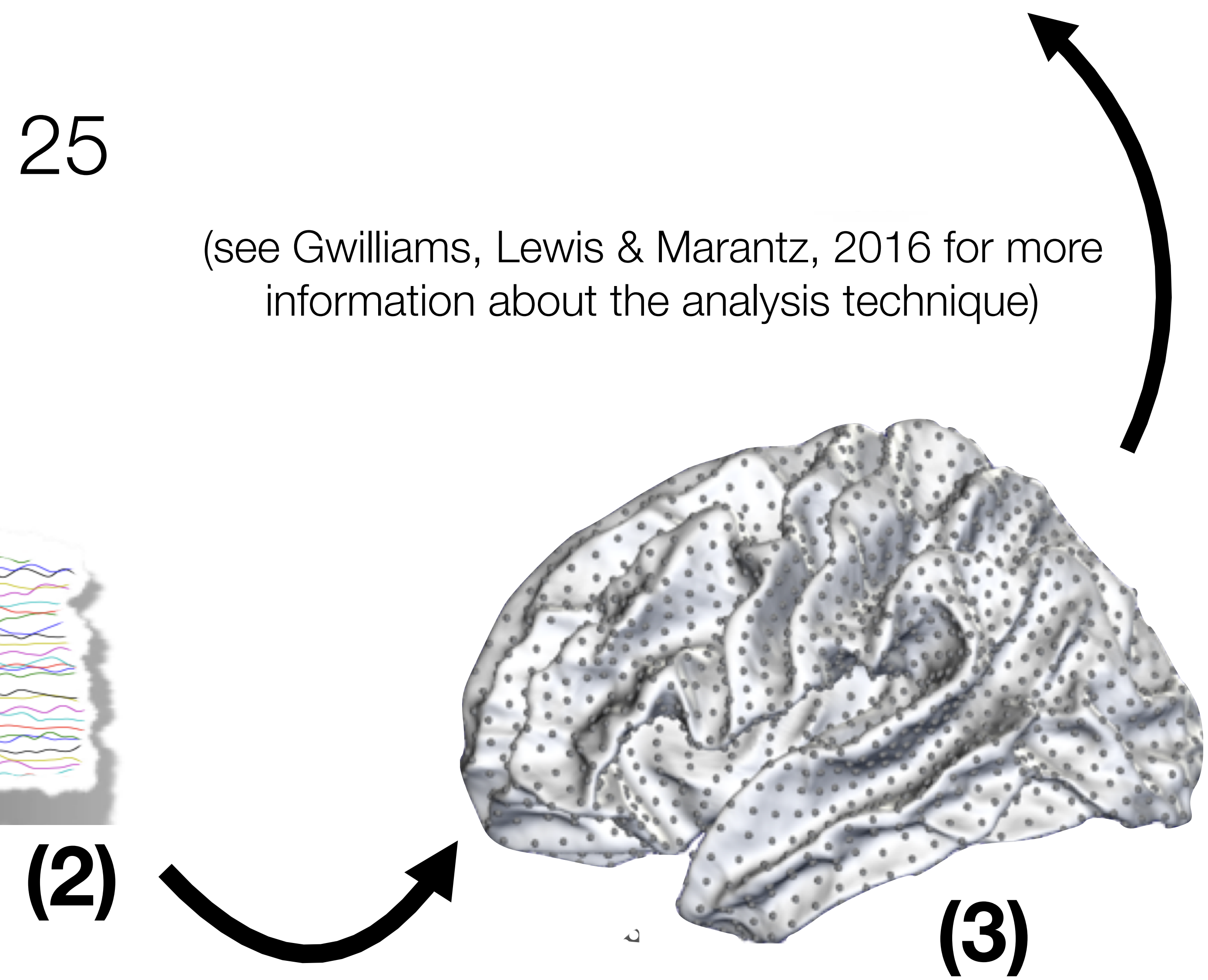
Design & Materials



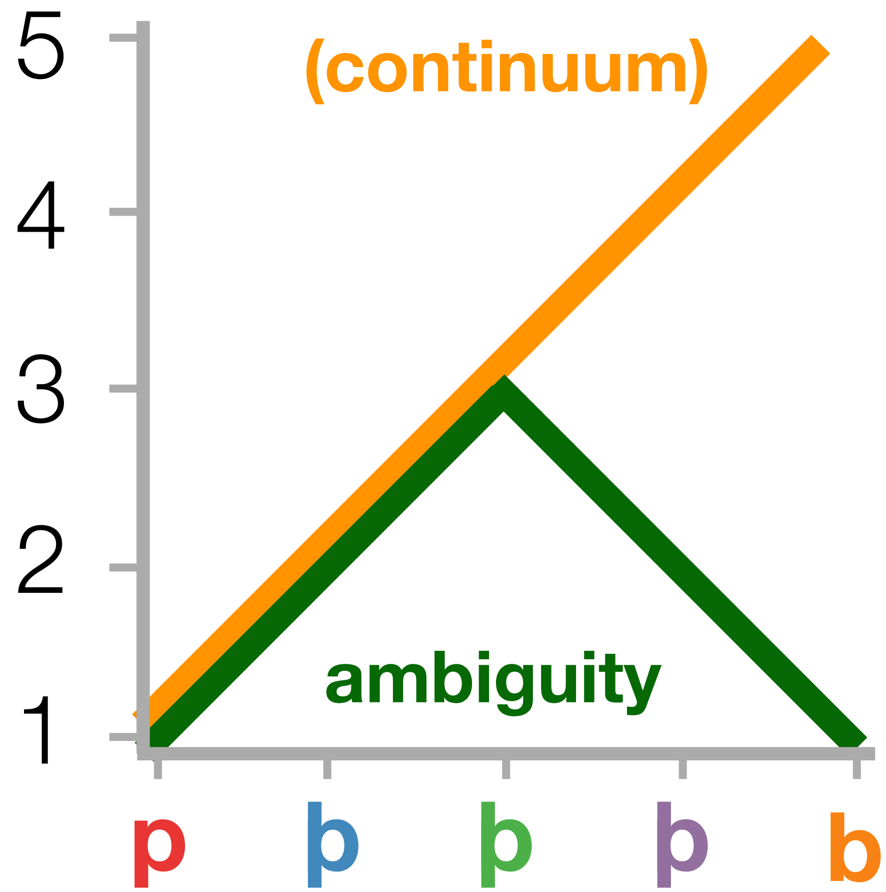
Procedure & Analysis

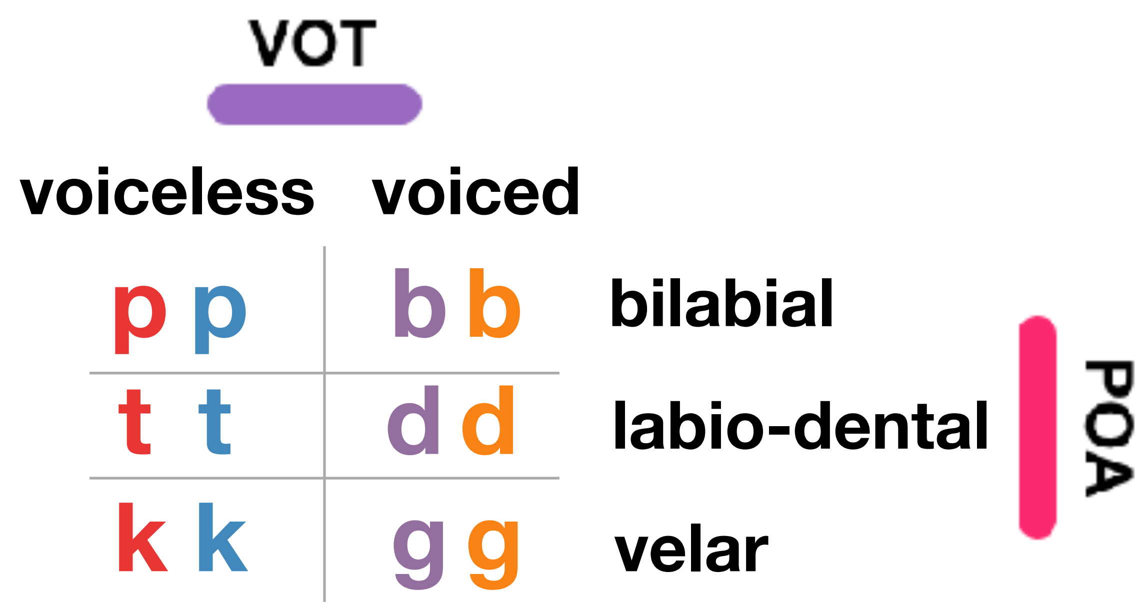
(see Gwilliams, Lewis & Marantz, 2016 for more information about the analysis technique)





Three Experimental Variables





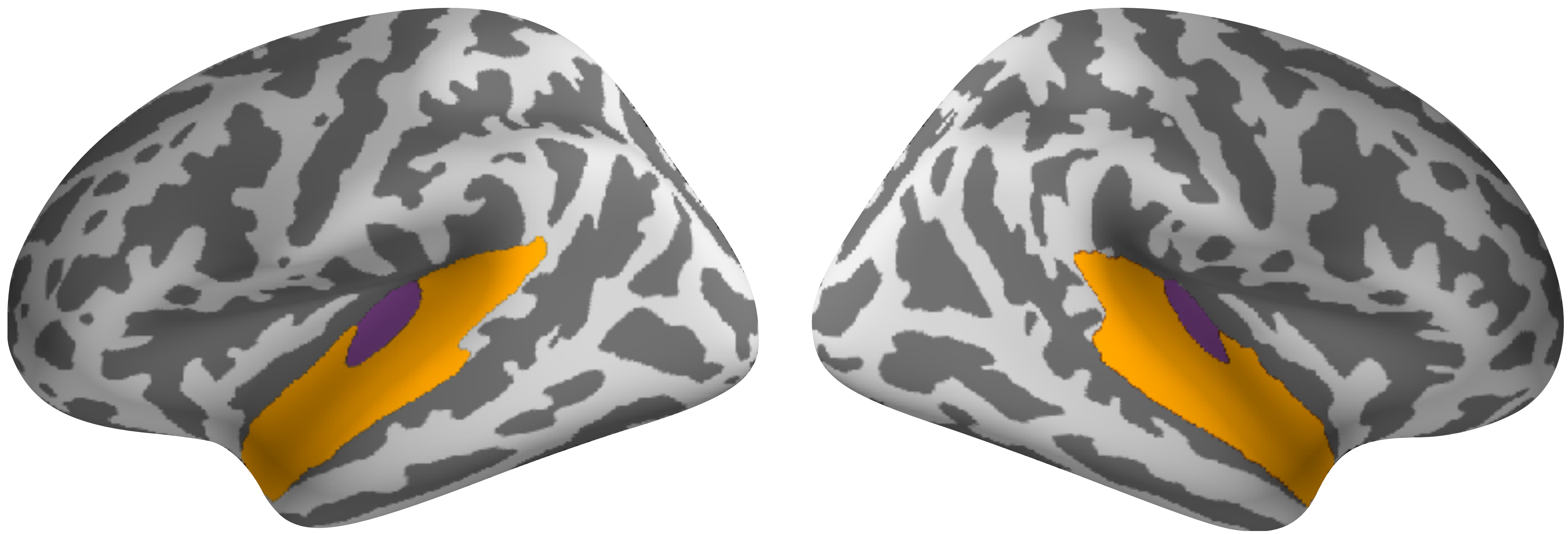
Today's Questions

How does the auditory cortex **respond** to phonological ambiguity?

Sensitivity to phonetic features ~100 ms after onset in superior temporal gyrus:

Simos et al. 1998, Ackermann et al. 1999, Obleser et al. 2003, Papanicolaou et al. 2003, Obleser et al. 2004 Mesgarani et al. 2014, Di Liberto et al. 2015 p p b b b

