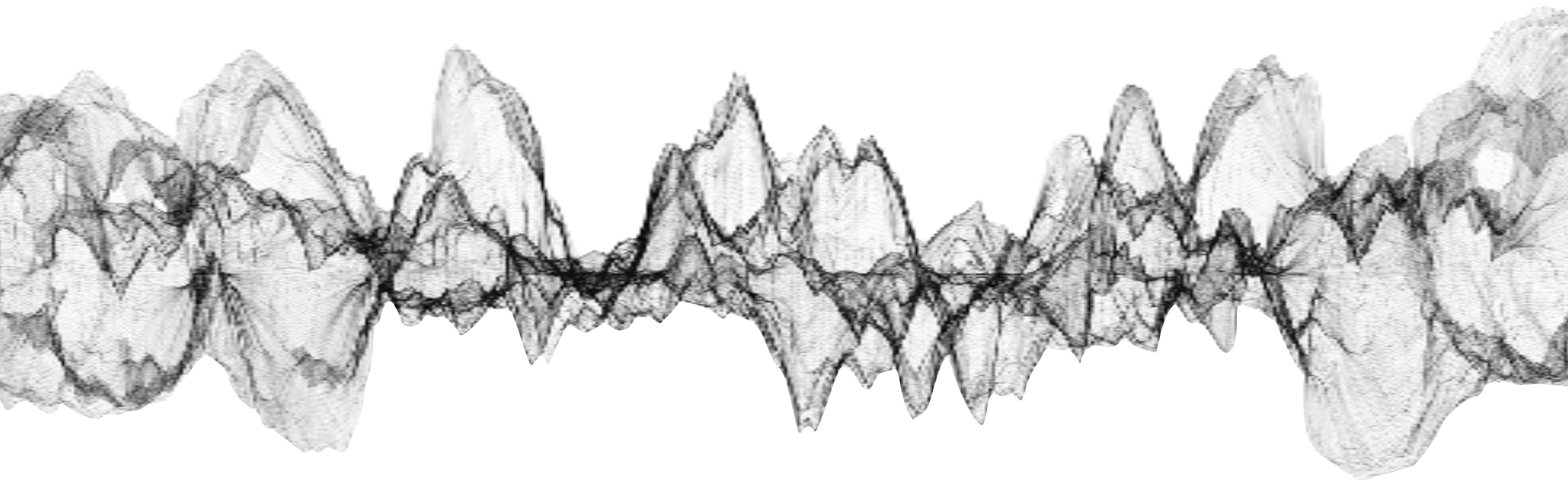




In spoken word recognition, the future predicts the past

Laura Gwilliams

5th June 2017



Road Map

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

Future Influences on Perception

Future Influences on Perception

- Speech is an inherently **noisy and ambiguous** signal

Future Influences on Perception

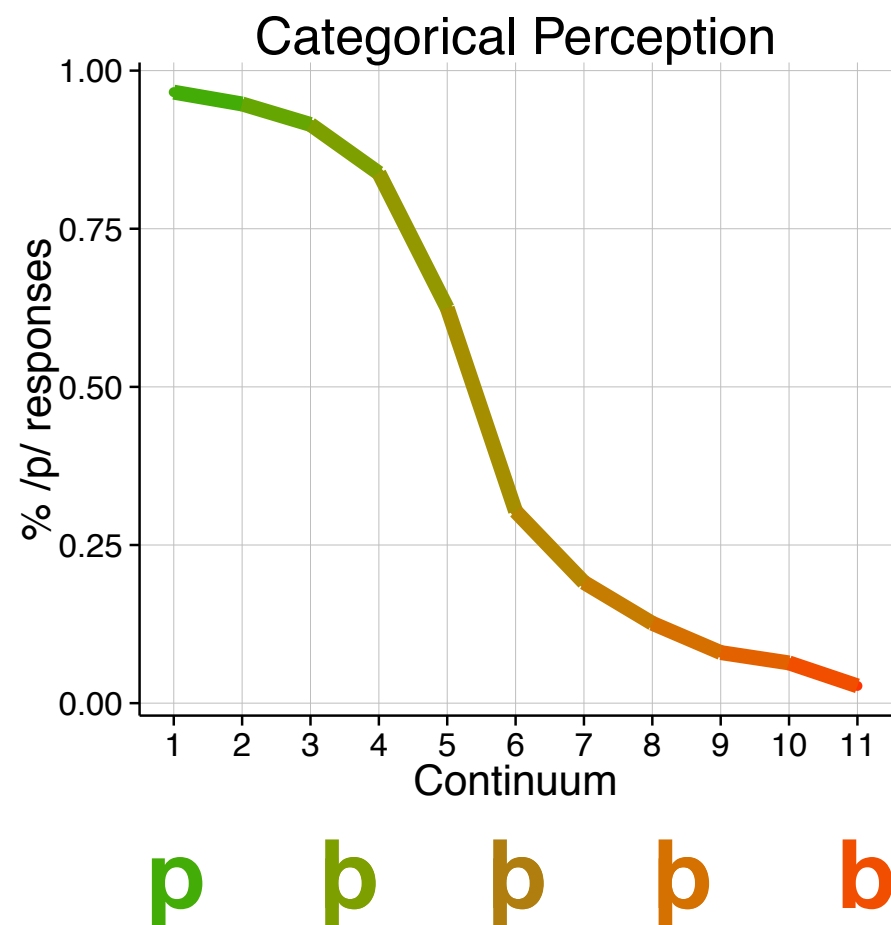
- Speech is an inherently **noisy and ambiguous** signal
- To fluently derive meaning, listeners must **integrate top-down** contextual information to guide their interpretation

Future Influences on Perception

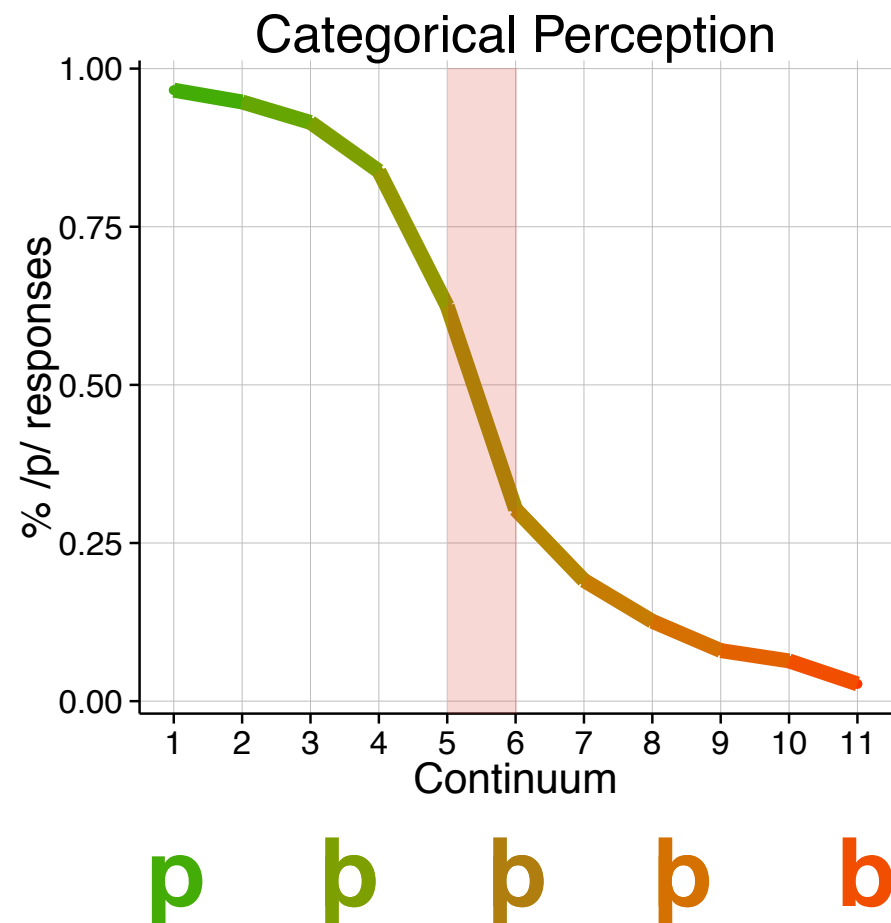
- Speech is an inherently **noisy and ambiguous** signal
- To fluently derive meaning, listeners must **integrate top-down** contextual information to guide their interpretation
- Top-down input occurring *after* an acoustic signal can be integrated to **affect the perception of earlier sounds**
(Bicknell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