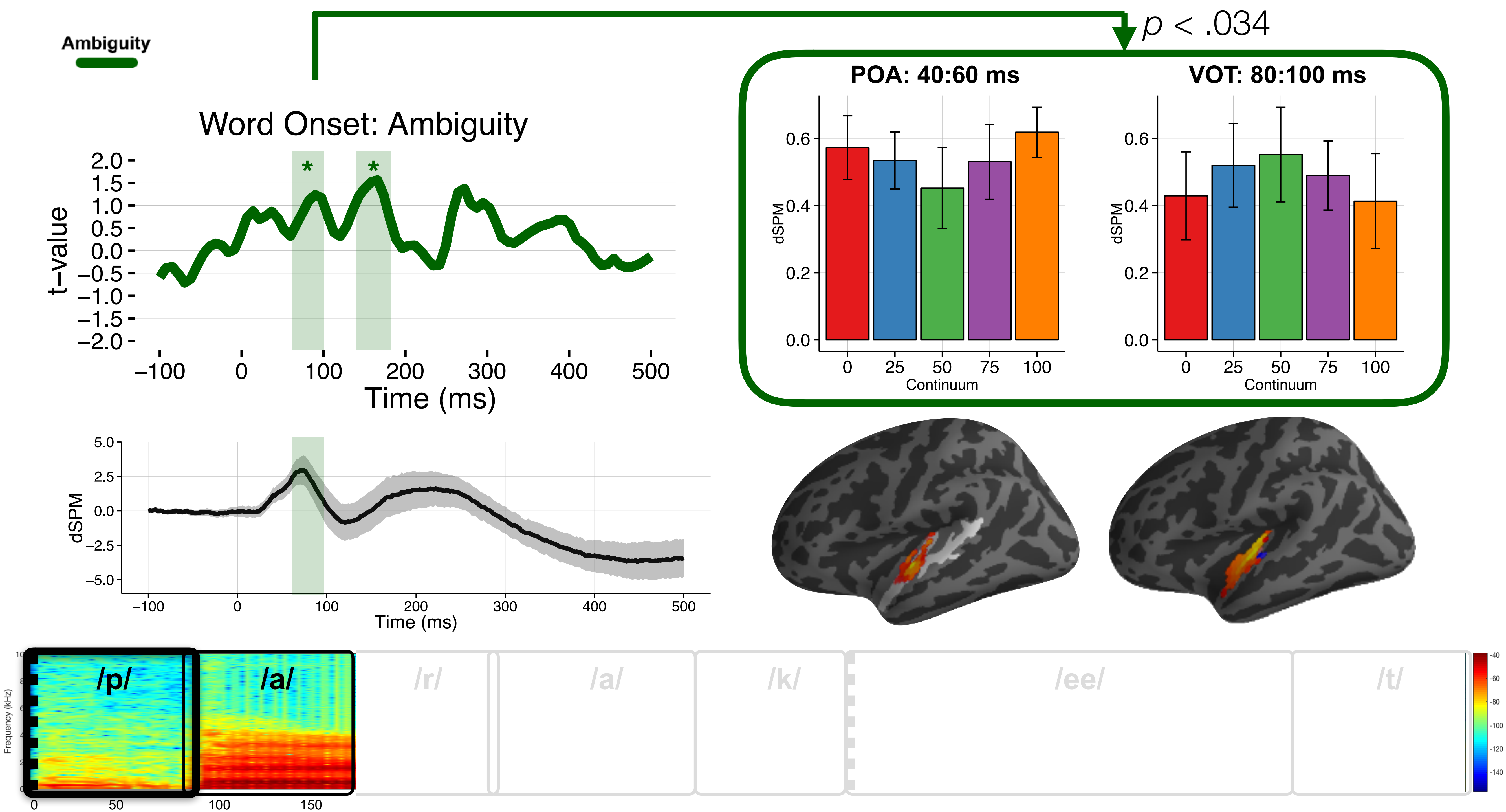
Ambiguity at Onset



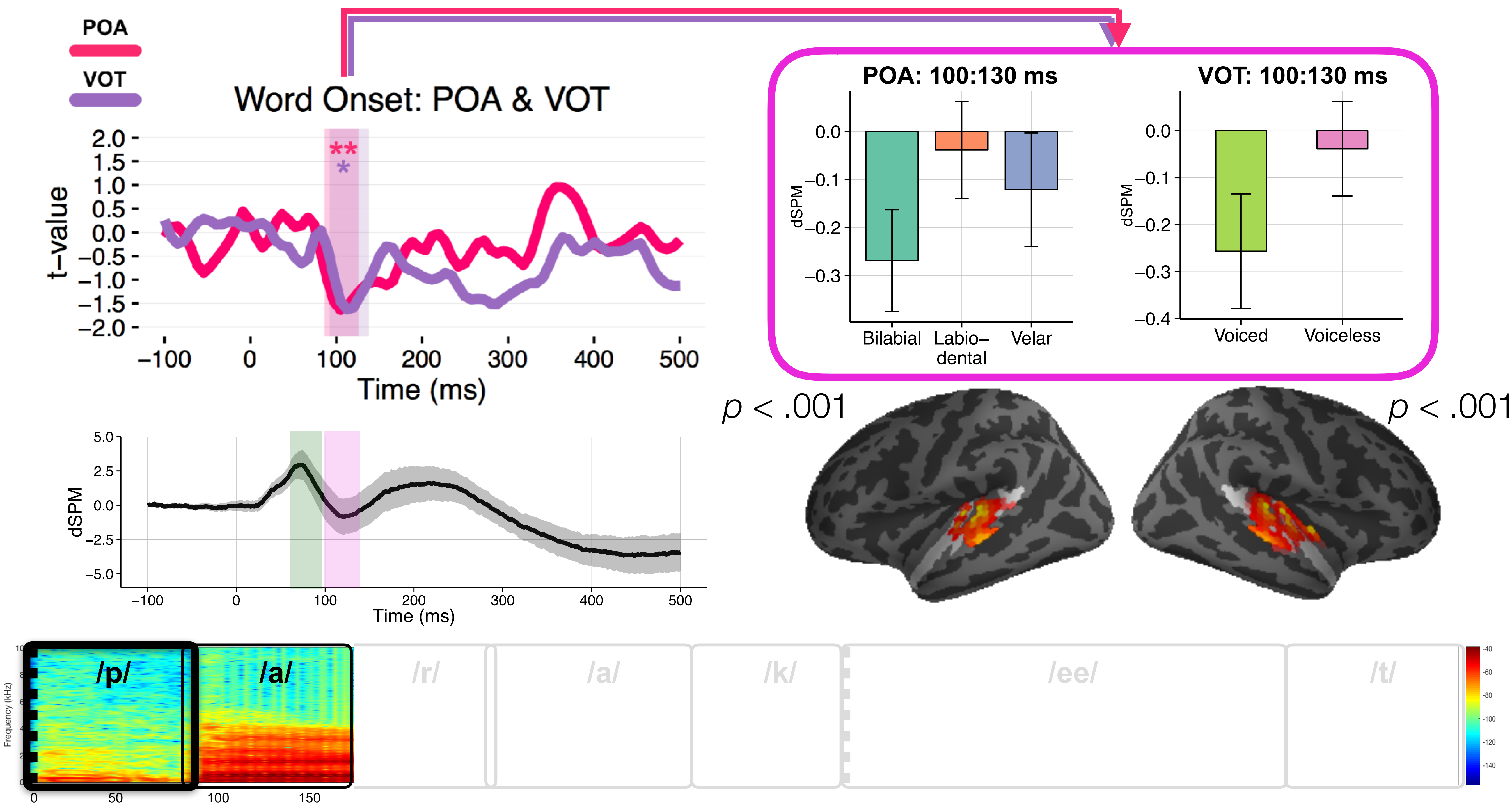
- Time-window: 0-200 ms after word onset
- Region: **Heschl's gyrus** & **superior temporal gyrus** bilaterally

p p b b b

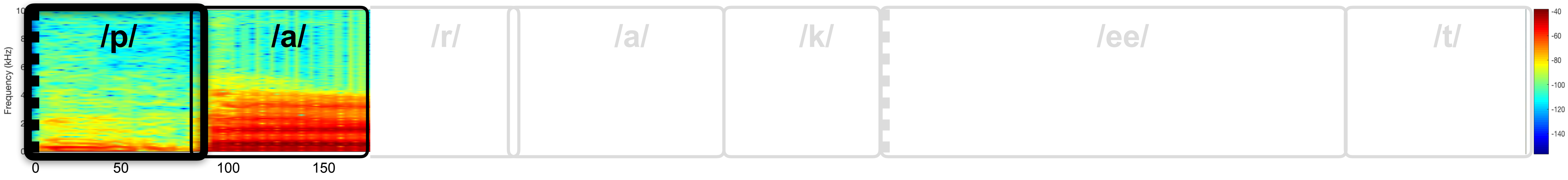
Ambiguity at Onset



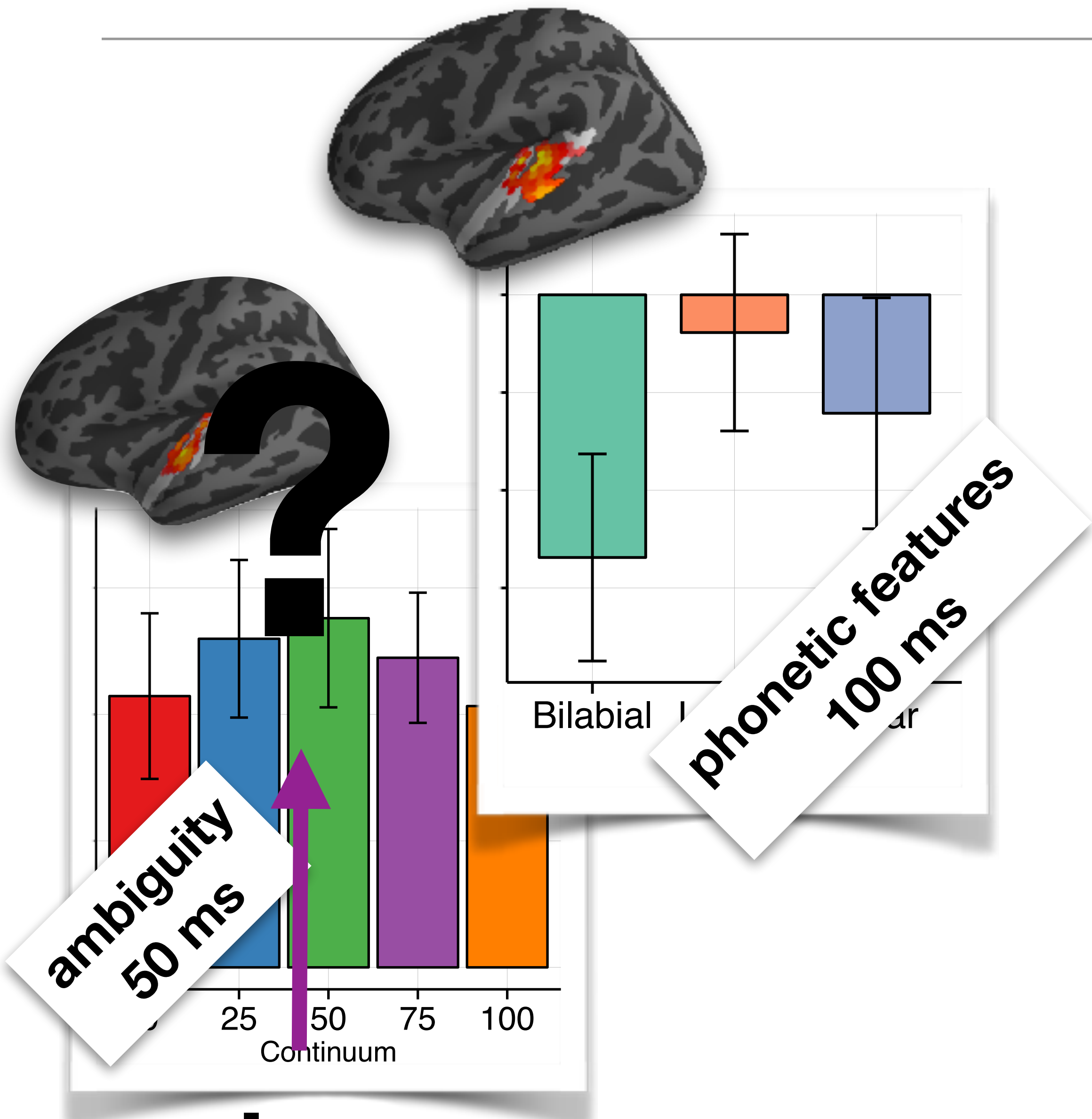
Ambiguity at Onset



Ambiguity at Onset



Interim Conclusion



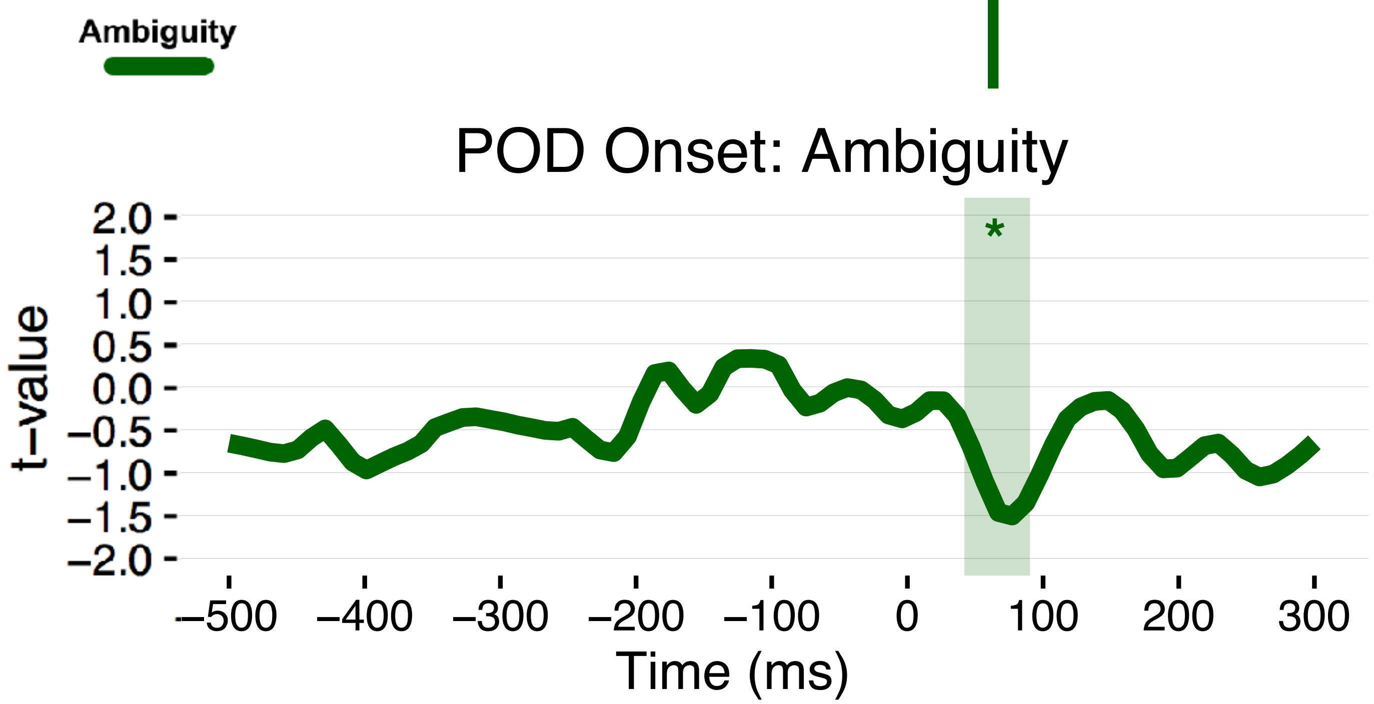
?

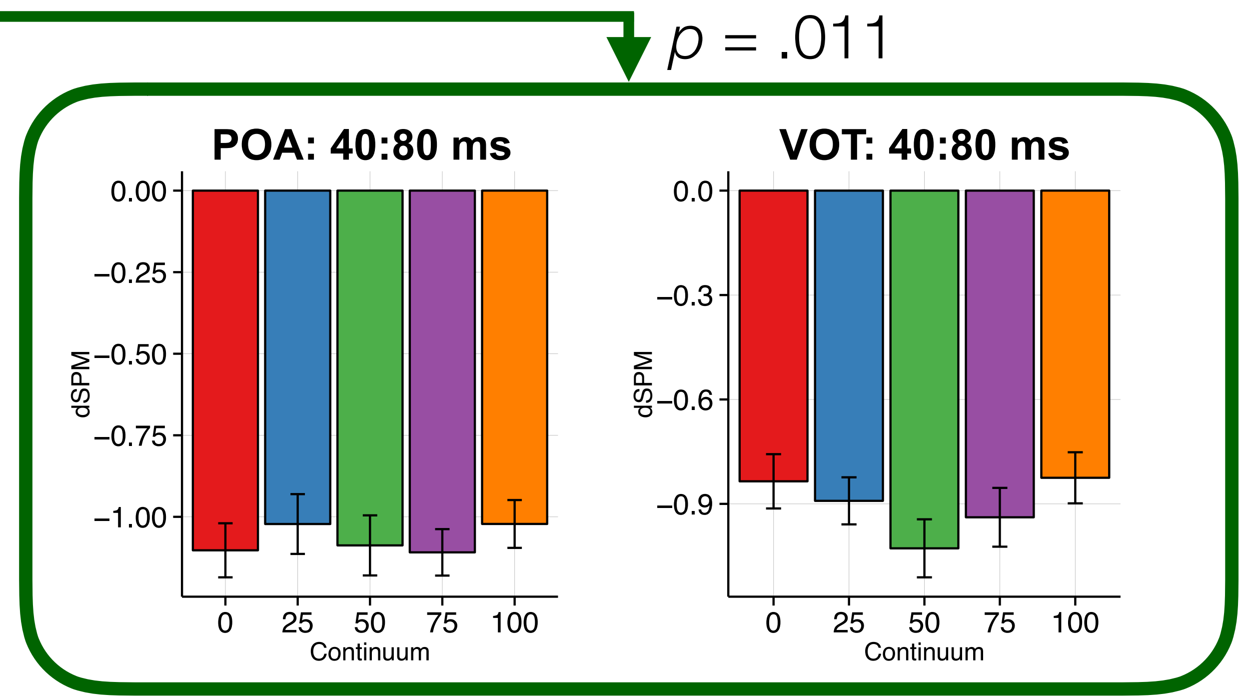
b a r a k **ee** t

Today's Questions

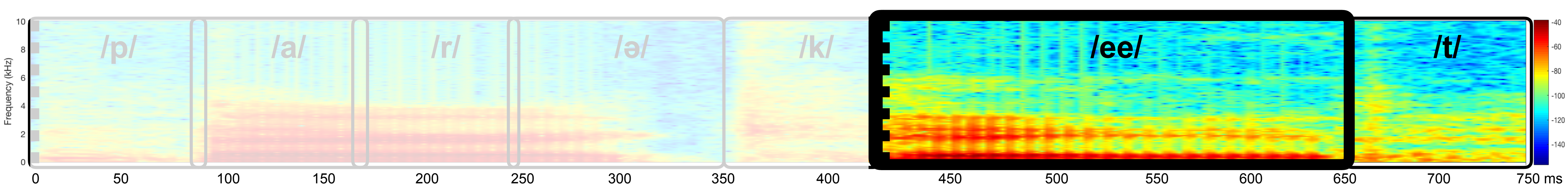
What are the neural signatures of ambiguity **resolution**?