Future Influences on Perception

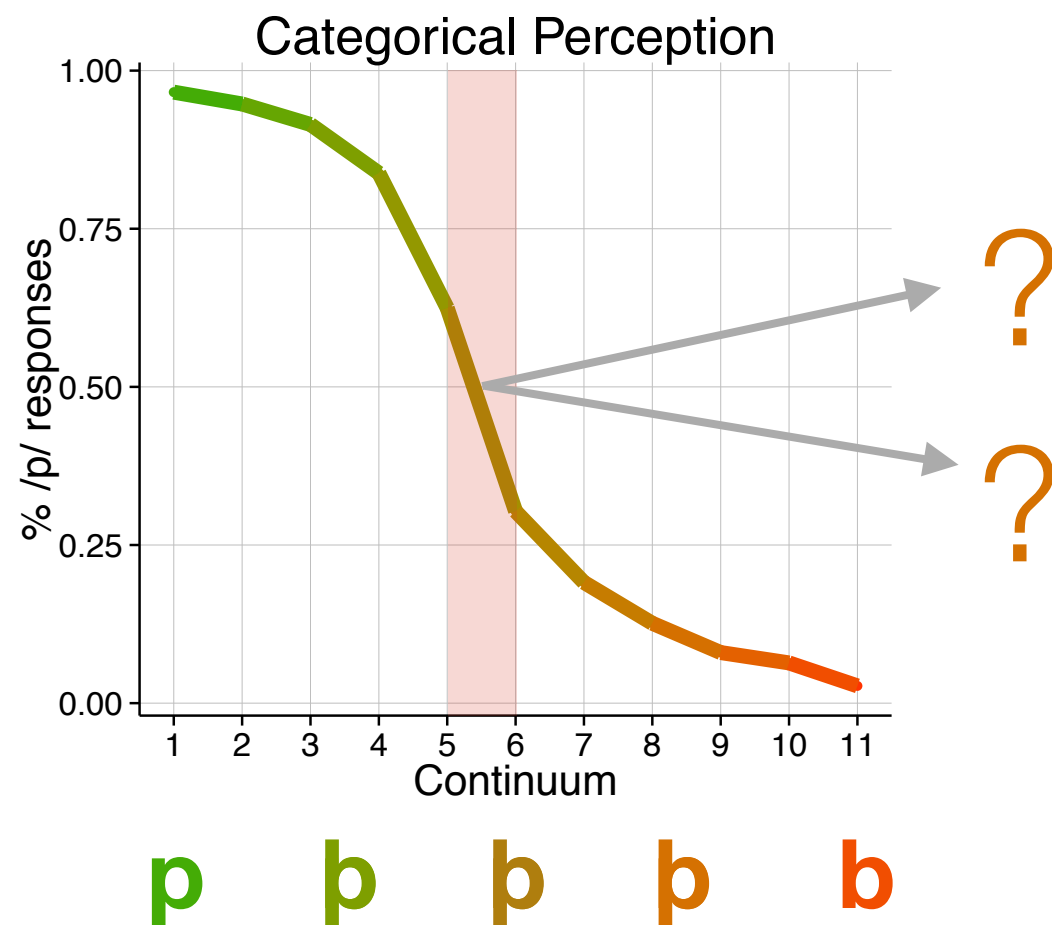
Future Influences on Perception



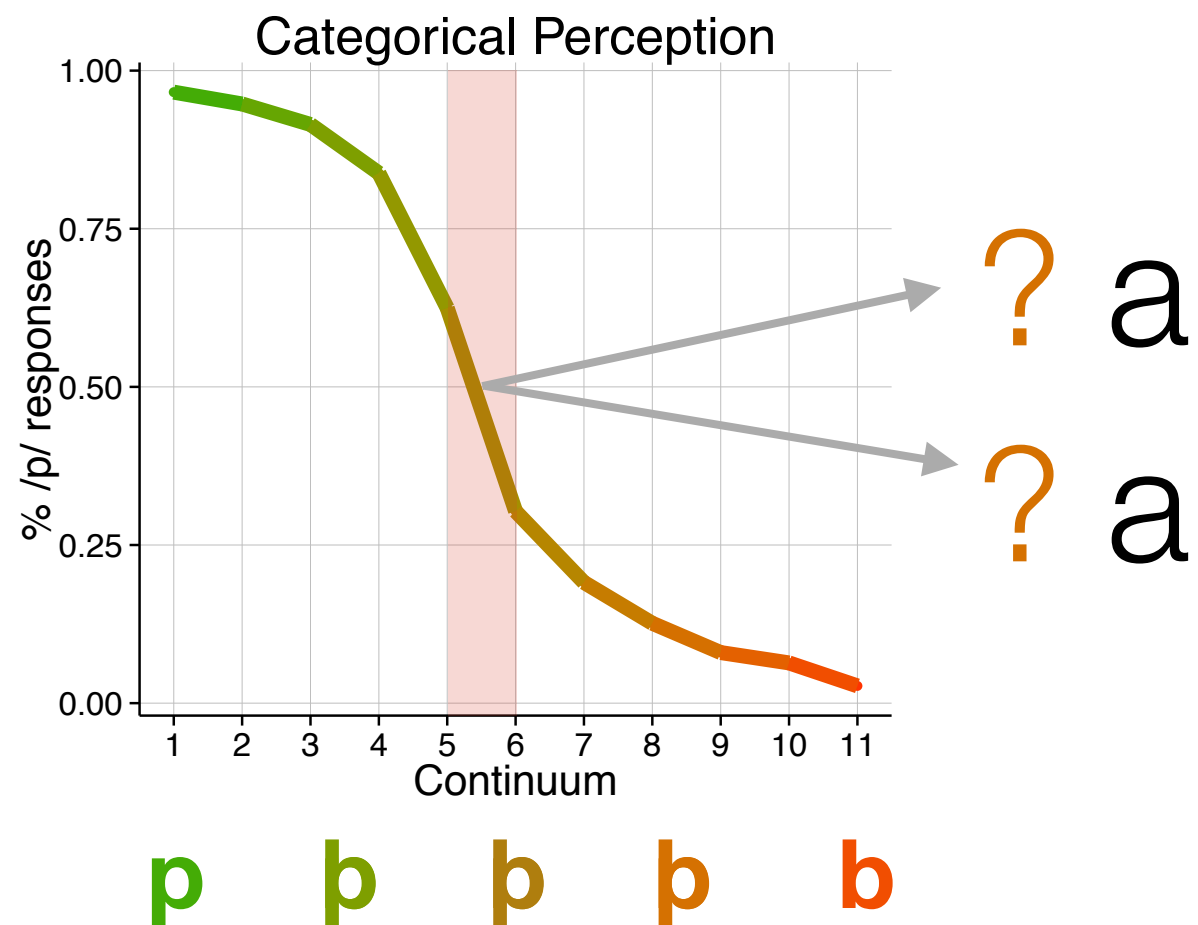
Future Influences on Perception



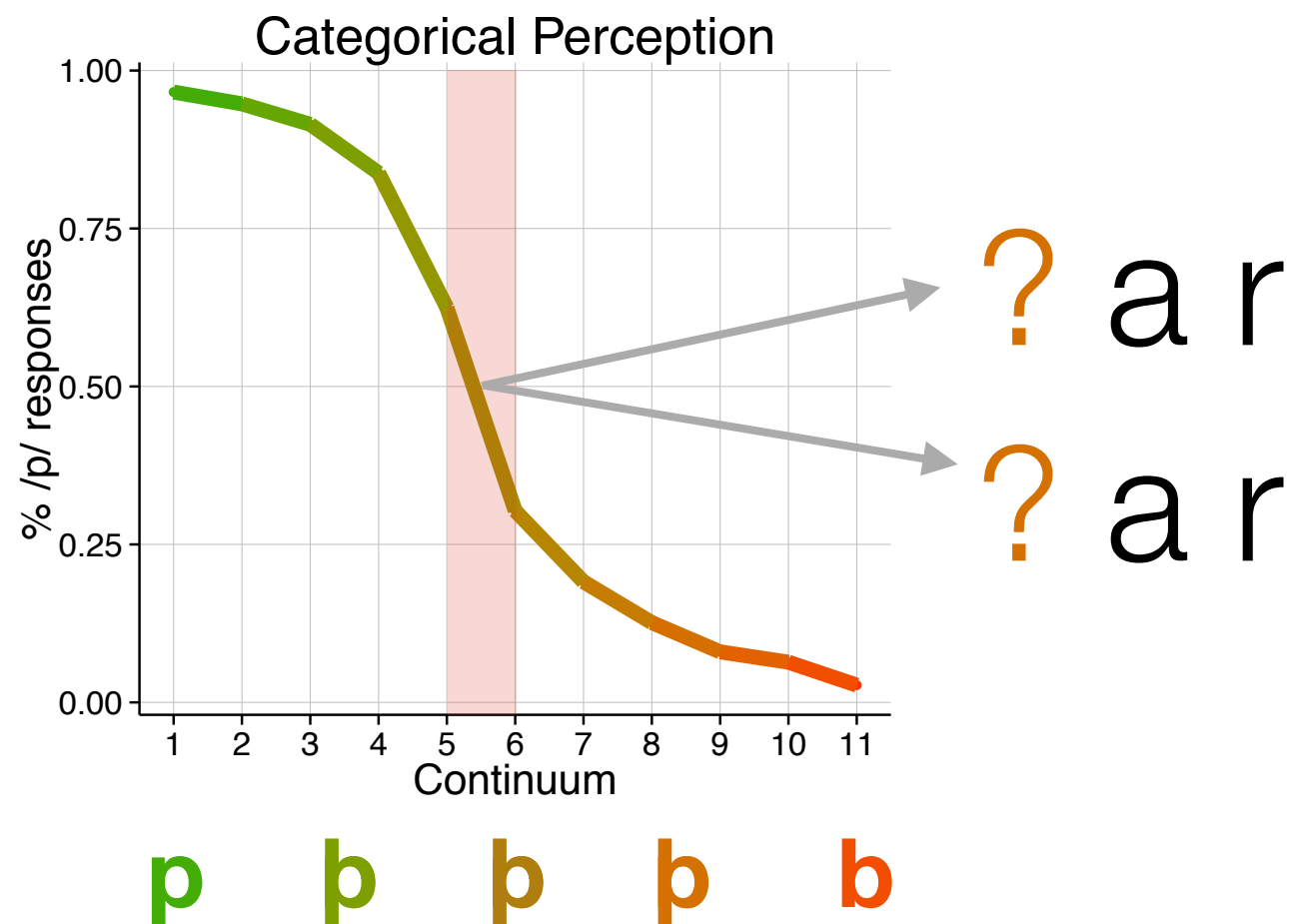
Future Influences on Perception



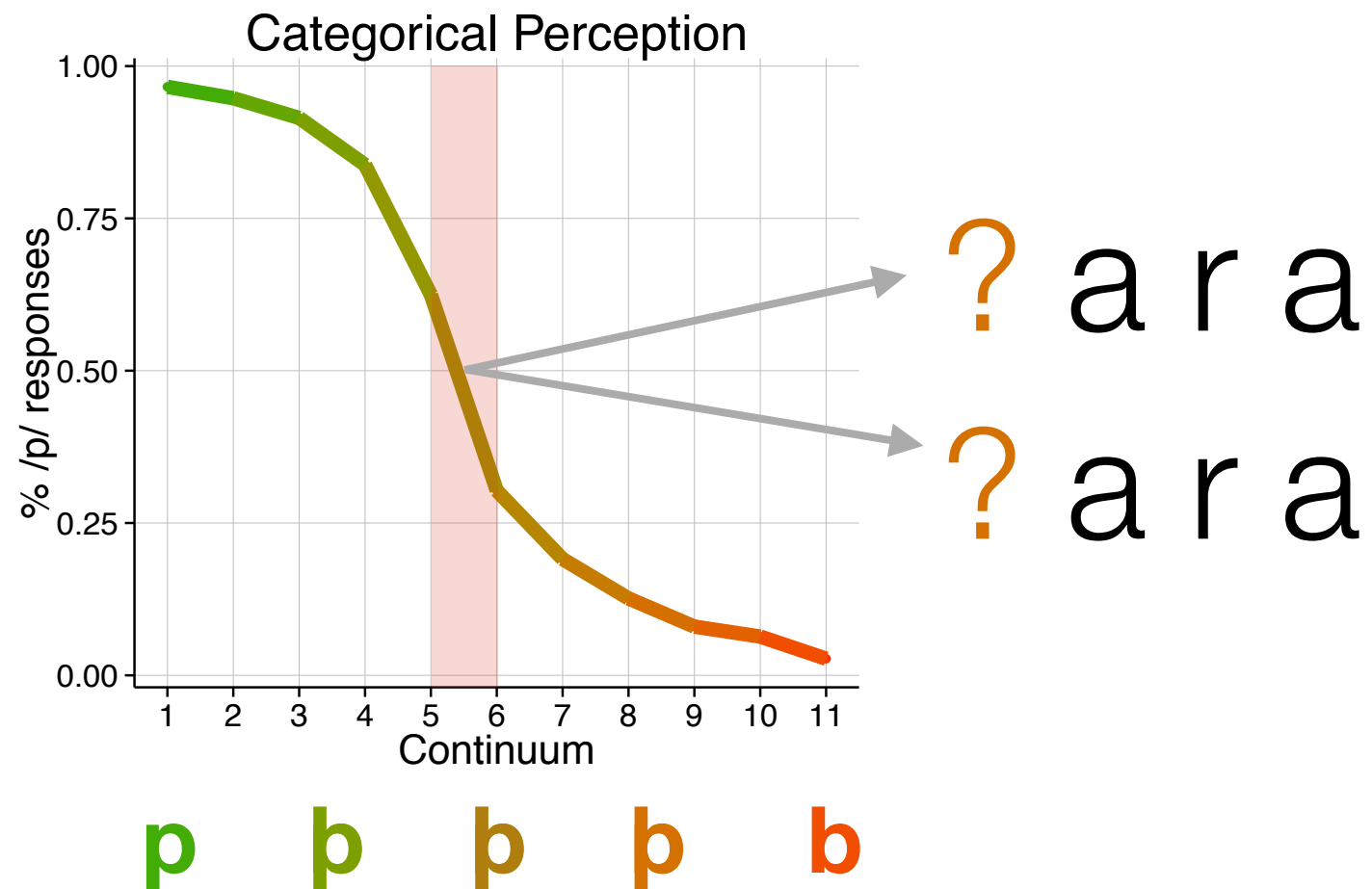
Future Influences on Perception



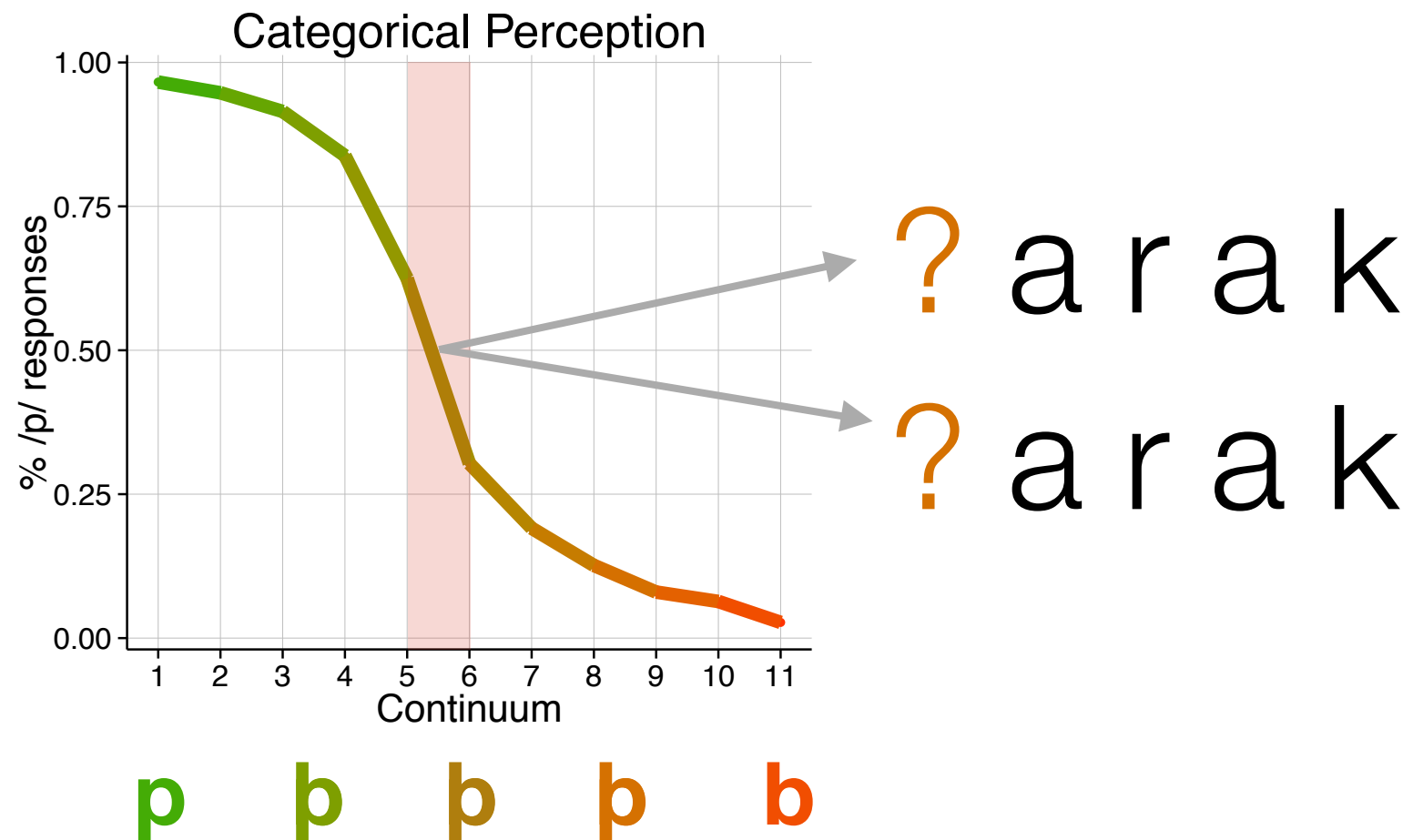
Future Influences on Perception



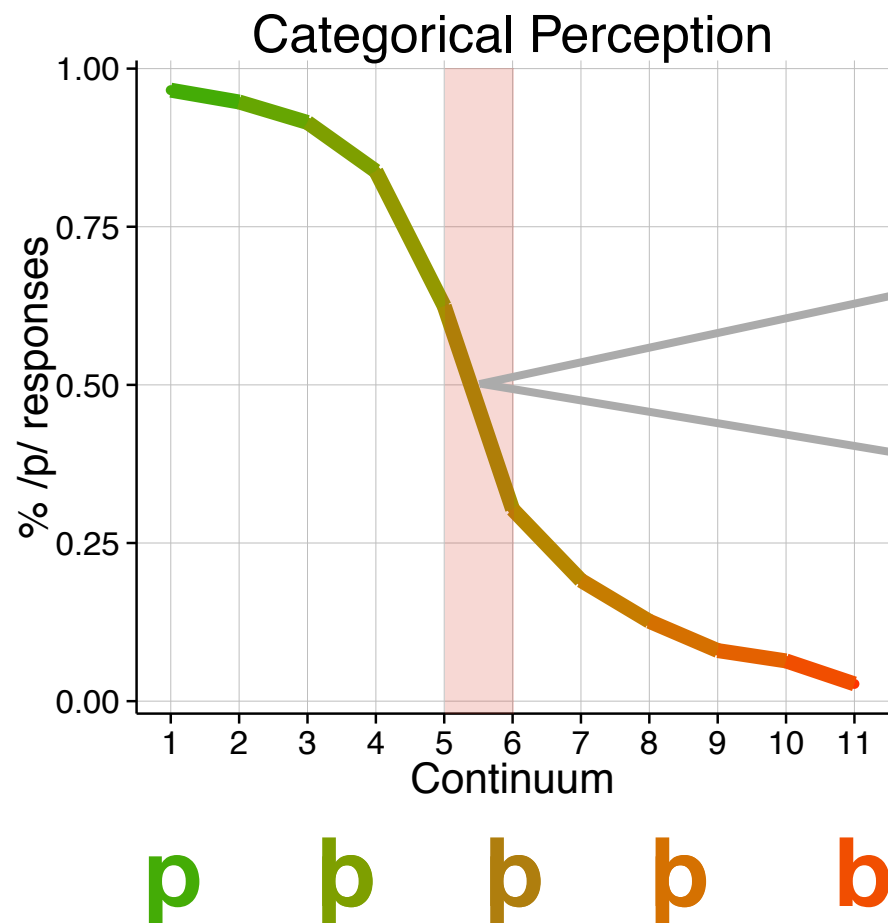
Future Influences on Perception



Future Influences on Perception



Future Influences on Perception

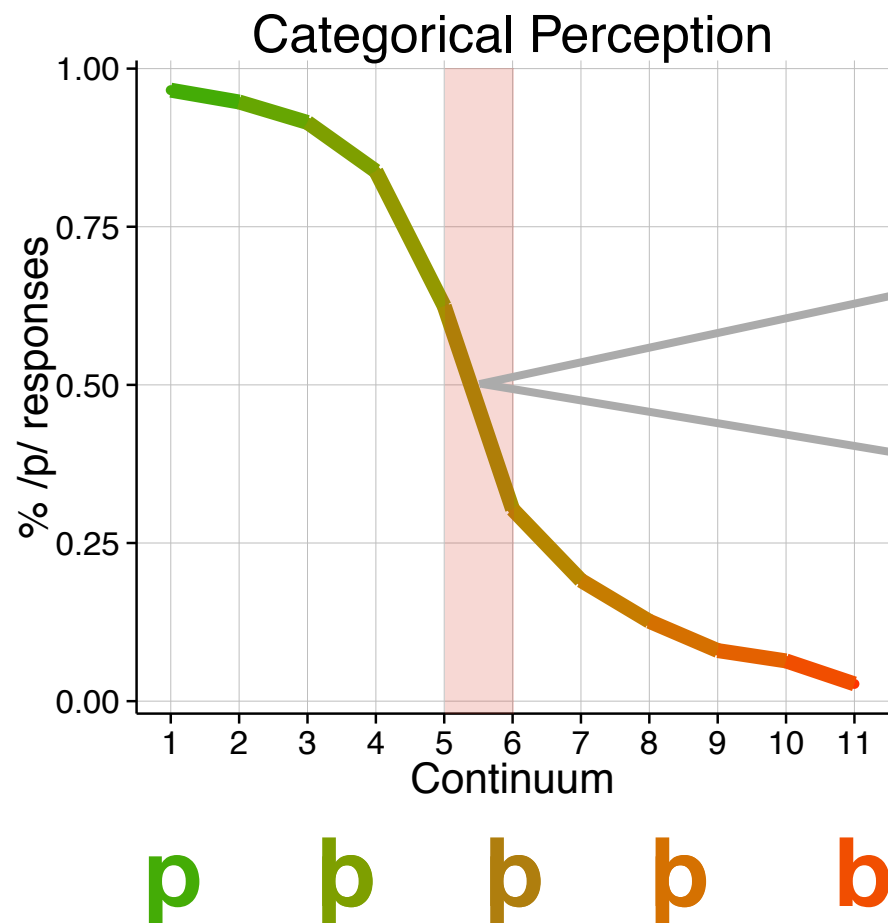


? a r a k **ee** t

? a r a k **ai** d

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

Future Influences on Perception



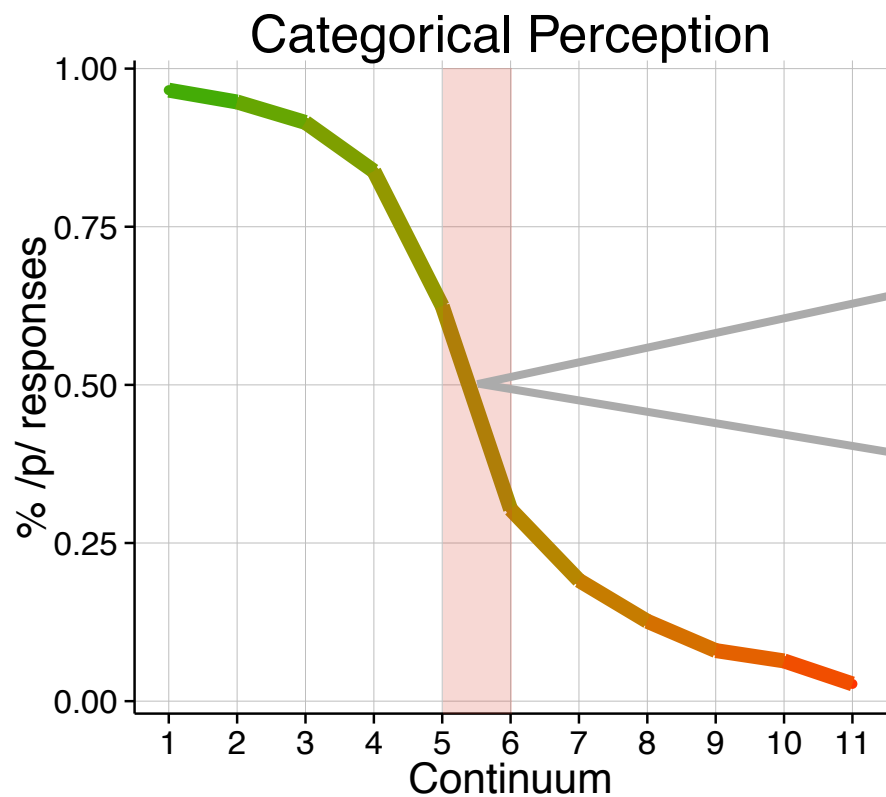
? a r a k **ee** t

? a r a k **ai** d

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

Future Influences on Perception

(this is a parakeet)