Ambiguity at POD

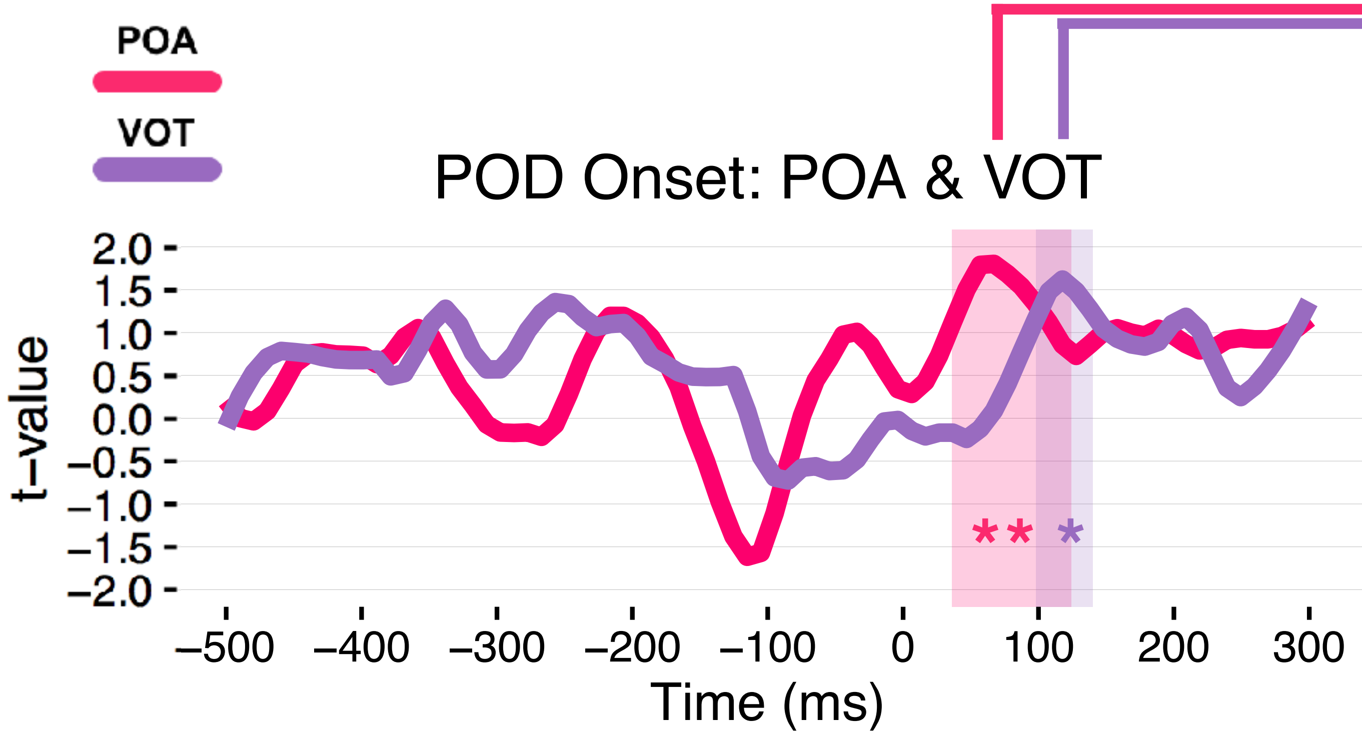


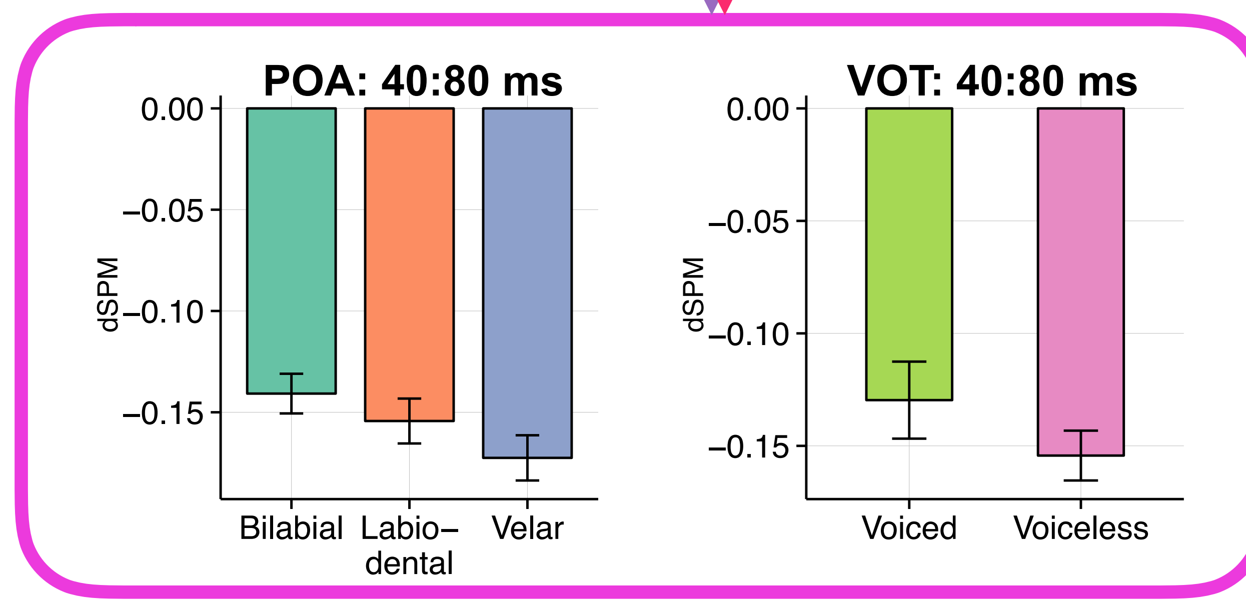


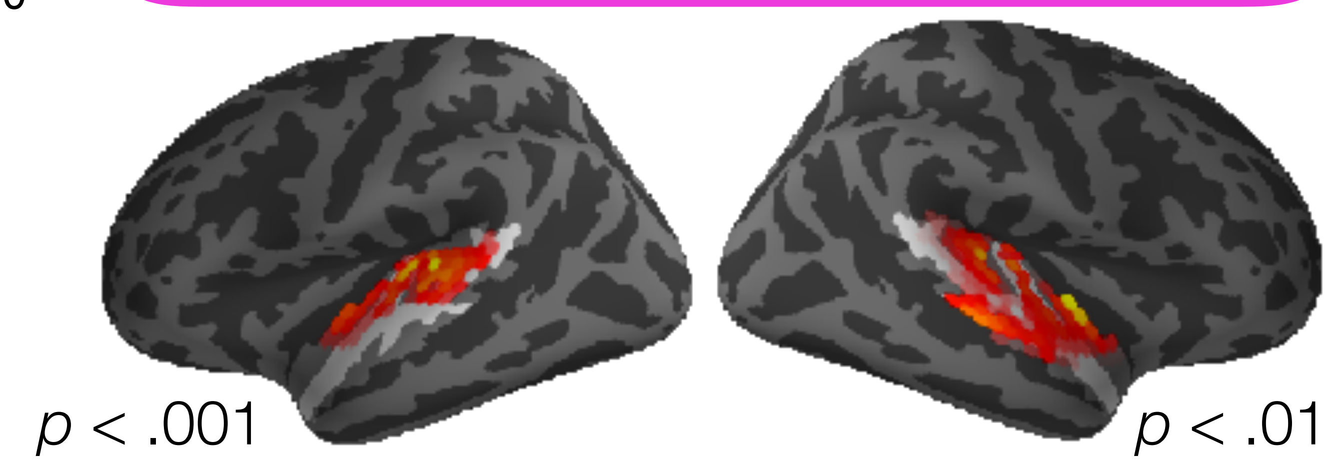
- Time-window: 0-200 ms after POD onset
- Region: **Heschl's gyrus** & **superior temporal gyrus** bilaterally

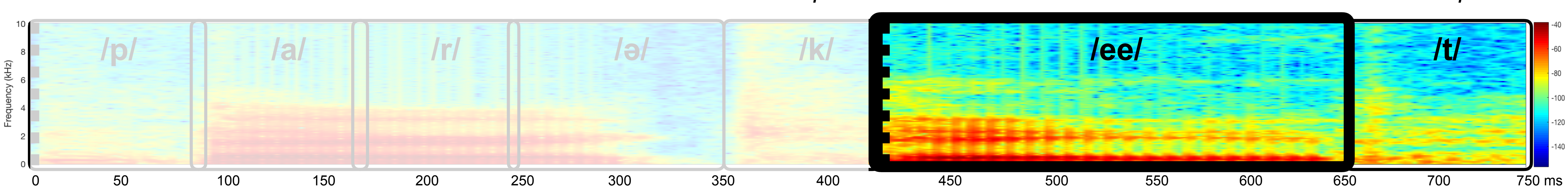


Ambiguity at POD

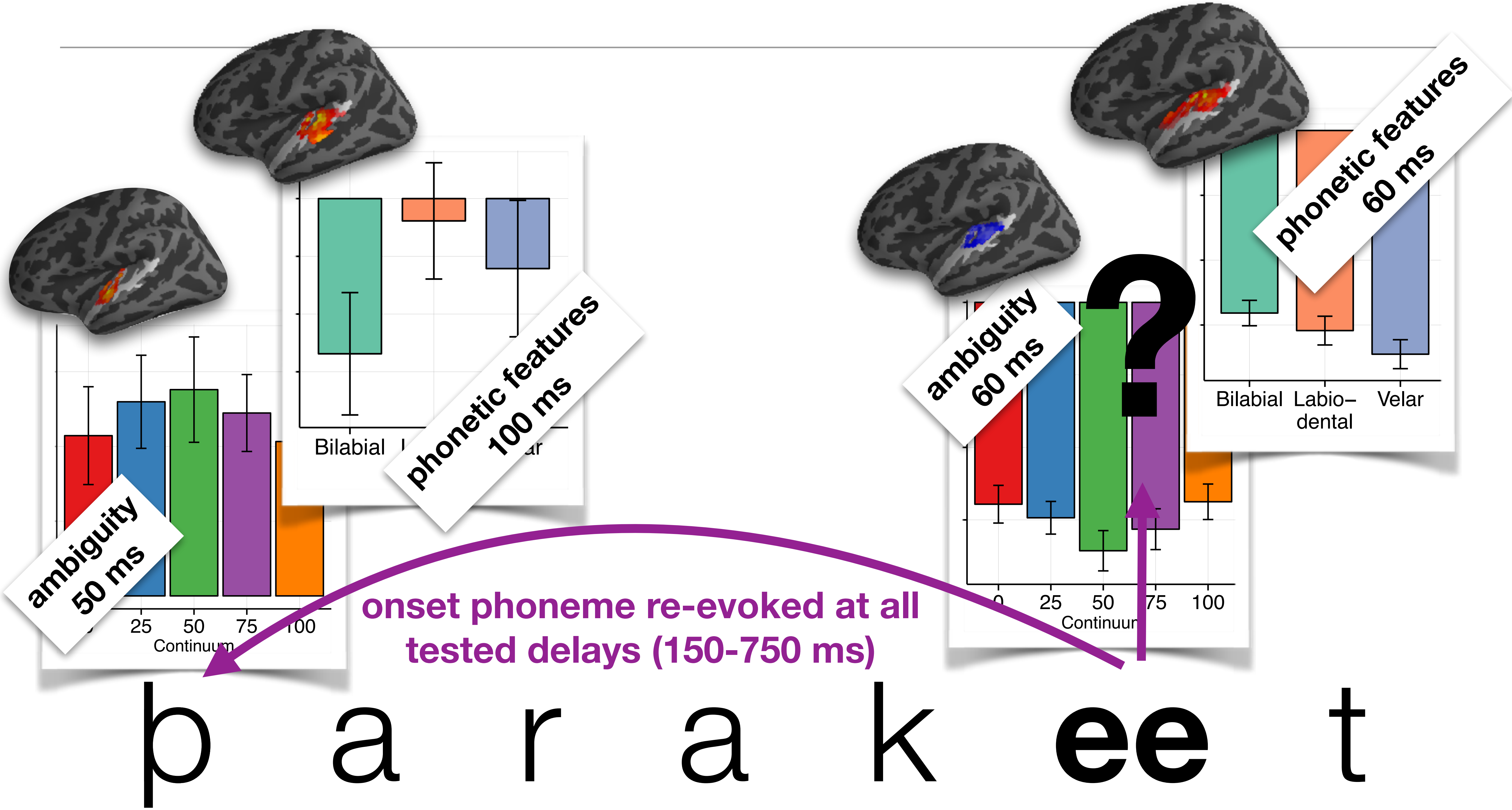




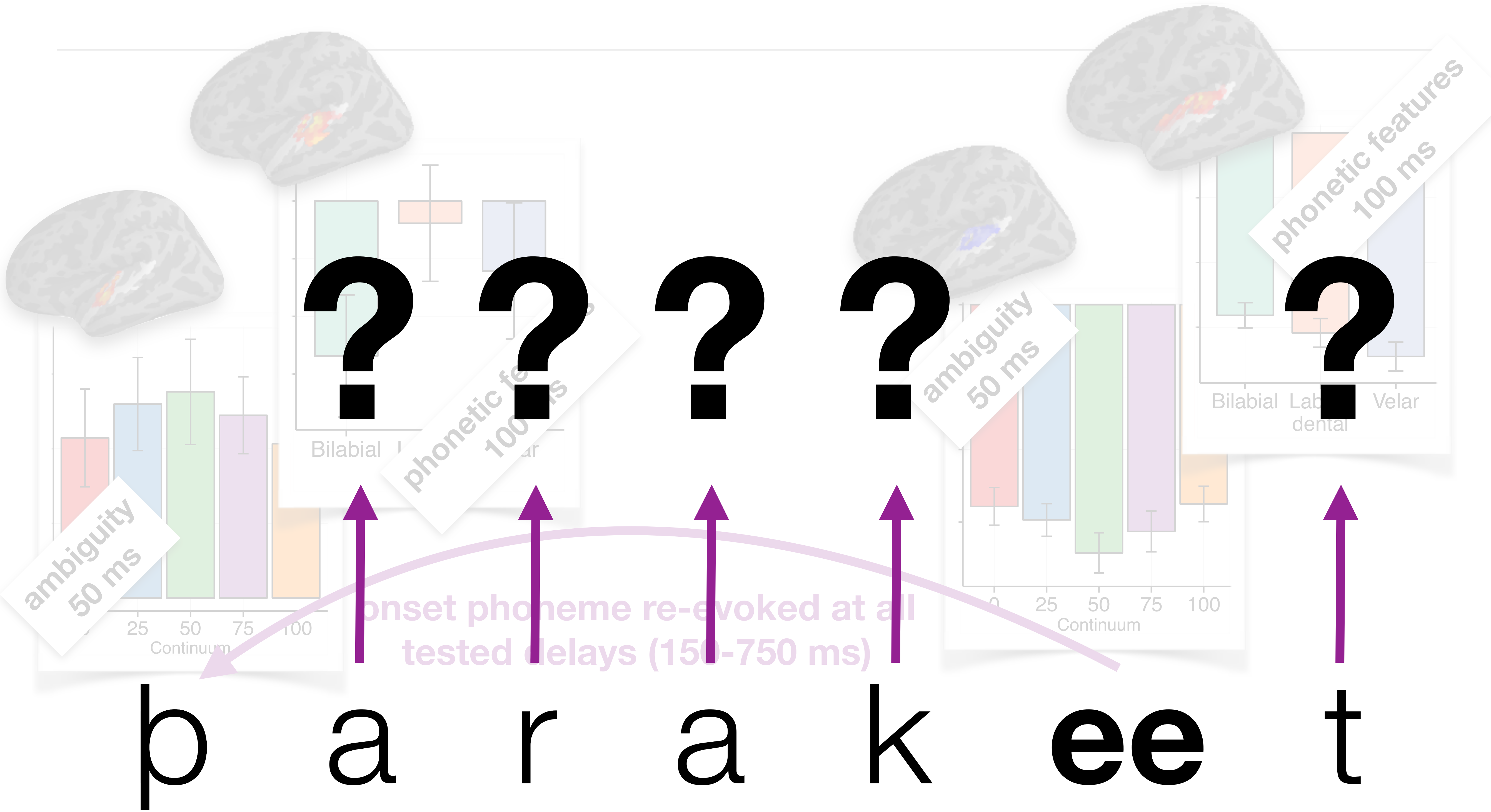




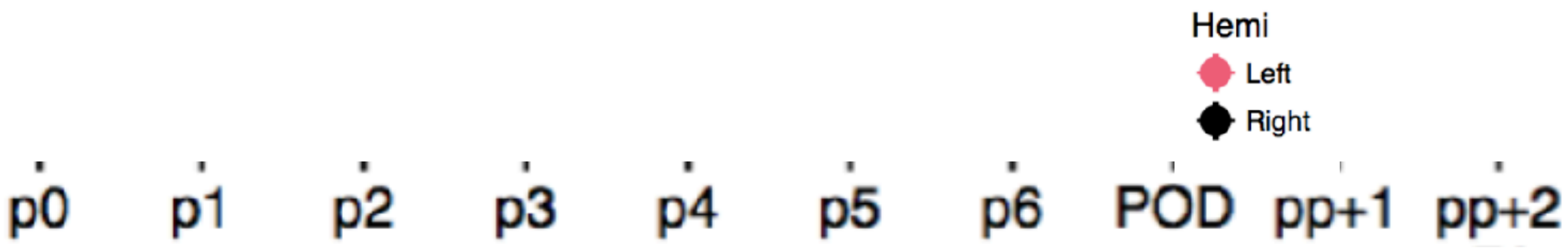
Interim Conclusion



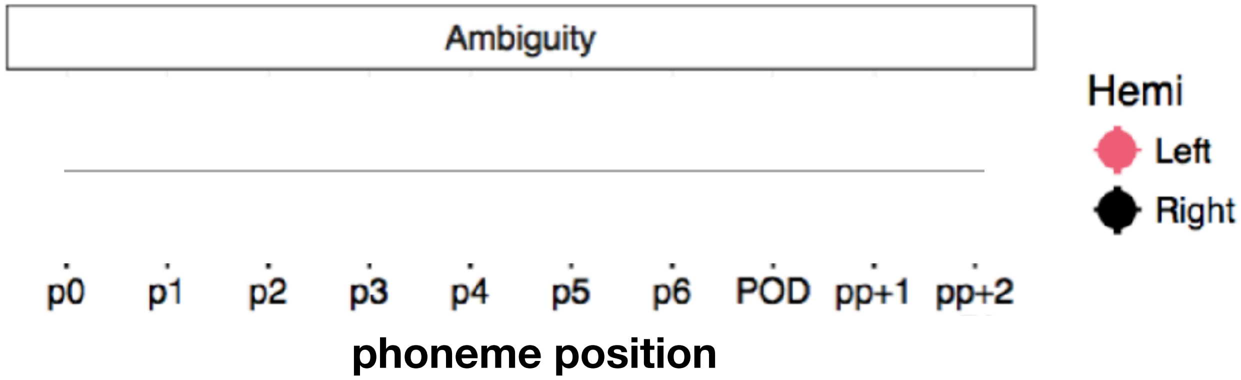
Interim Conclusion



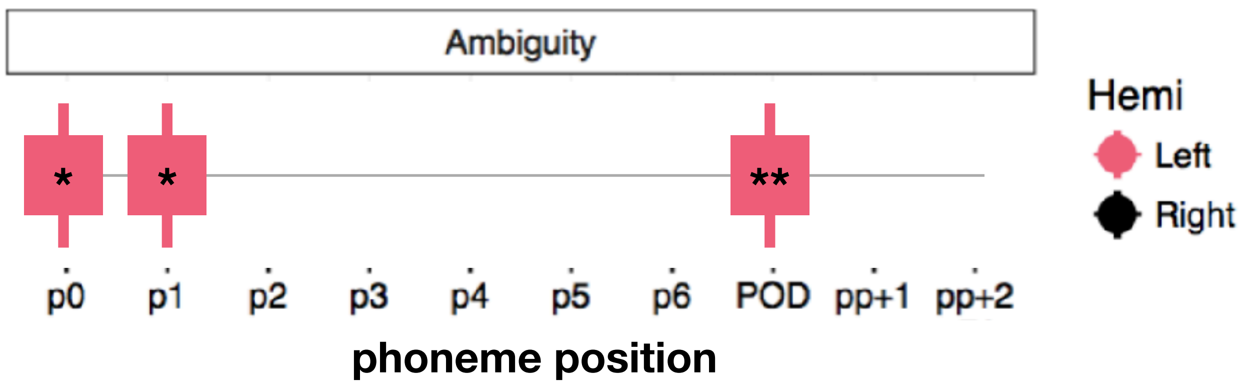
Reactivation in Intermediate Positions



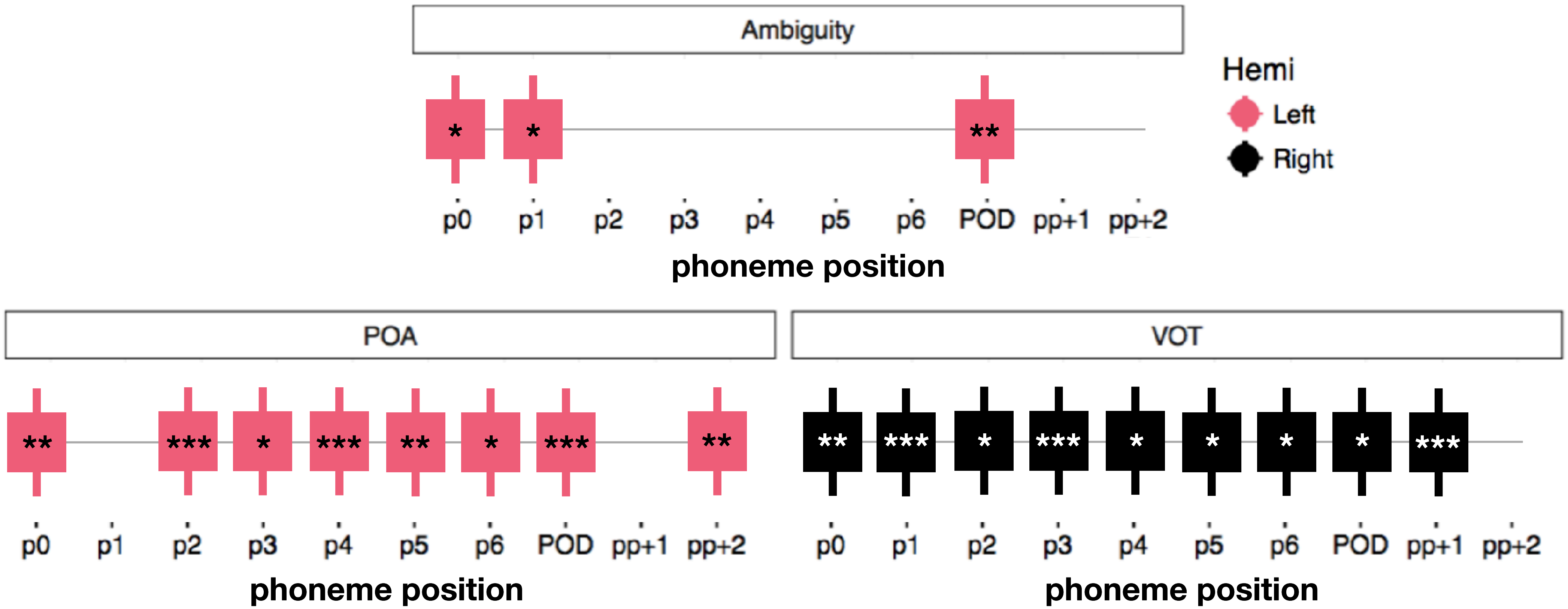
Reactivation in Intermediate Positions



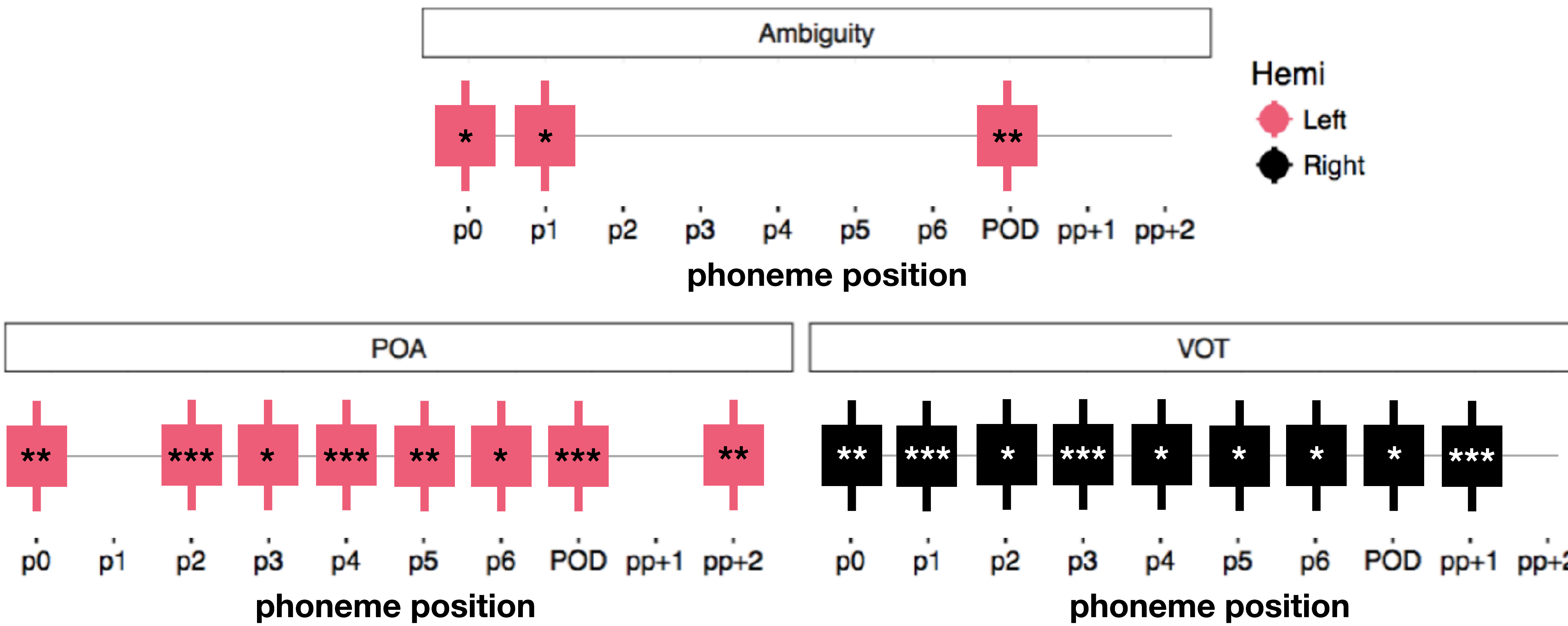
Reactivation in Intermediate Positions



Reactivation in Intermediate Positions

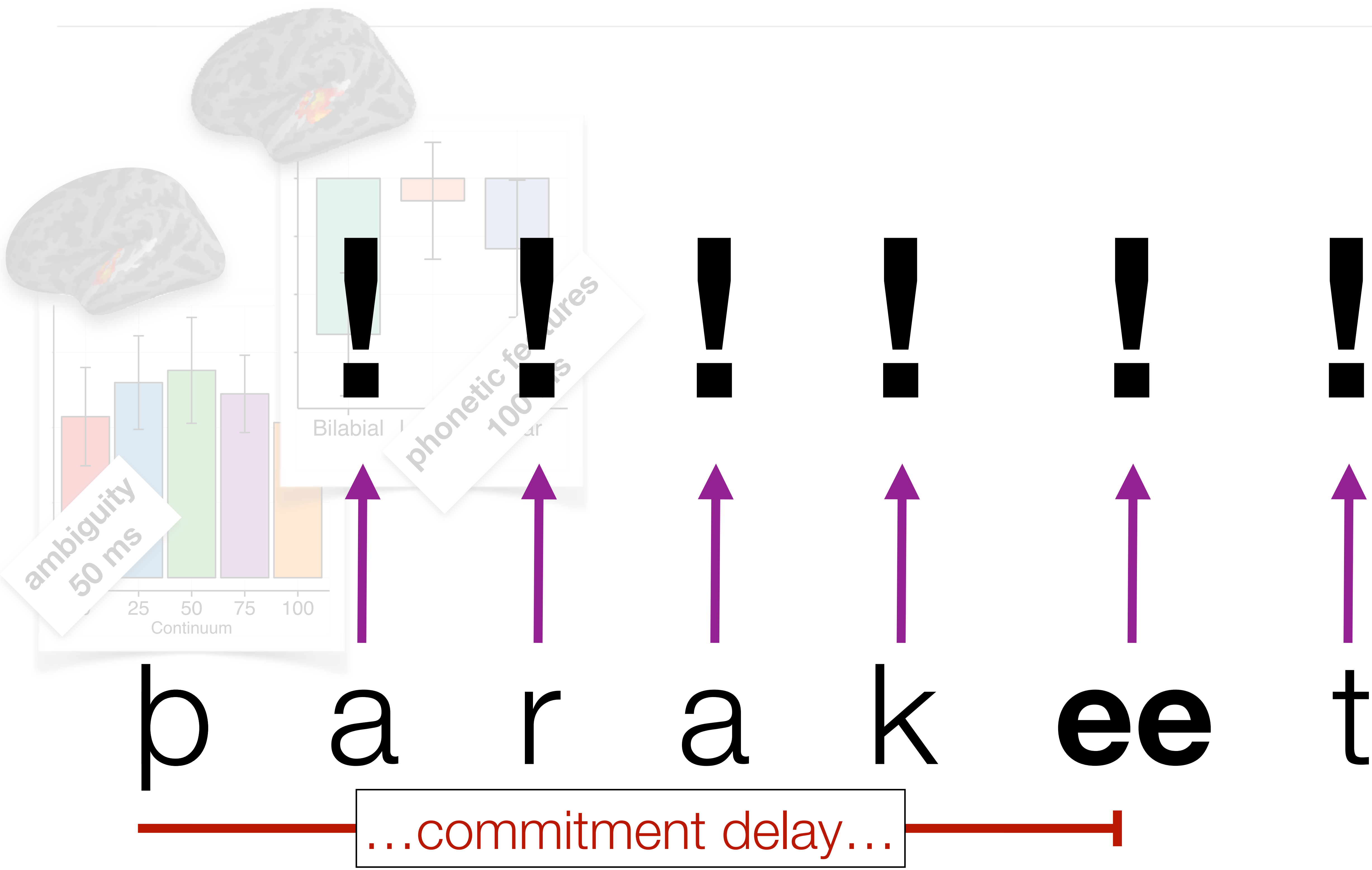


Reactivation in Intermediate Positions



- Information is re-evoked in auditory cortex
- Specifically time-locked to the onset of subsequent phonemes
  - Not driven by residual information in the acoustic signal
- Not specific to the ambiguous tokens — general to language processing

Interim Conclusion



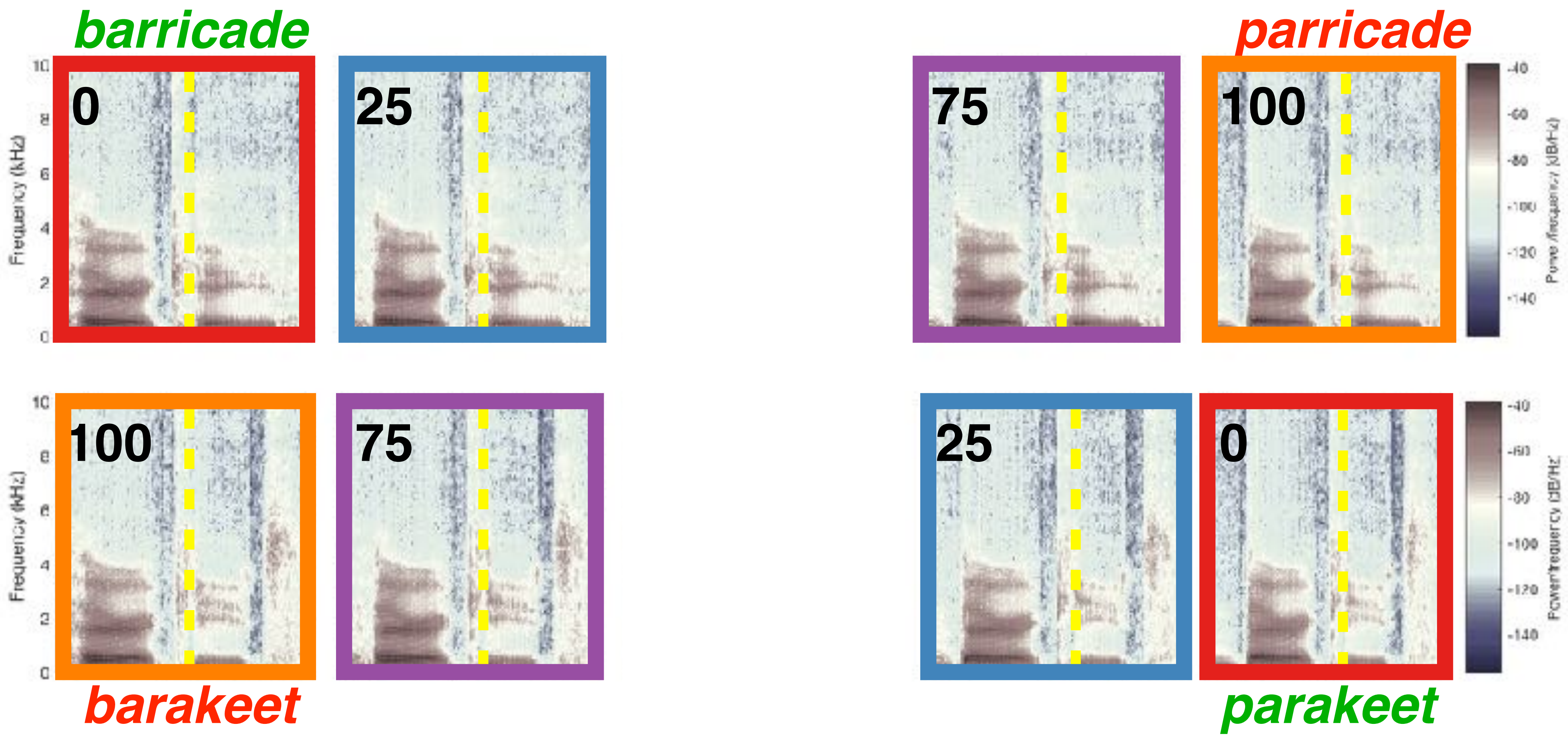
Today's Questions

How long can the system delay **phonological commitment**?