p a r a k e e t

b a r a k a i d

p p p p b

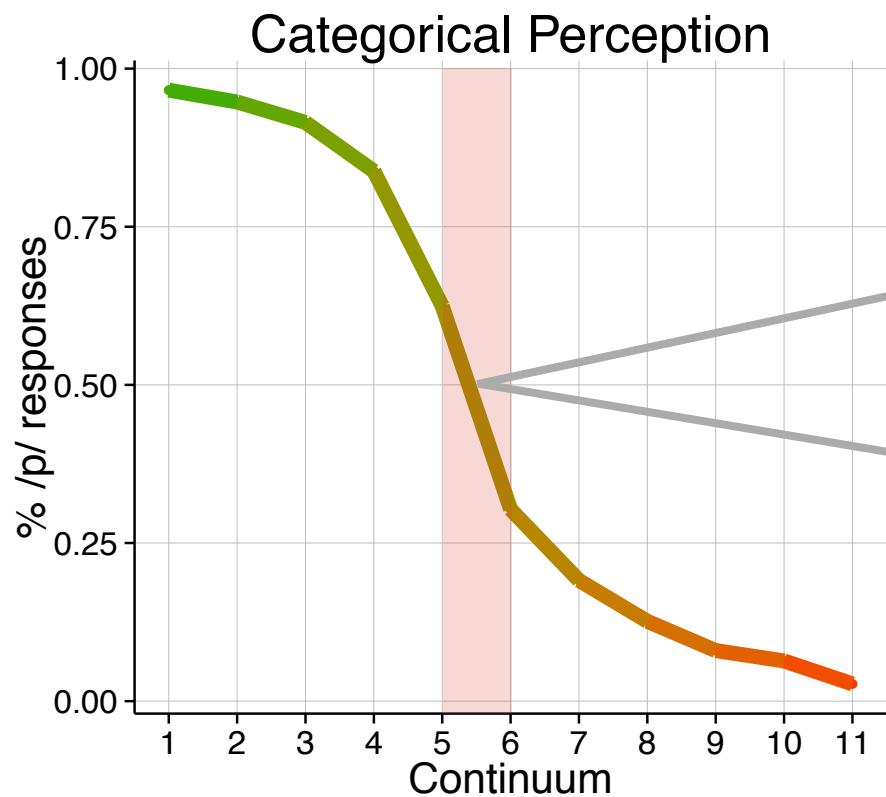


(this is a barricade)

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

Future Influences on Perception

(this is a parakeet)



p b p b b

p a r a k e e t

b a r a k a i d

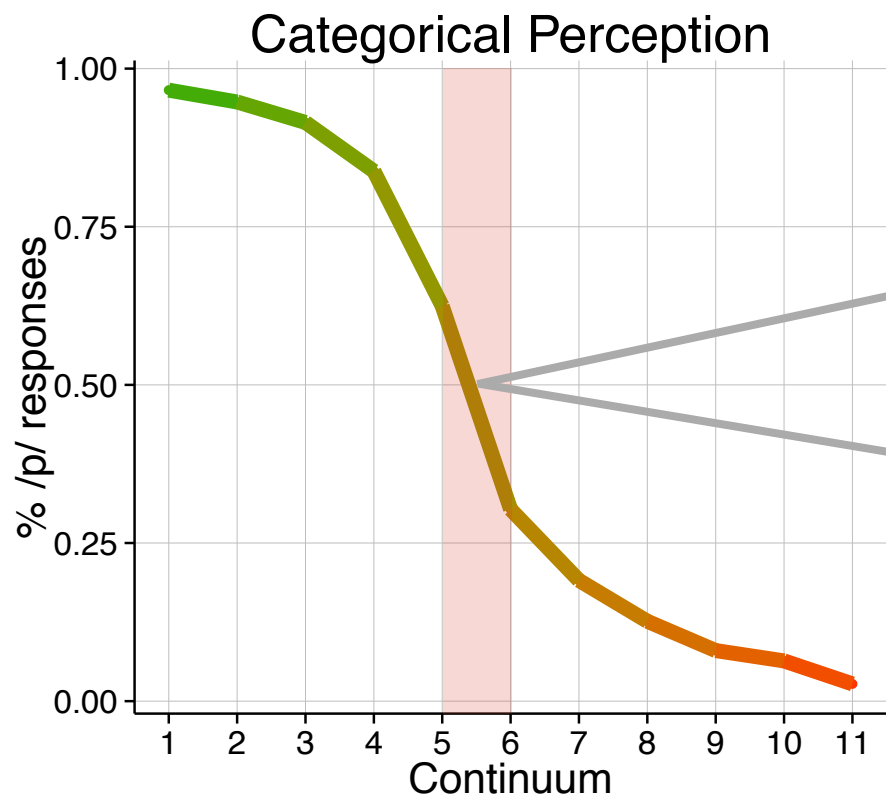


(this is a barricade)

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

Future Influences on Perception

(this is a parakeet)



p p p p b

p a r a k e e t

b a r a k a i d

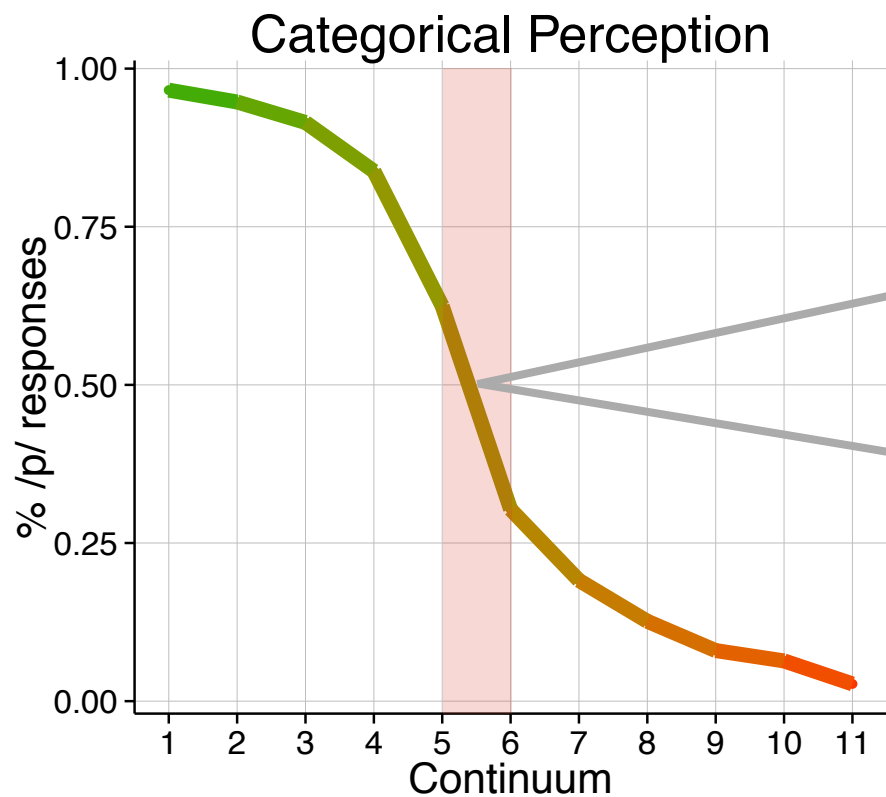


(this is a barricade)

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

Future Influences on Perception

(this is a parakeet)



p a r a k **ee** t

b a r a k **ai** d

p b p b b



(this is a barricade)

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

Today's Questions

p a r a k e e t

Today's Questions

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

?



p a r a k e e t

Today's Questions

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

?



p

a

r

a

k

What are the neural
signatures of ambiguity
resolution?

?