Psycholinguistic investigations into this question:

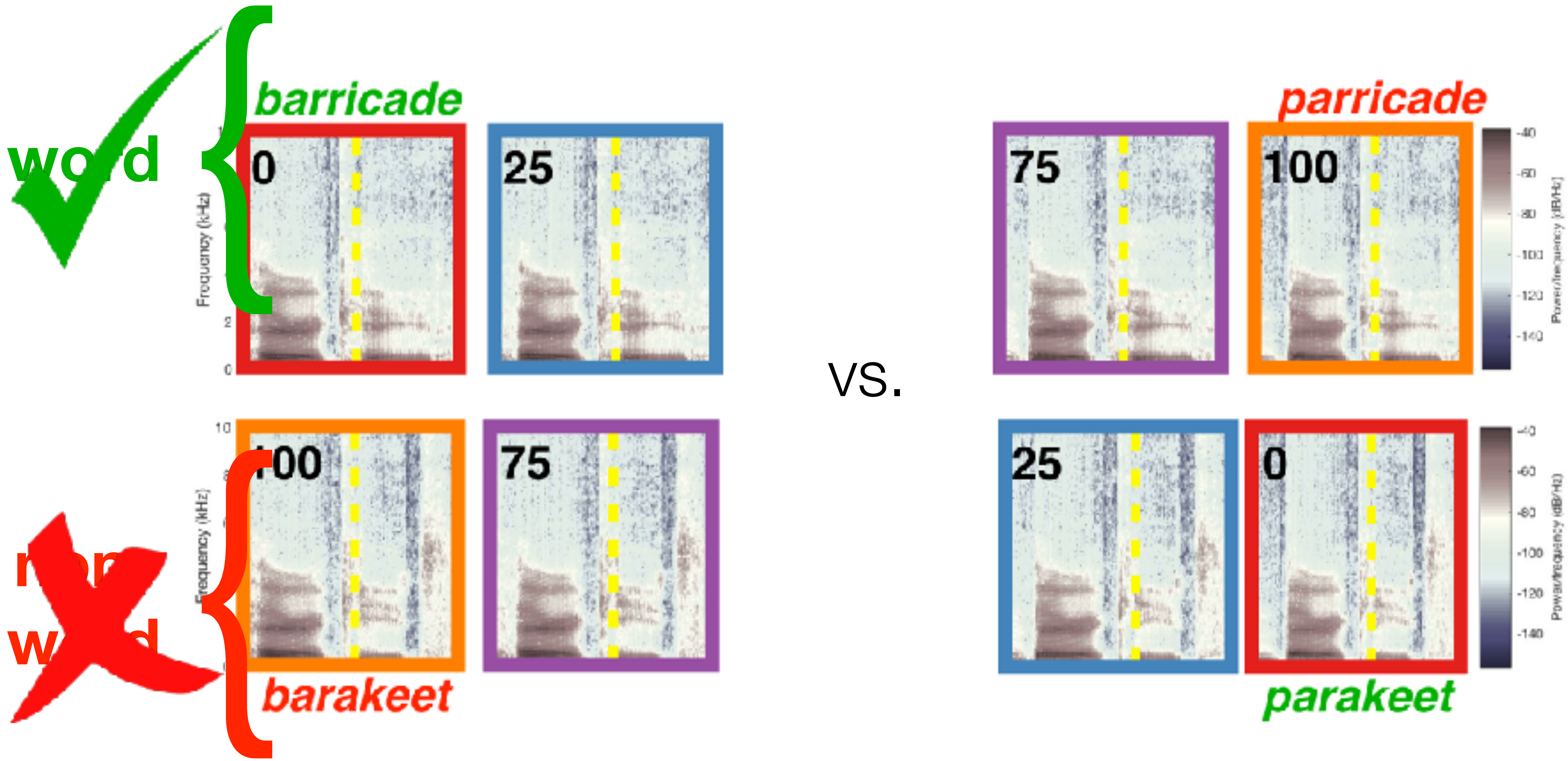
Connine et al. 1991; Samuel 1991; McMurray et al. 2009; Szostak and Pitt 2013

Example Continuum Pair



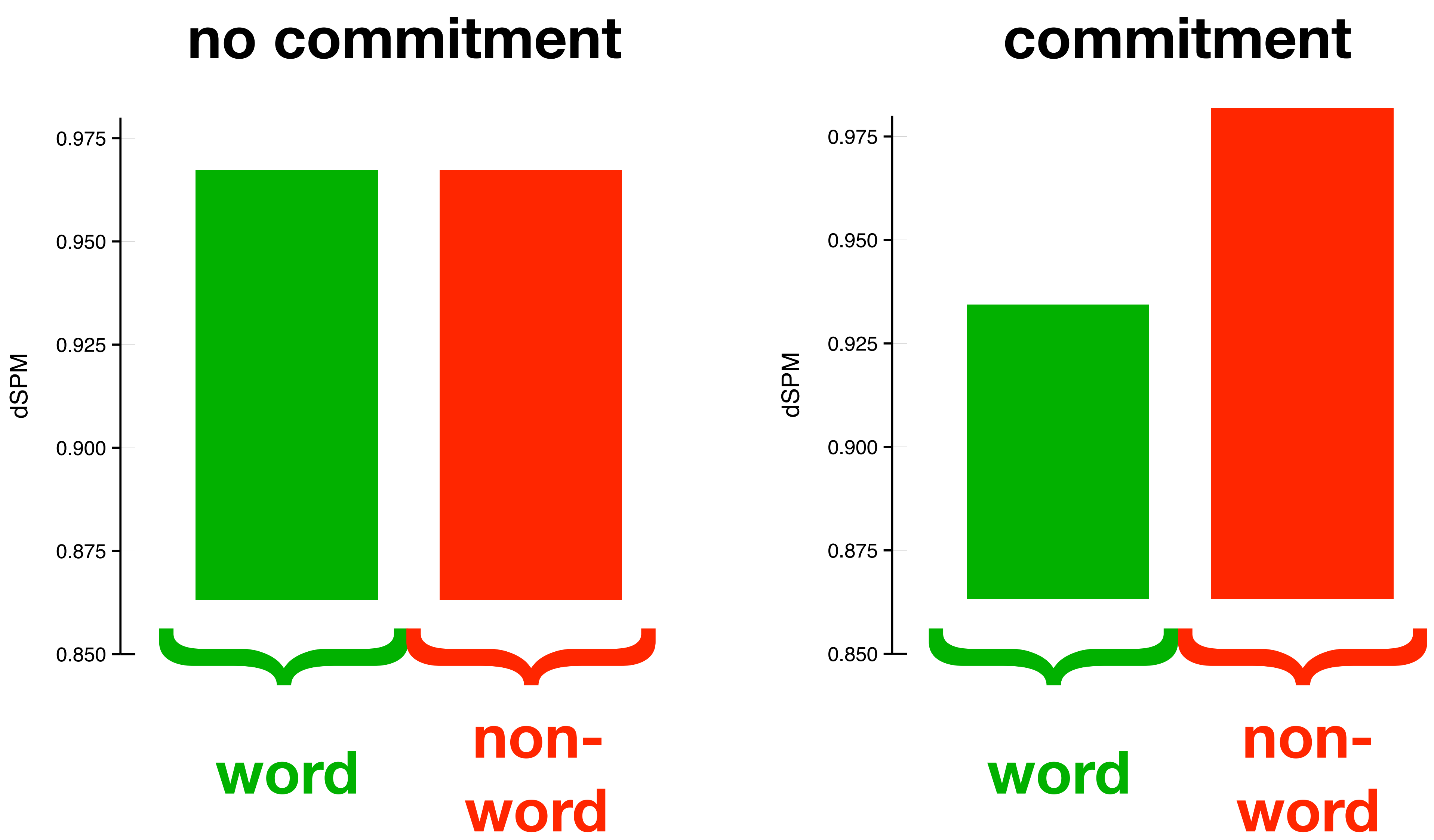
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Commitment Before POD

< 450 ms

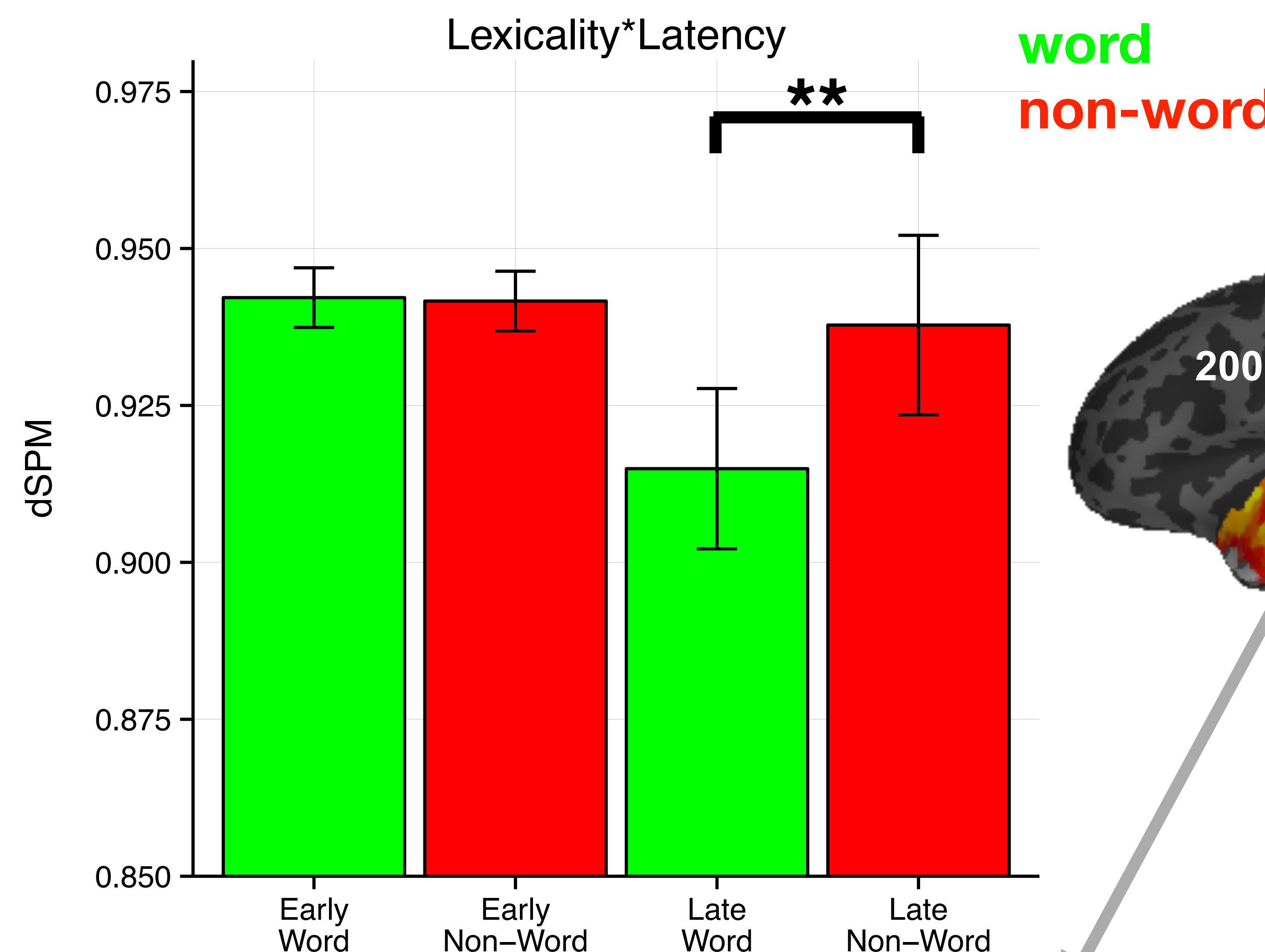
þ POD

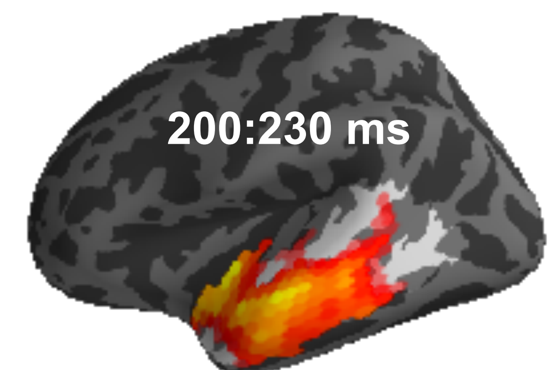
Early: POD <u>earlier</u> than 450 ms after word onset

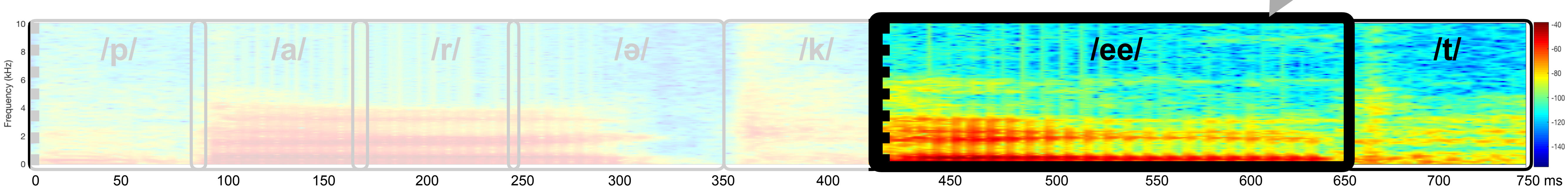
Late: POD <u>later</u> than 450 ms after word onset

þ POD

> 450 ms

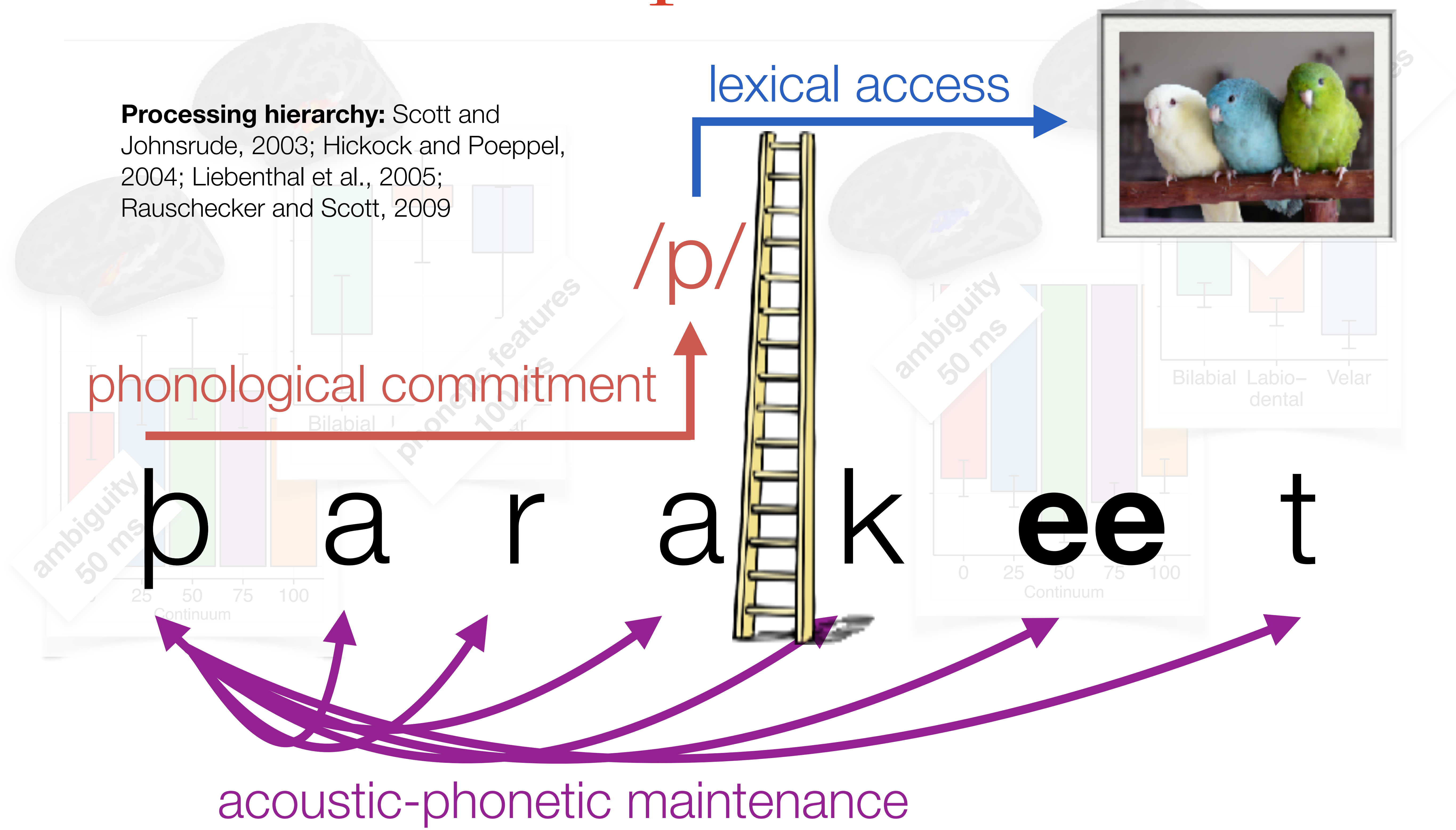






Interpretation

Processing hierarchy: Scott and Johnsrude, 2003; Hickock and Poeppel, 2004; Liebenthal et al., 2005; Rauschecker and Scott, 2009