ee

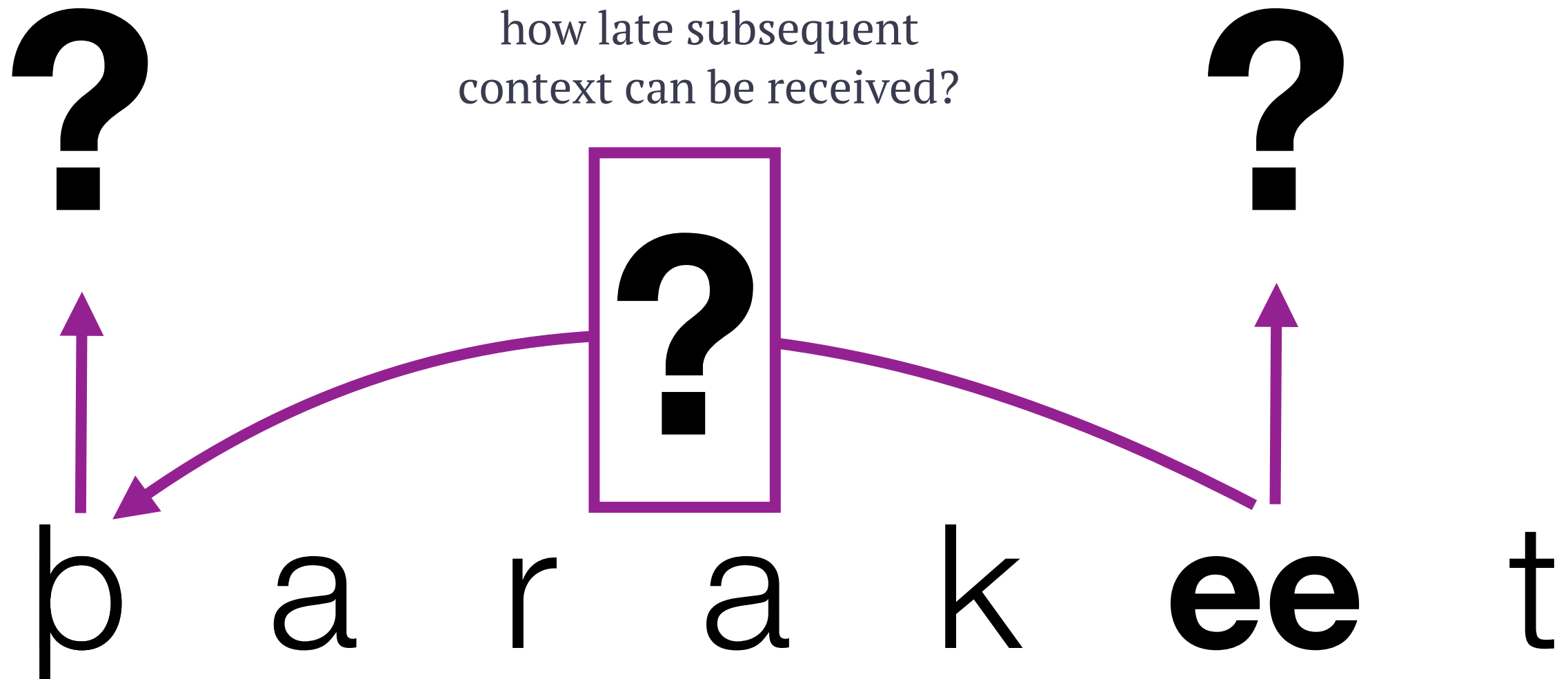
t

Today's Questions

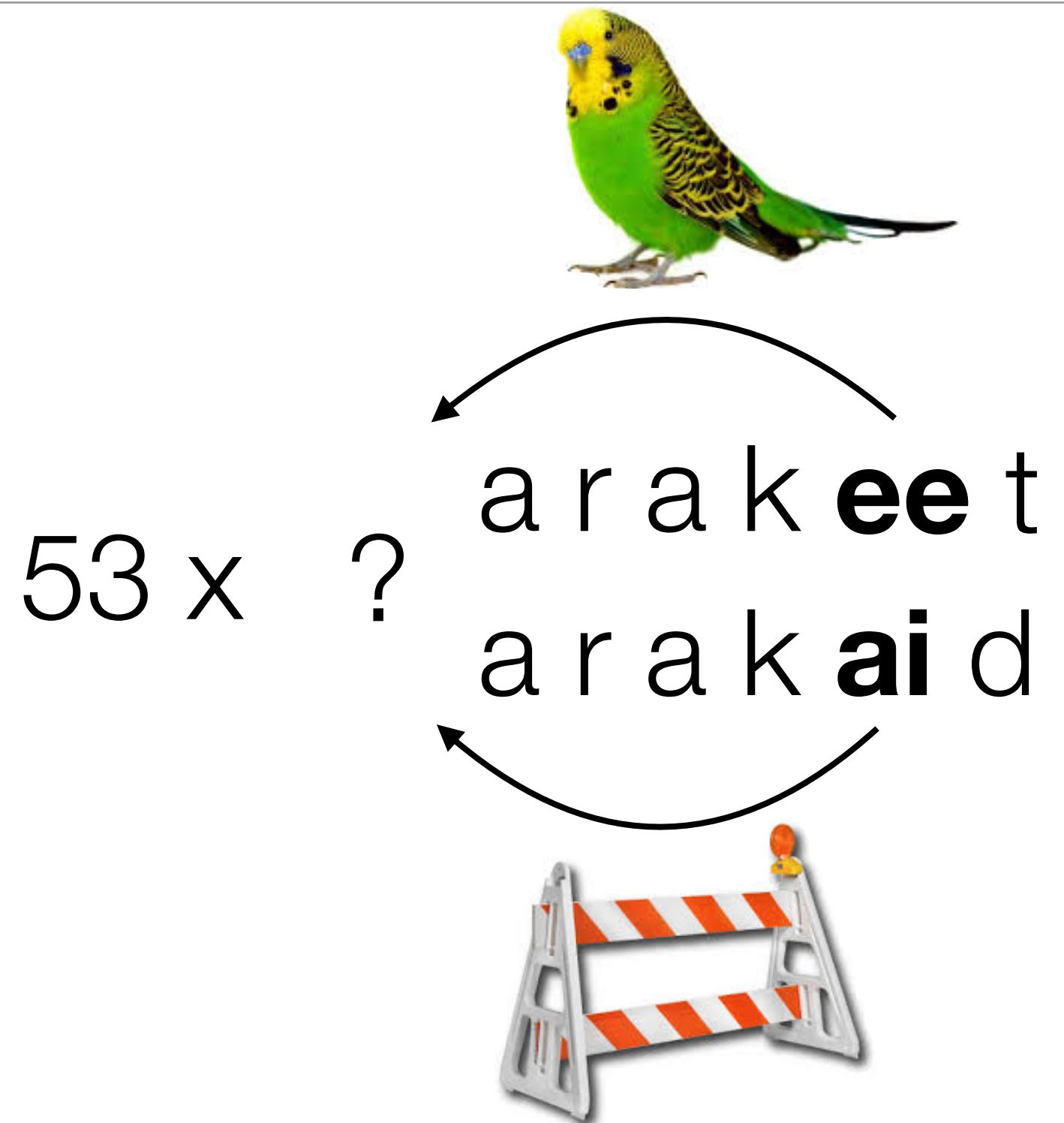
How does the auditory cortex **respond** to phonological ambiguity?

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

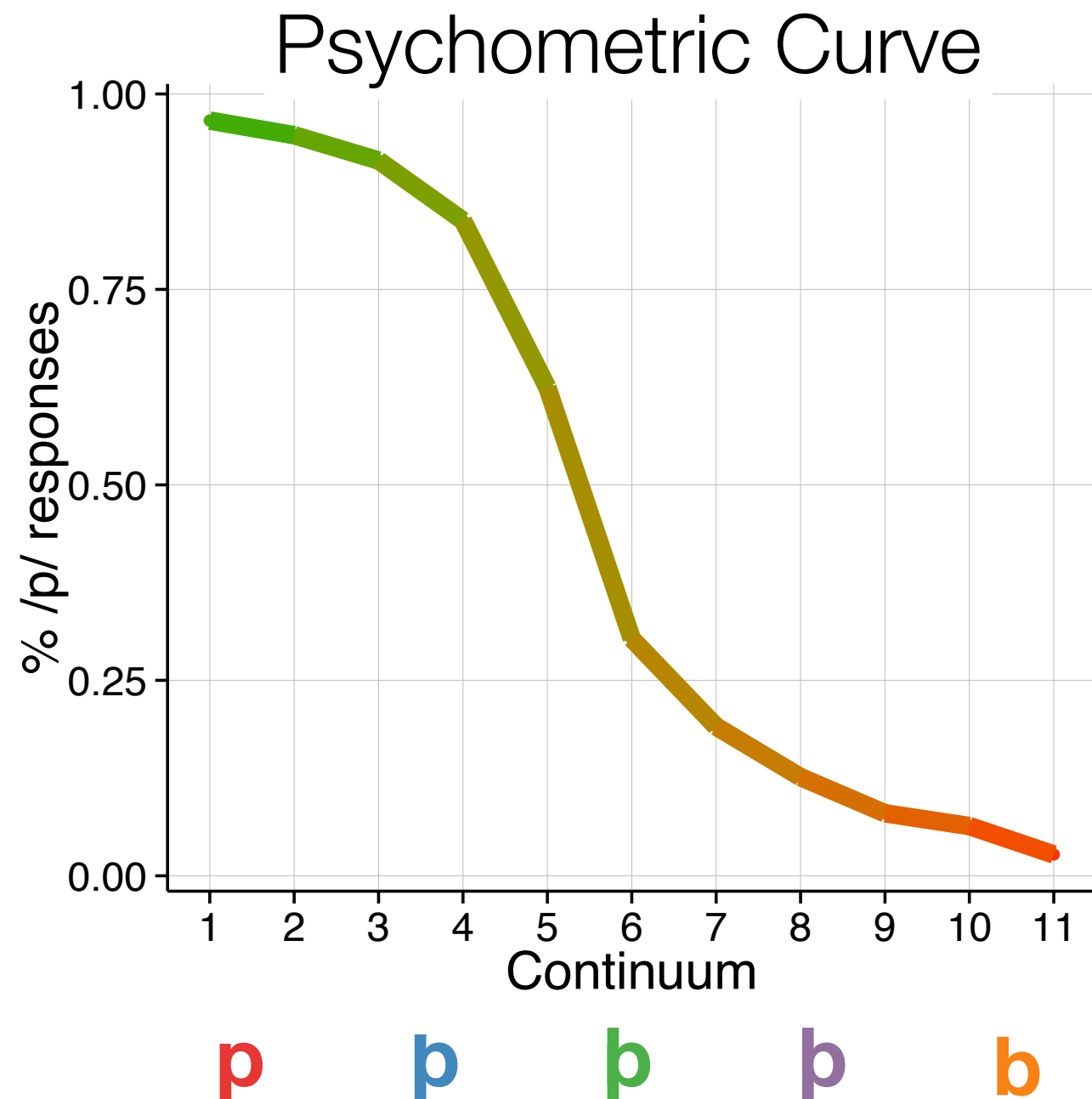
What is the **time-limit** on how late subsequent context can be received?