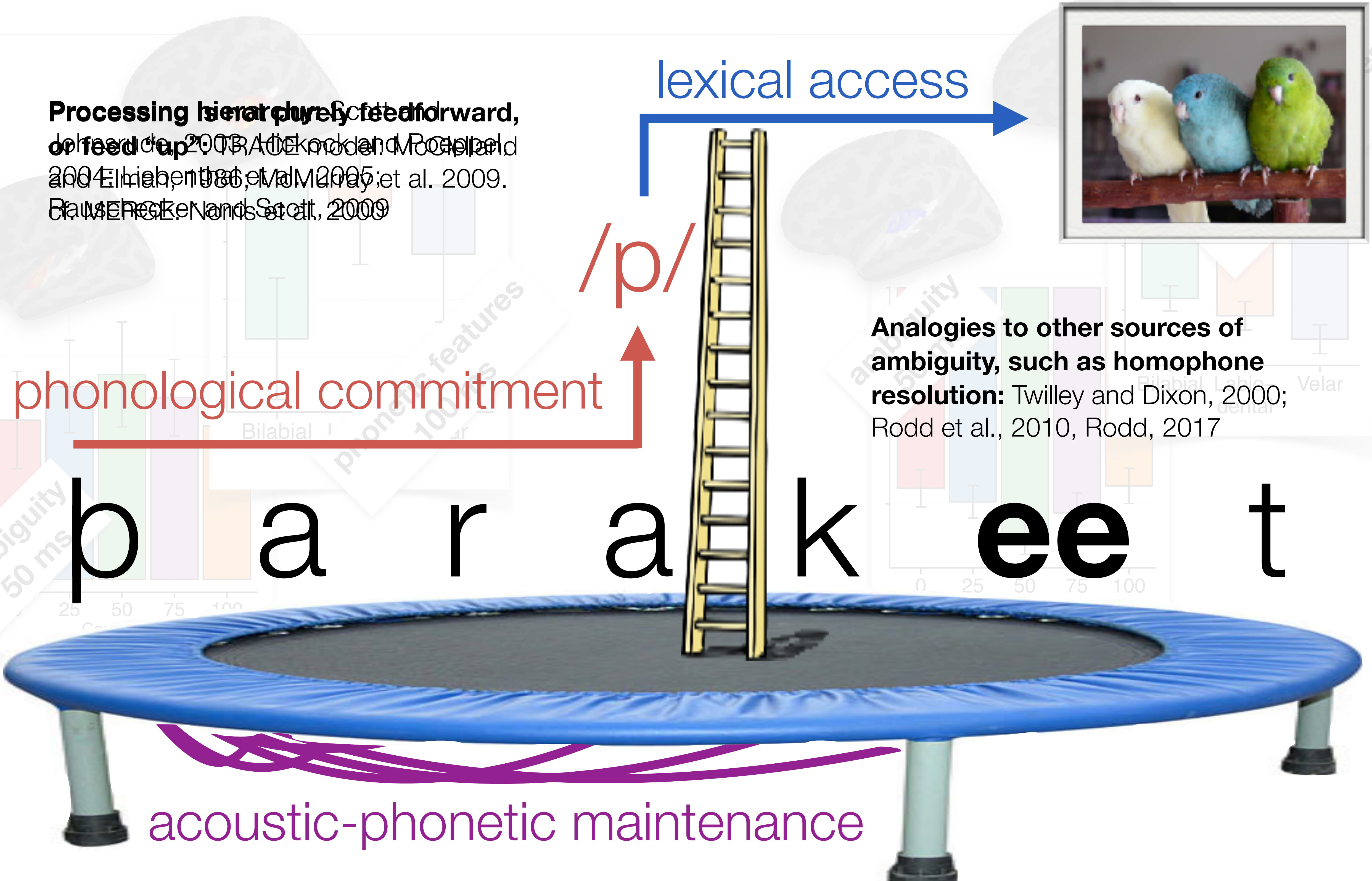


Interpretation

Processing hierarchy: Scott and Johnsrude, 2003; Hickok and Poeppel, 2004; Liebenthal et al. 2005; Rauschecker and Scott, 2009

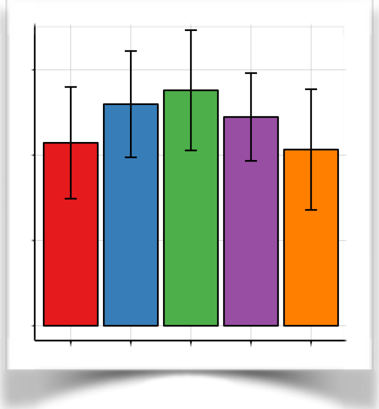
Is it purely feedforward, or feed "up"? TRACE model: McClelland and Elman, 1986; McMurray et al. 2009. cf. MERGE: Norris et al. 2000

Analogies to other sources of ambiguity, such as homophone resolution: Twilley and Dixon, 2000; Rodd et al., 2010, Rodd, 2017

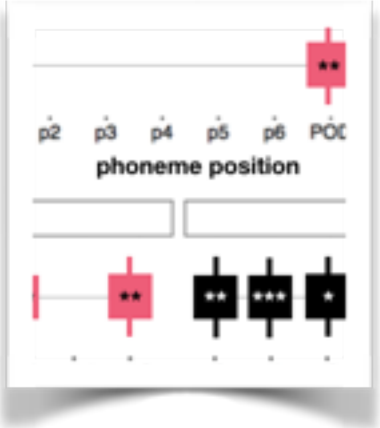


Conclusion Part 1

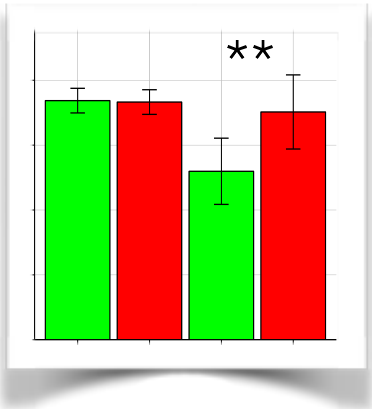
- Sensitivity to phoneme ambiguity ~50 ms after onset in primary auditory cortex



- Subphonemic detail is maintained over long time-scales (+700 ms) and re-evoked at subsequent phoneme positions



- Phonological commitment resolves ~450 ms after phoneme onset in superior temporal gyrus



Future Directions

Applying machine-learning analysis tools to uncover the dynamics of phonological processing

Research Question

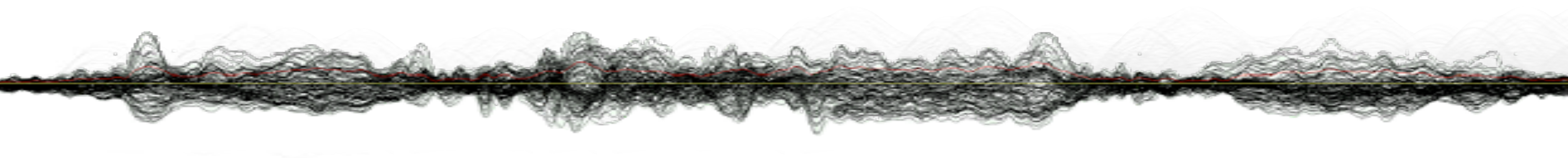
How is sub-phonemic information maintained when listening to continuous speech?

Collaborator

ME **JEAN-RÉMI KING**

Setup

we take continuous speech, and annotate it for phoneme boundaries and phonetic information



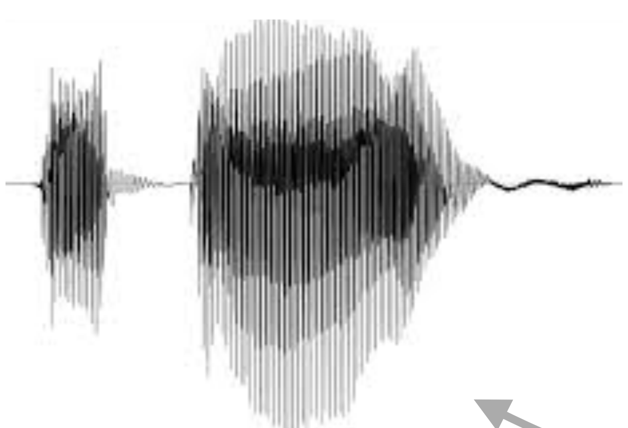
- 24 participants
- 1 hour recording
- ~40,000 phonemes per participant