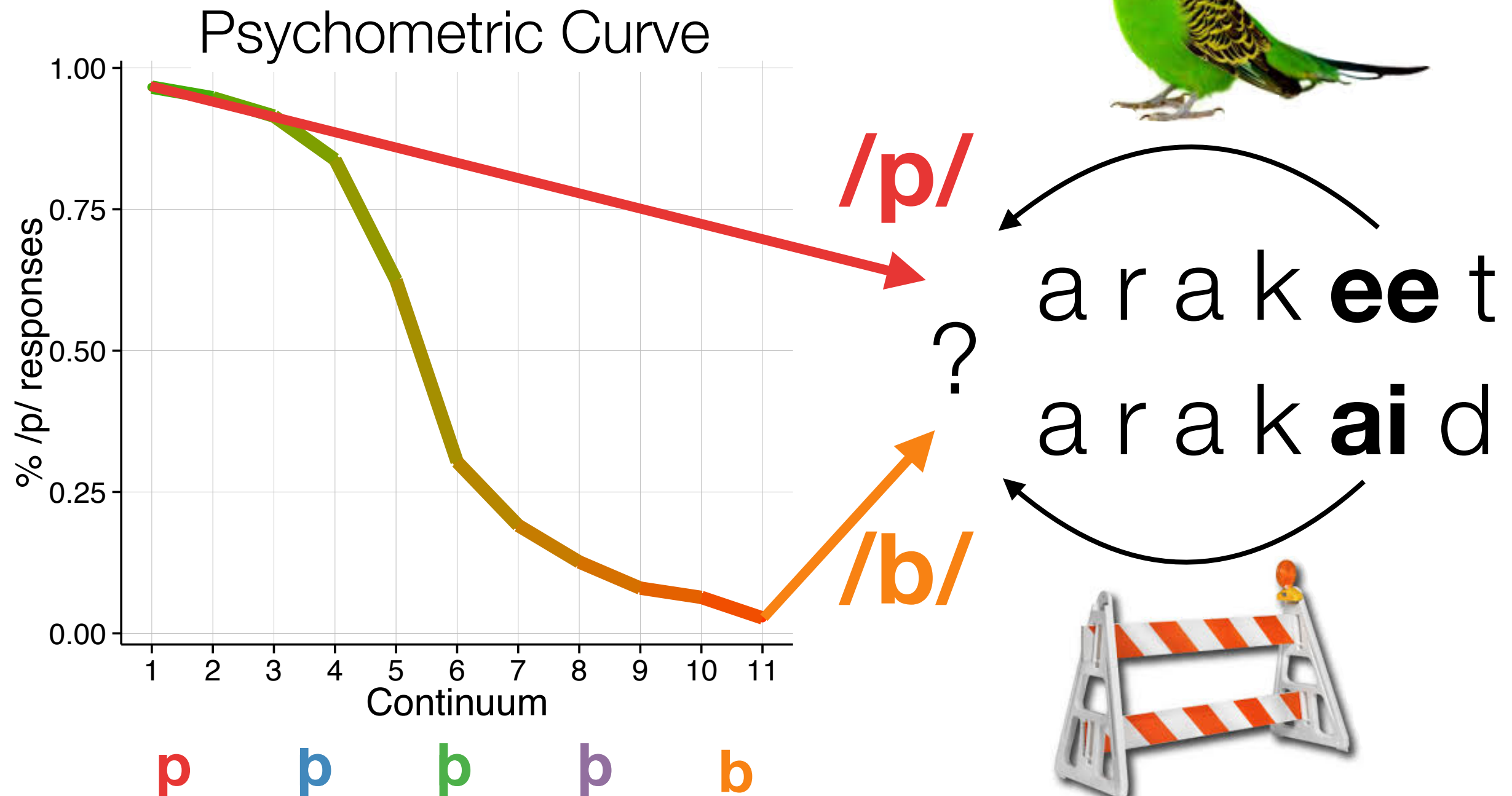
Design & Materials



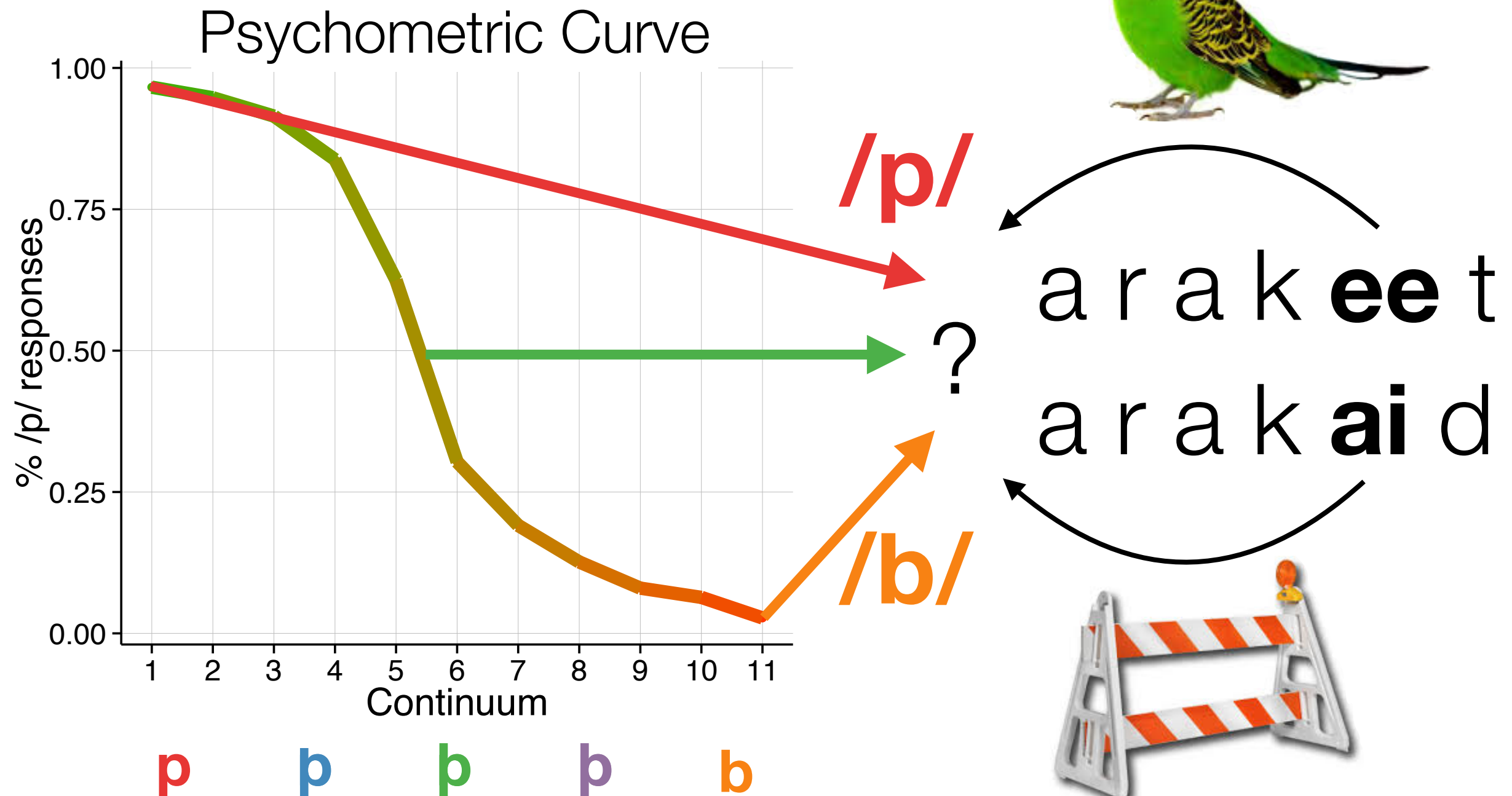
Design & Materials



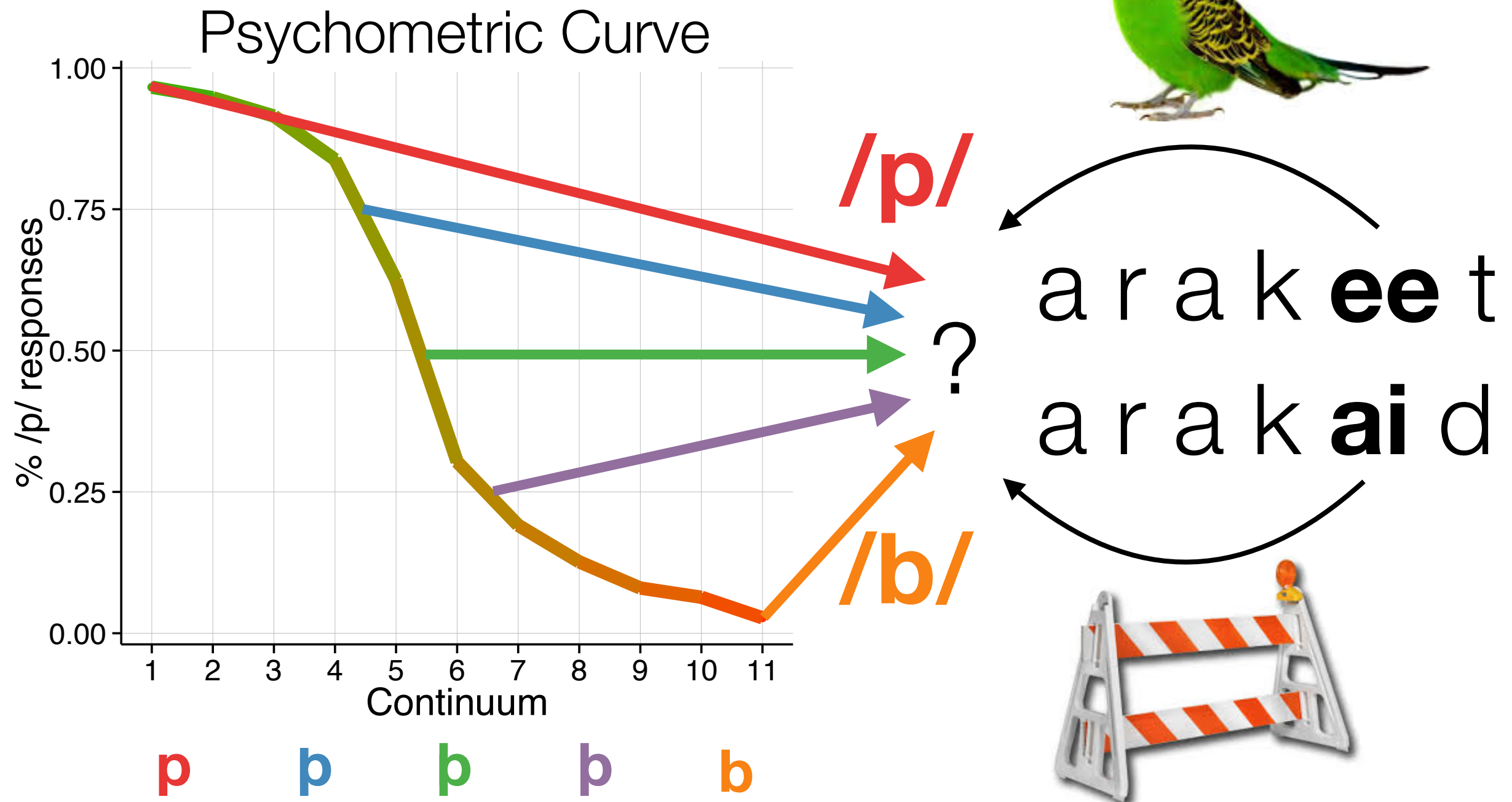
Design & Materials



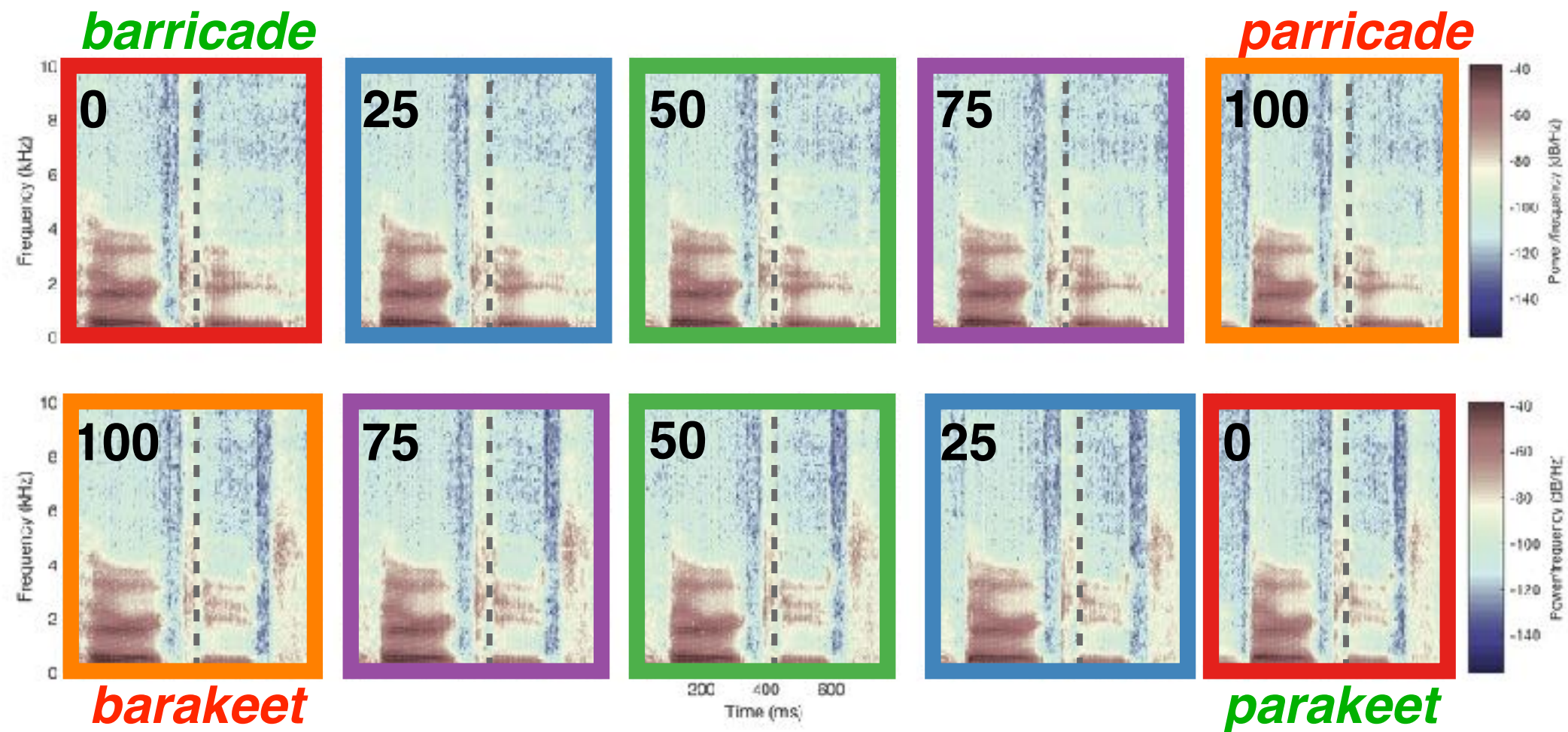
Design & Materials



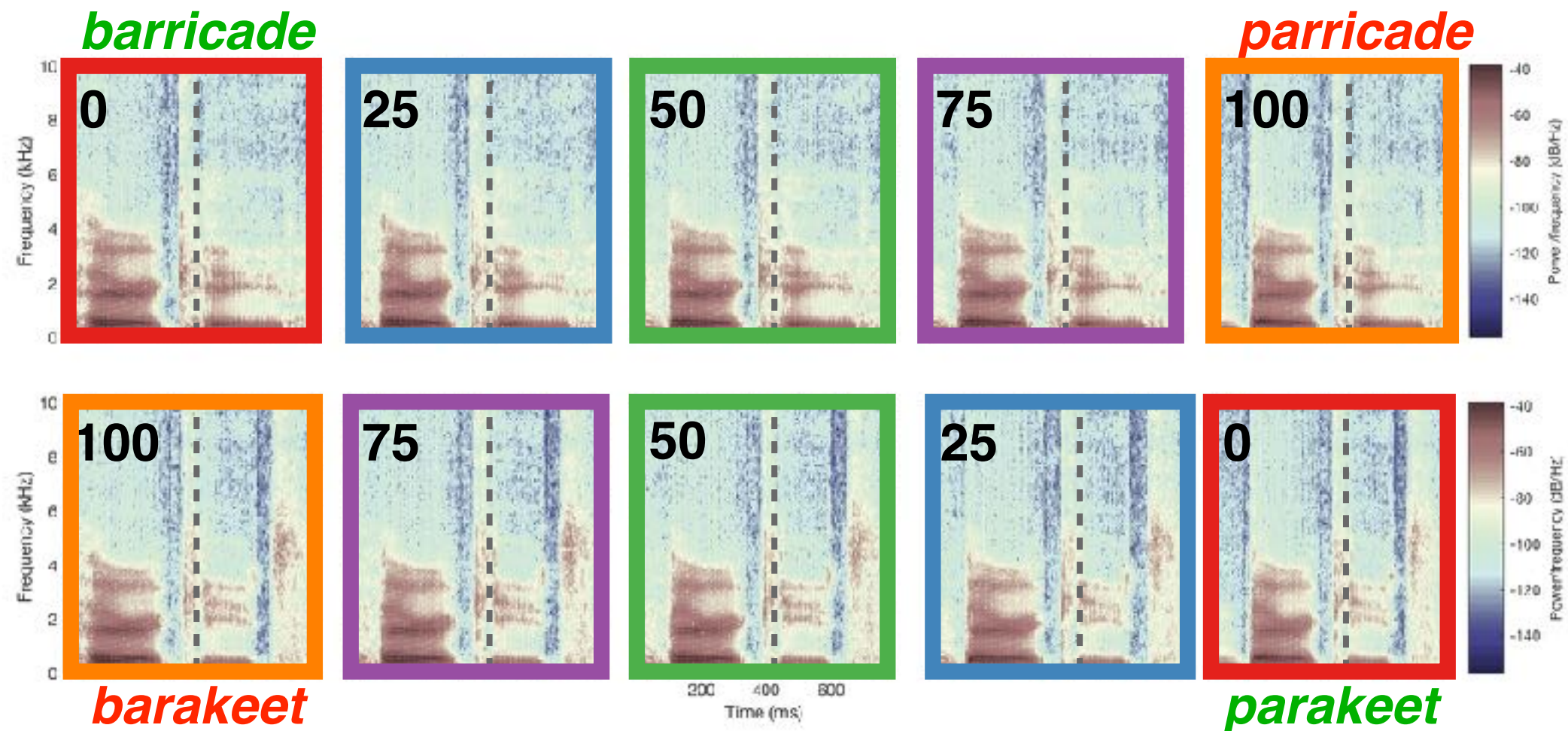