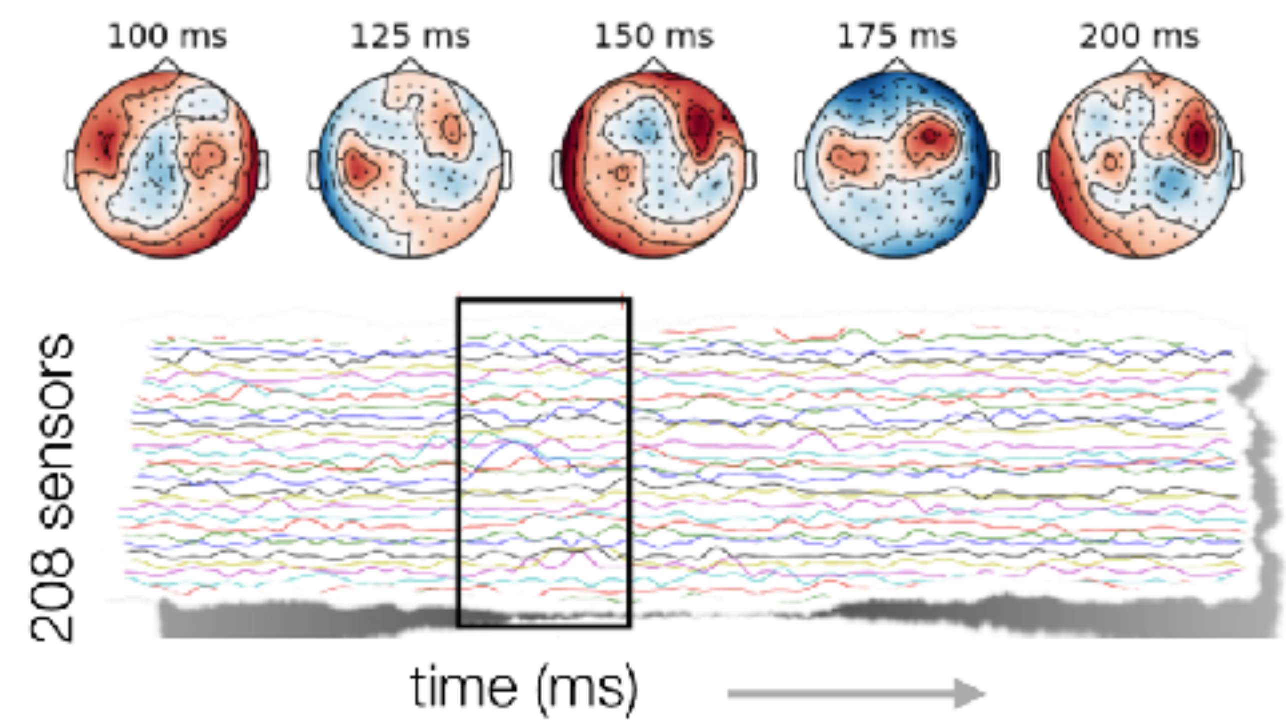
Decoding from the MEG Signal

/b/ /f/
/m/

+bilabial
+ fricative
- voiced

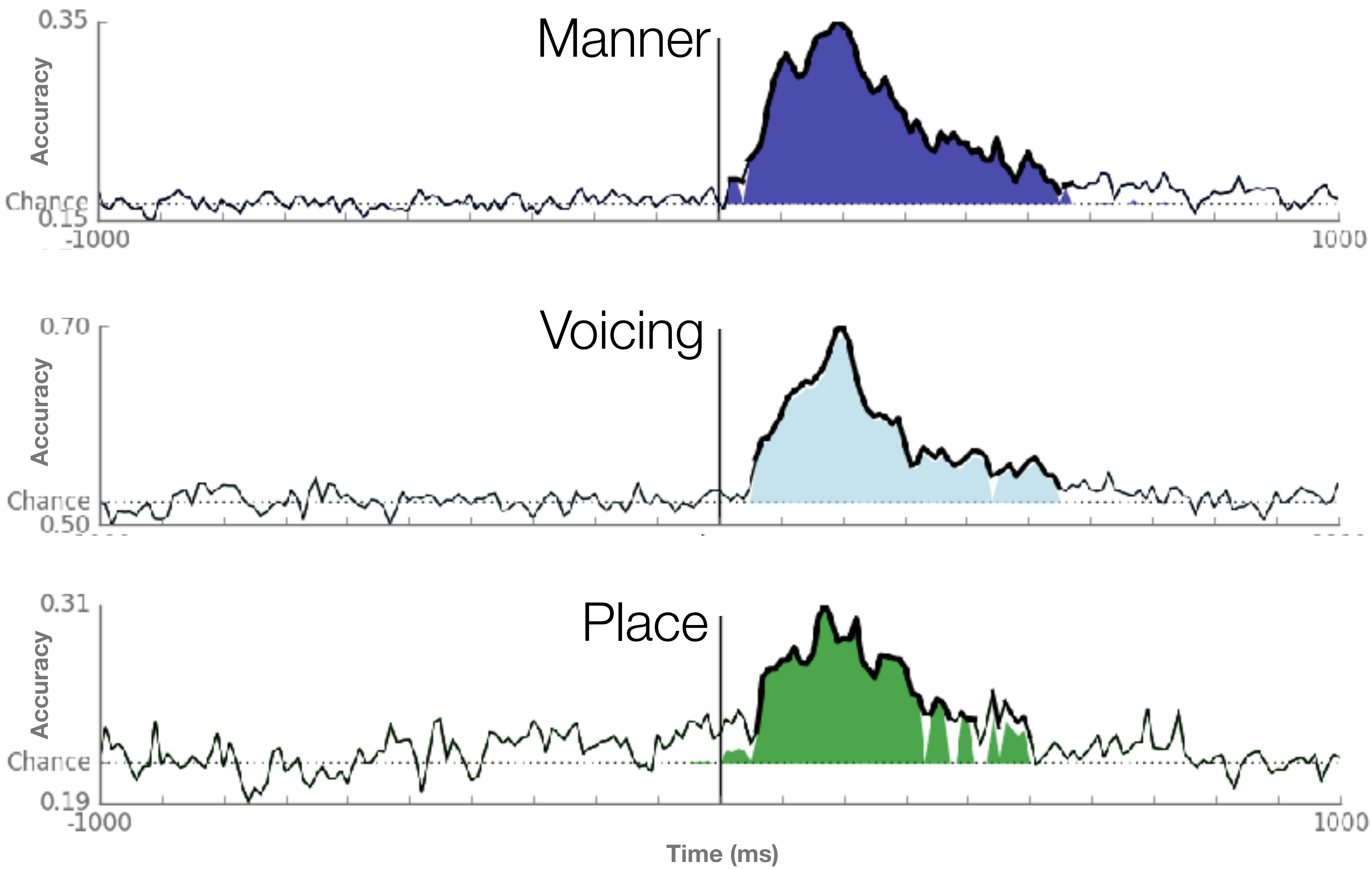
encode
stimulus



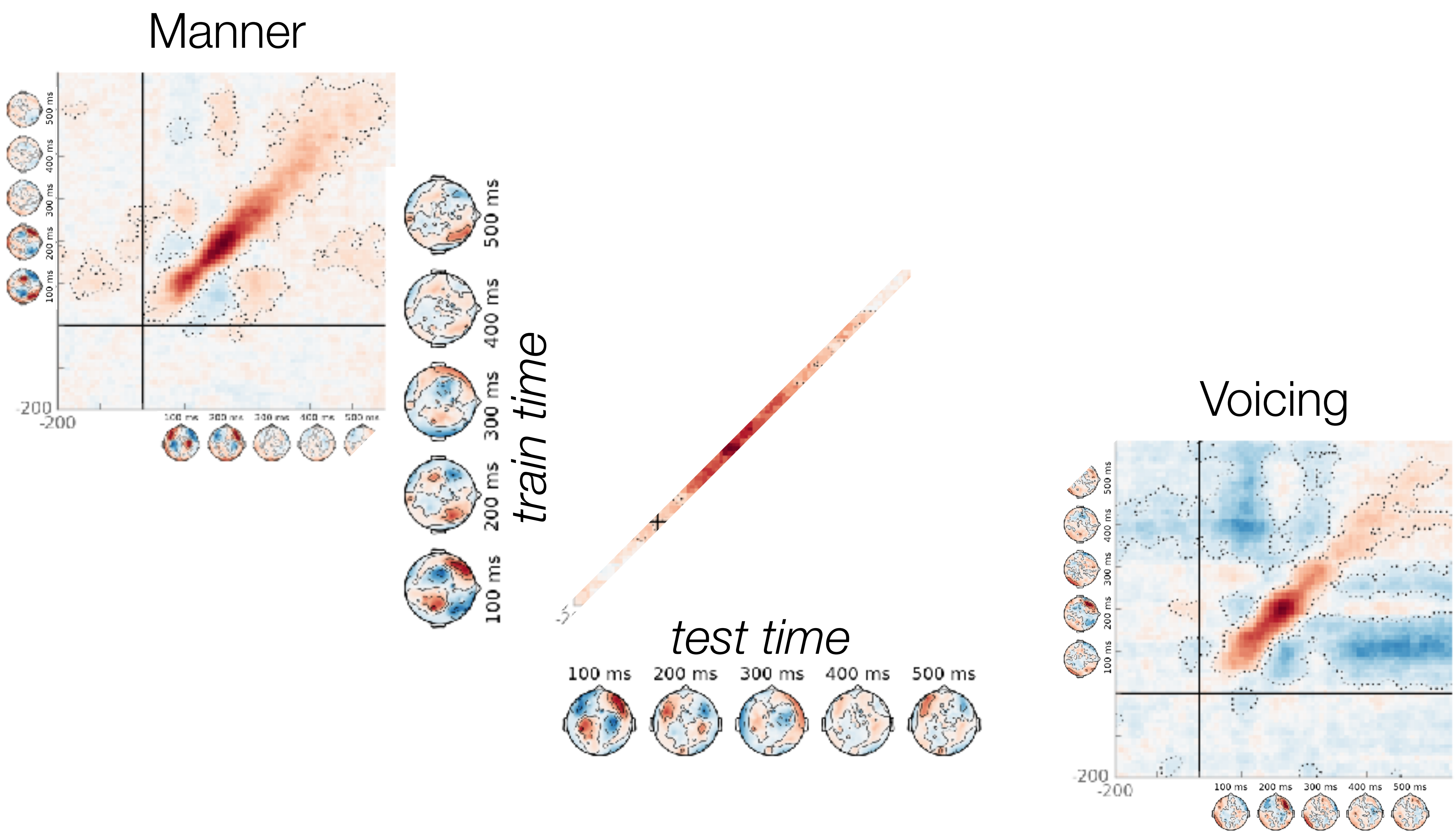


decode
stimulus

Decode Phonetic Features from the MEG Signal

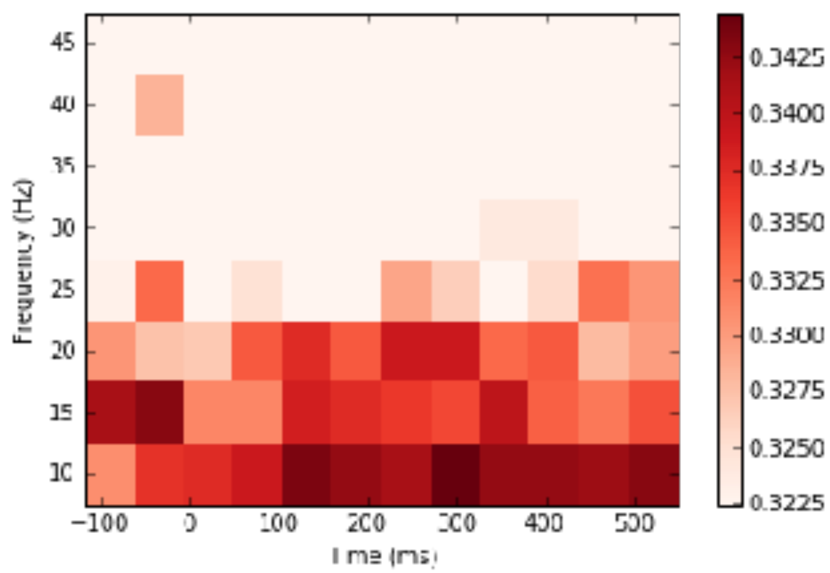


Decode the *Dynamics* of Phonetic Feature Processing

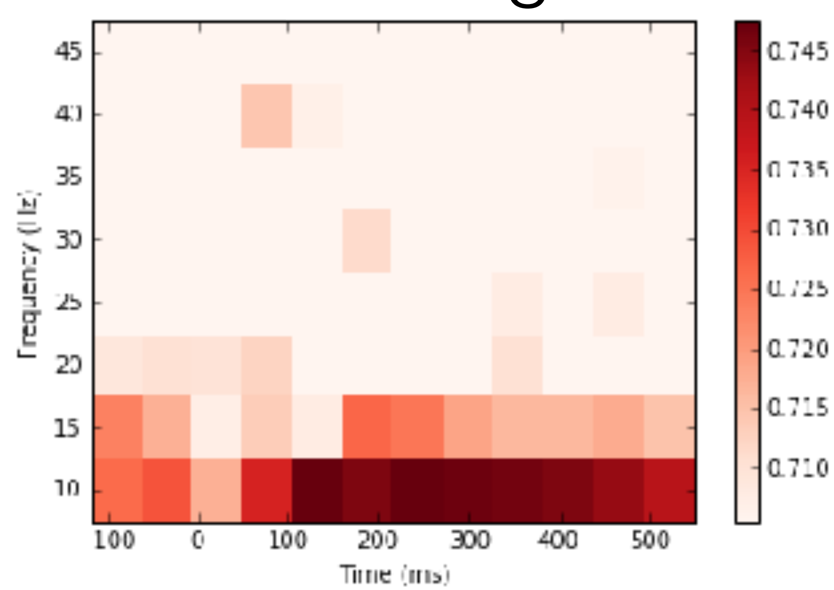


Decode the *Frequency* of Phonetic Feature Processing

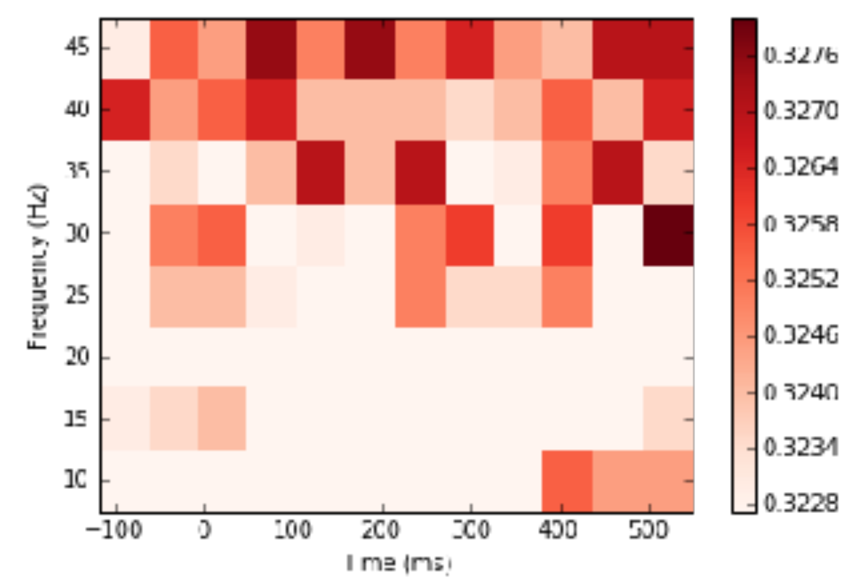
Manner



Voicing

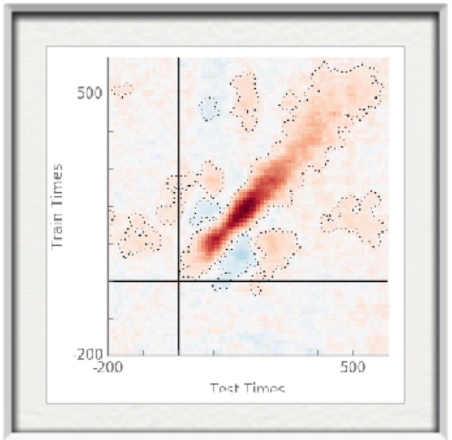


Place

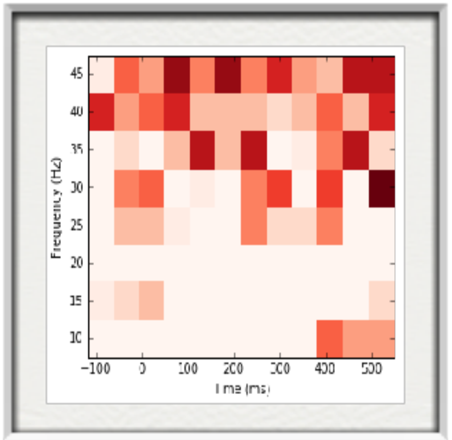


Promising Work in Progress

- Phonetic features appear to elicit different temporal dynamics



- And different spectral profiles



- There is great utility in applying machine-learning analyses to spatiotemporally resolved MEG data

Finish Line

Completed Research:

- Sensitivity to **phonological ambiguity** is reflected in the very initial stages (~50 ms) of processing a speech sound
- Sub-phonemic information is **maintained** for long periods of time, and is **re-evoked** at subsequent phoneme positions in the spoken word
- The system **commits to phonological interpretations** on a shorter time-scale in parallel to phonetic maintenance

Future Directions:

- Can we apply **machine-learning** analysis techniques to MEG data to unveil the dynamics with which sub-phonemic information is processed?

laura.gwilliams@nyu.edu
@GwilliamsL

With big thanks to:

- My supervisors, **Alec Marantz** and **David Poeppel**, as well as everyone in the **Neuroscience of Language Lab** and **Poeppel Lab**!


Funding: G1001 Abu Dhabi Institute

laura.gwilliams@nyu.edu
@GwilliamsL

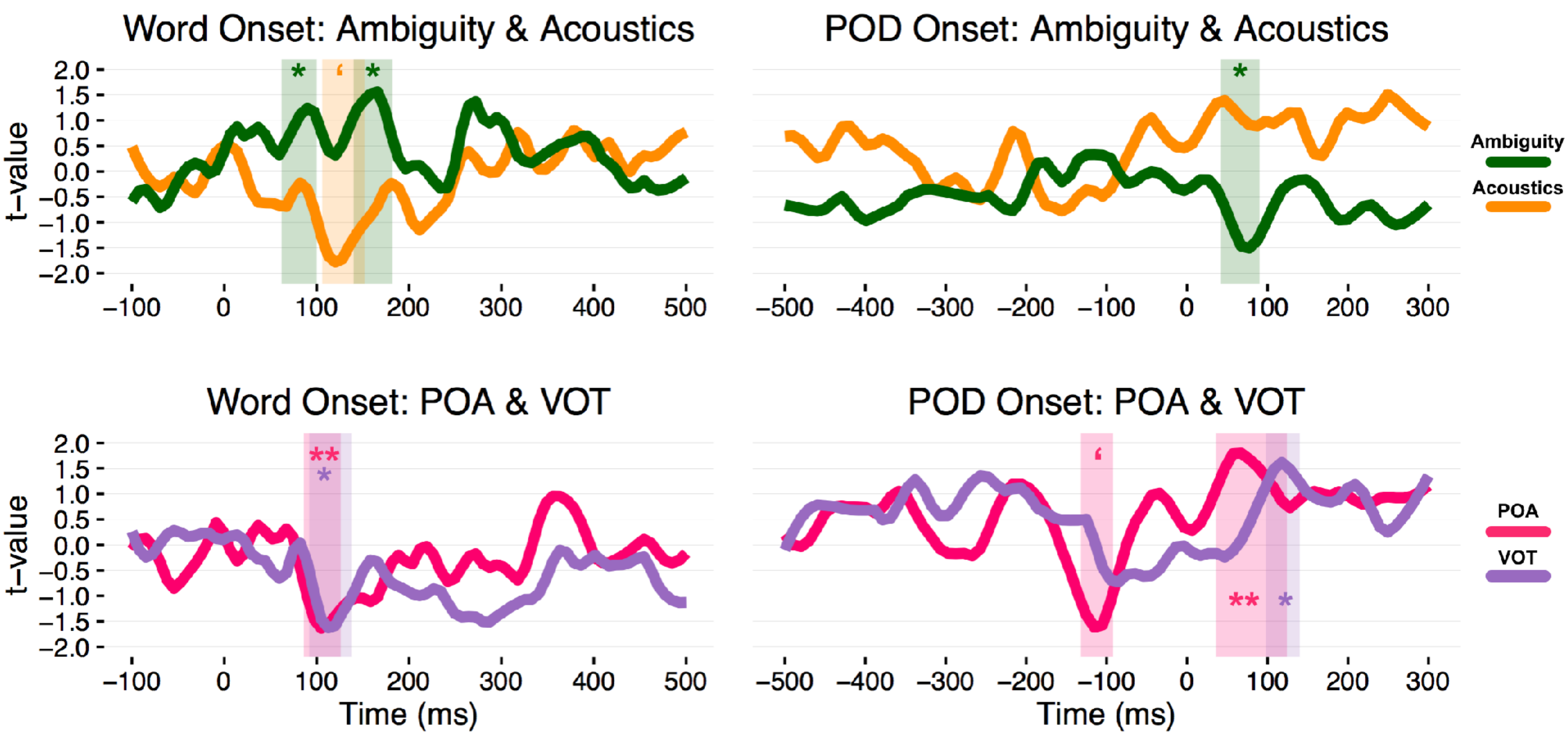
Thank you!

NEW YORK UNIVERSITY

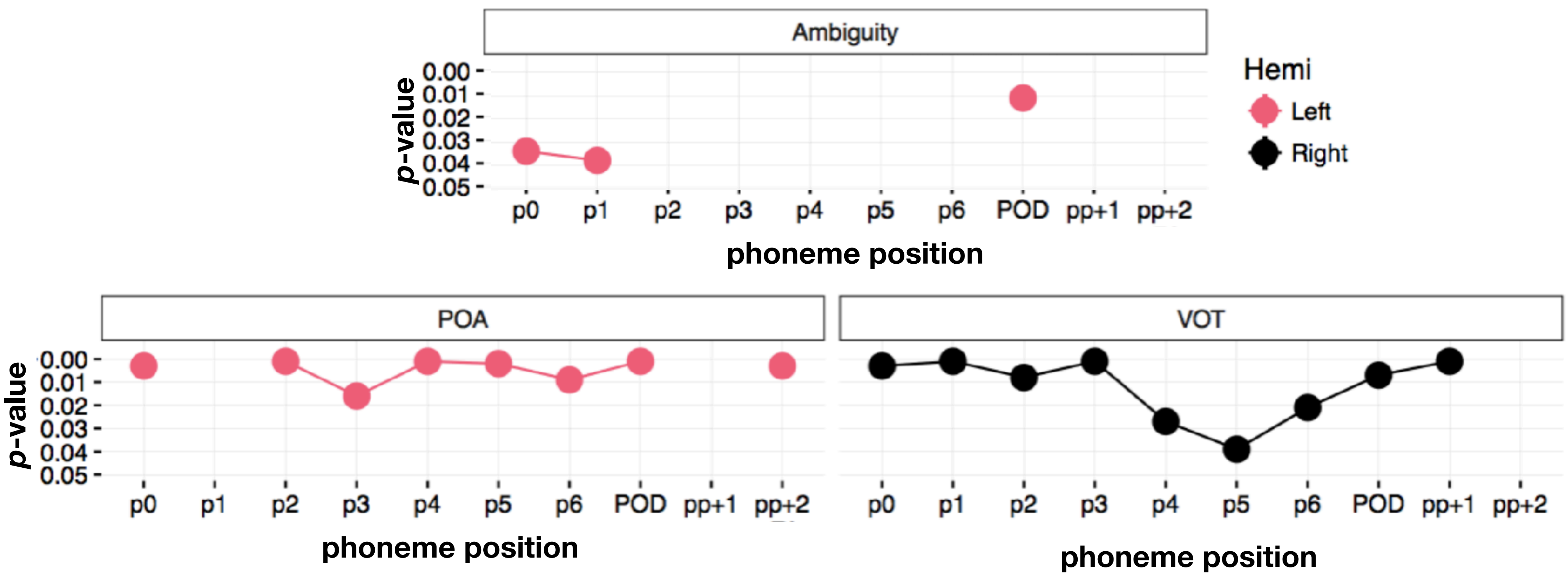
Example Stimuli Pairs

potent	total
dwindle	twinkle
bubbly	publish
primal	triumph
direct	tirade
crash	grasp
democratic	temporary
chemically	temperature
commodity	tomorrow
badger	pageant
percolate	turkeys
crochet	grotesque

bazaar	position
choir	twilight
decades	technician
dreadlock	treadmill
delaware	telephone
capitalise	tapestries
curling	girlish
depositor	topography
balloon	pollute
caucuses	talkative
blunt	plunge
beneficial	penicillin

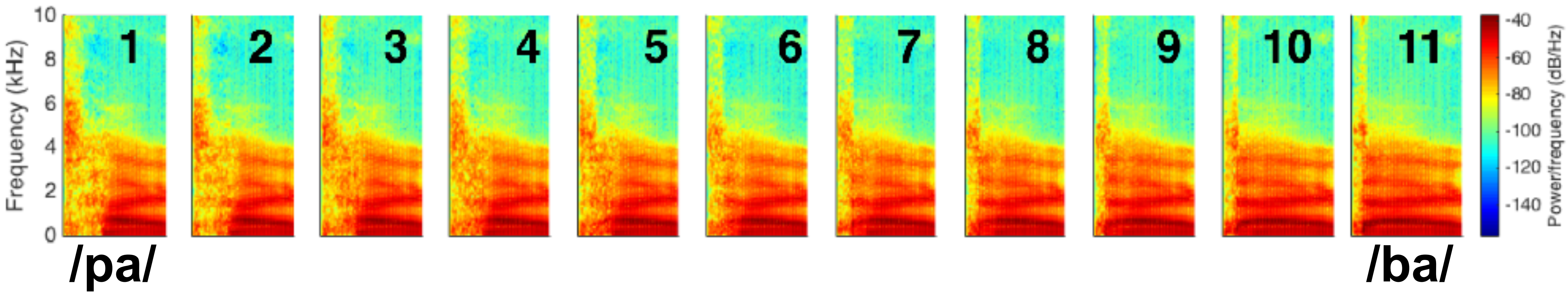

Word Onset: Ambiguity & Acoustics
POD Onset: Ambiguity & Acoustics
Word Onset: POA & VOT
POD Onset: POA & VOT
t–value
Time (ms)
Ambiguity
Acoustics
POA
VOT
**
*

Reactivation in Intermediate Positions

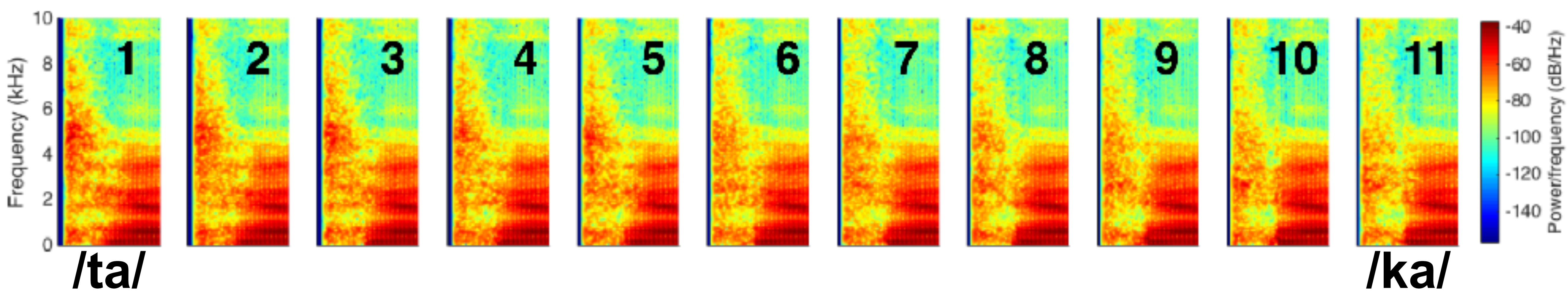


Experiment 1: Design & Materials

- Voice onset time (VOT) - {p-b, t-d, k-g}

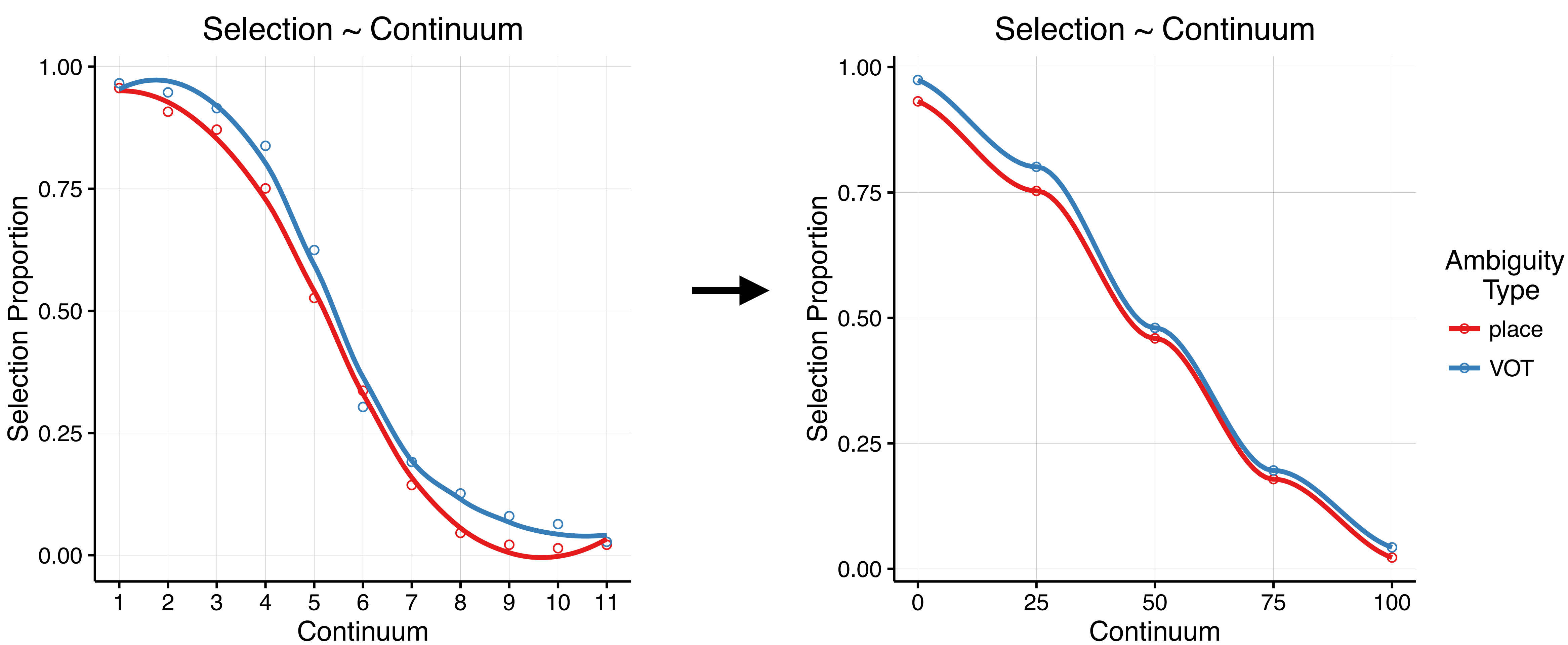


- Place of articulation (PoA) - {t-k, p-t}

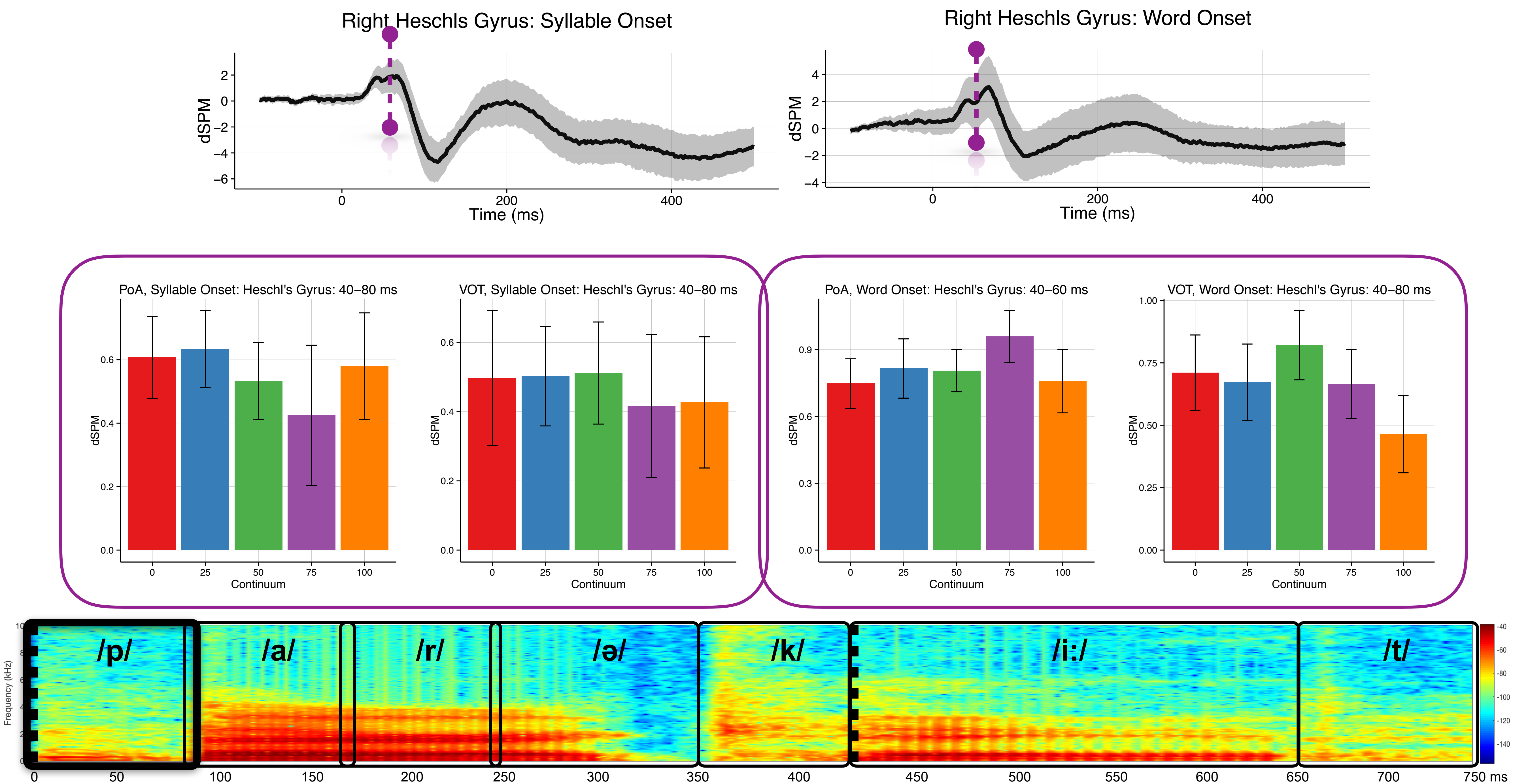


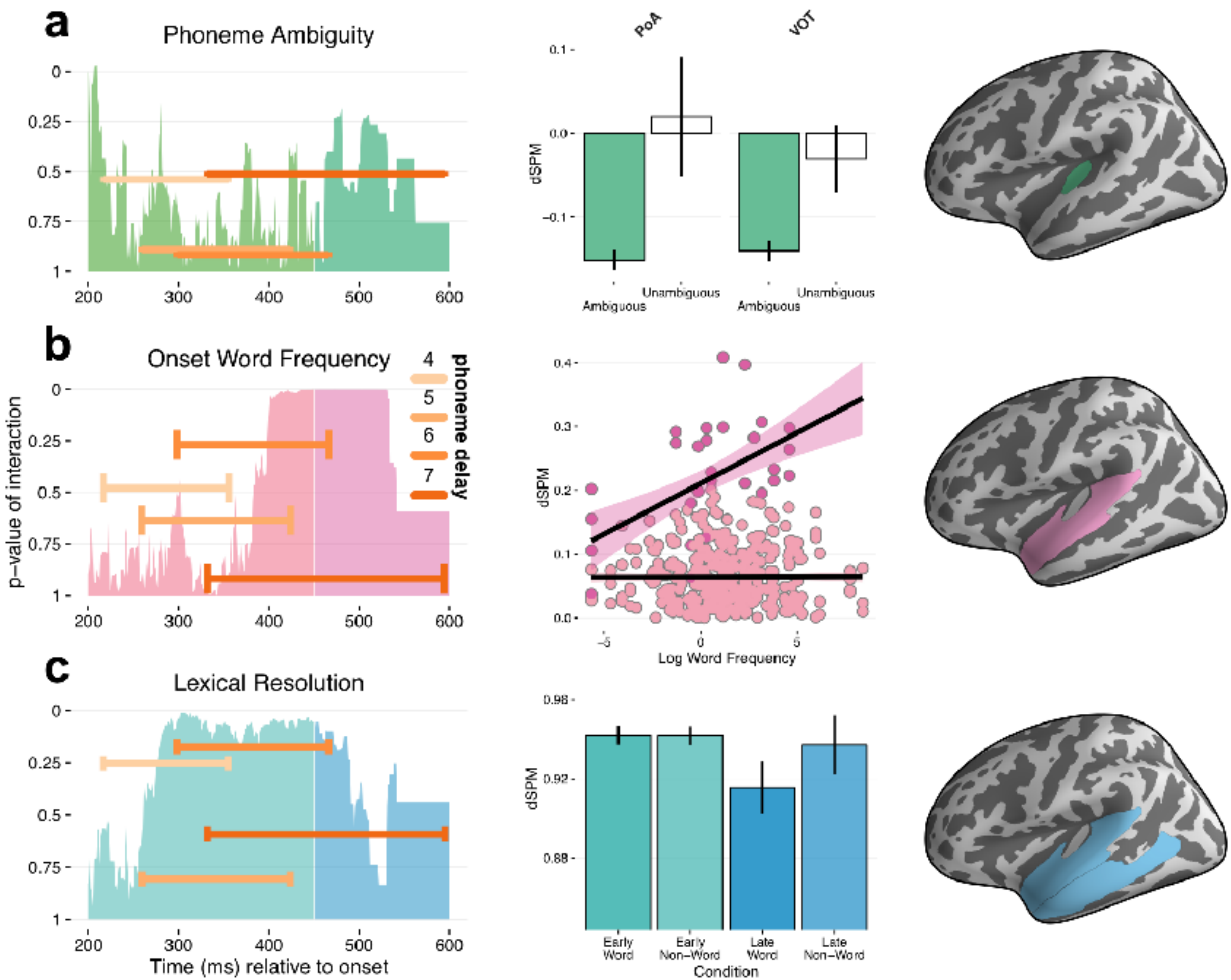
Experiment 1: Design & Materials

- Re-sampled the continuum to match perceptual categorisation



No Ambiguity Effect in Right Hemisphere





a
Phoneme Ambiguity
PoA
VOT
Ambiguous
Unambiguous
dSPM
b
Onset Word Frequency
phoneme delay
4
5
6
7
Log Word Frequency
c
Lexical Resolution
p−value of interaction
Time (ms) relative to onset
Early Word
Early Non−Word
Late Word
Late Non−Word
Condition