Design & Materials



Design & Materials

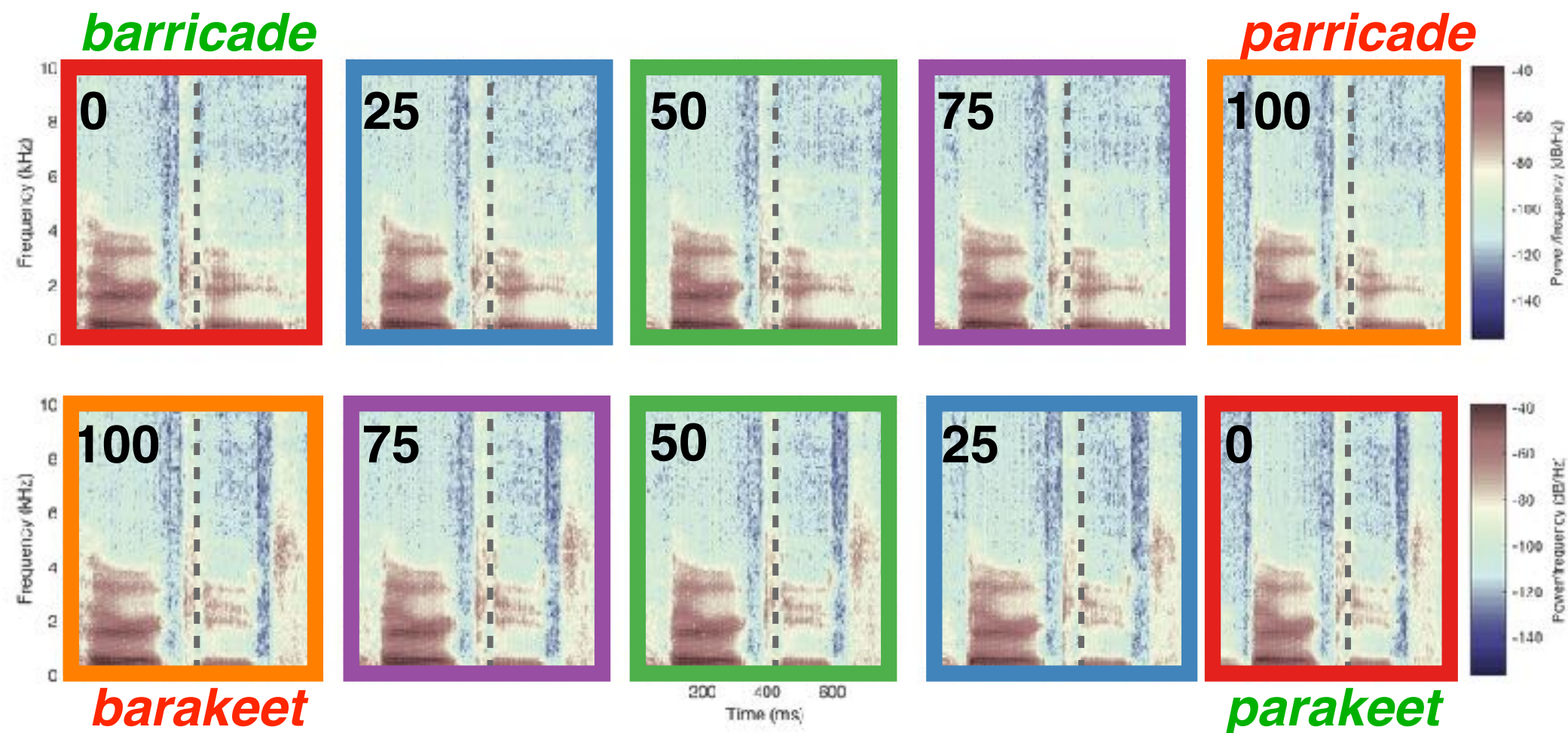


Design & Materials



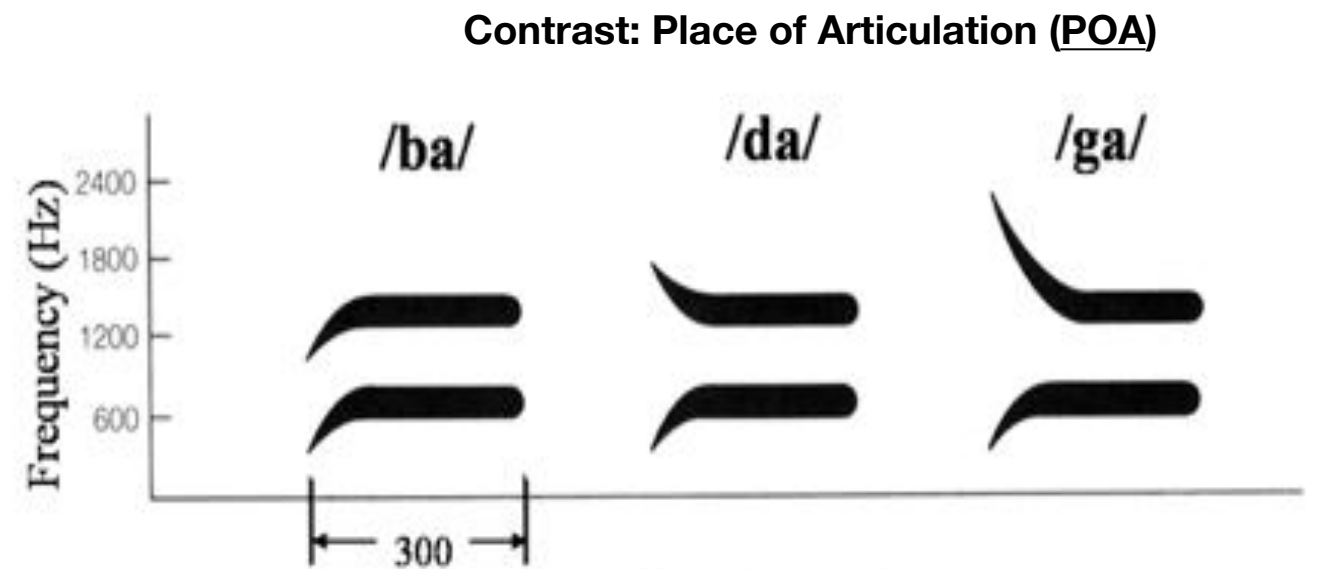
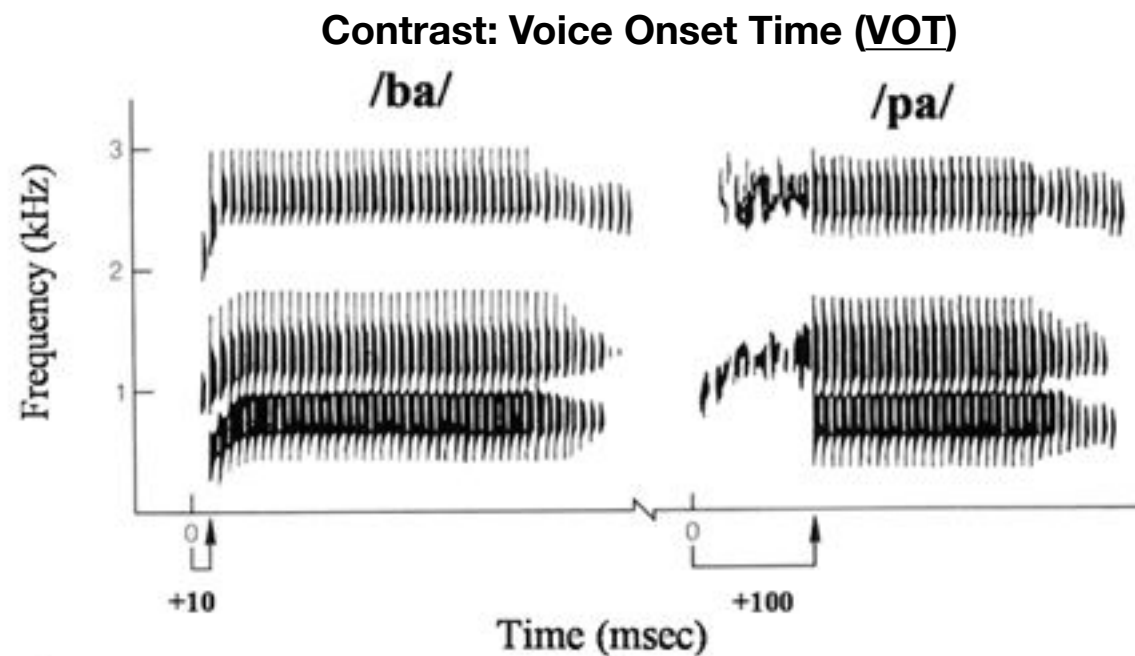
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

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}

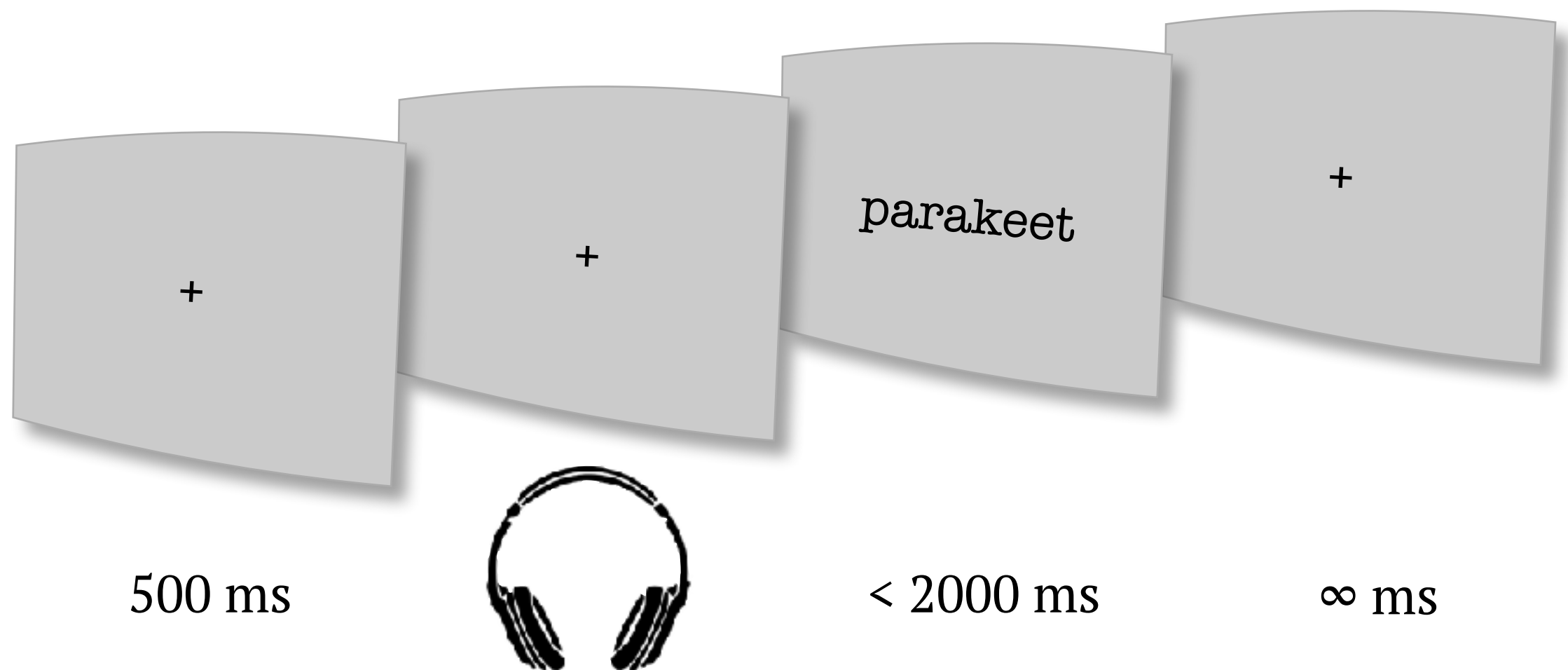
Design & Materials



- 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

Design & Materials



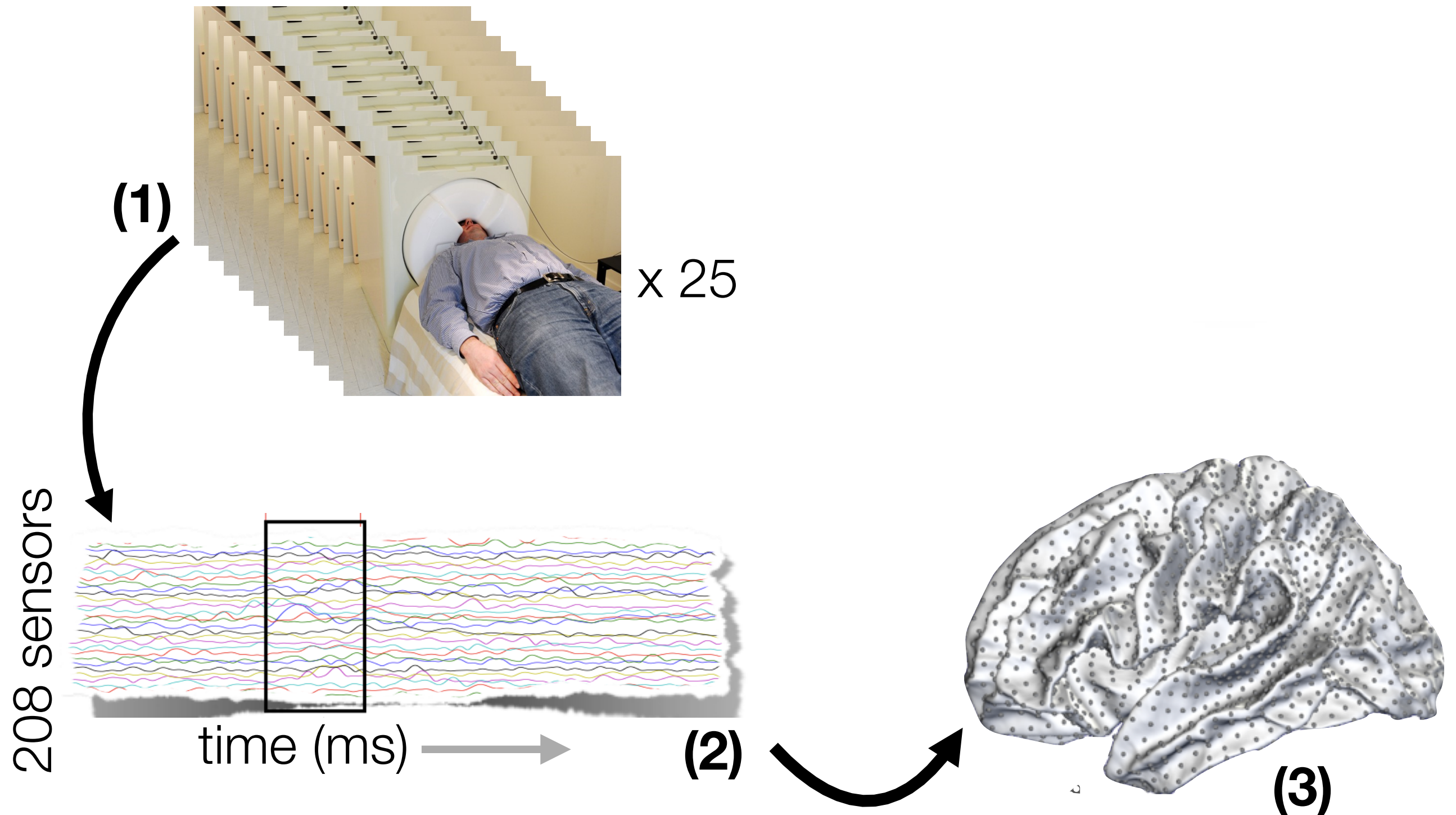
Procedure & Analysis

Procedure & Analysis

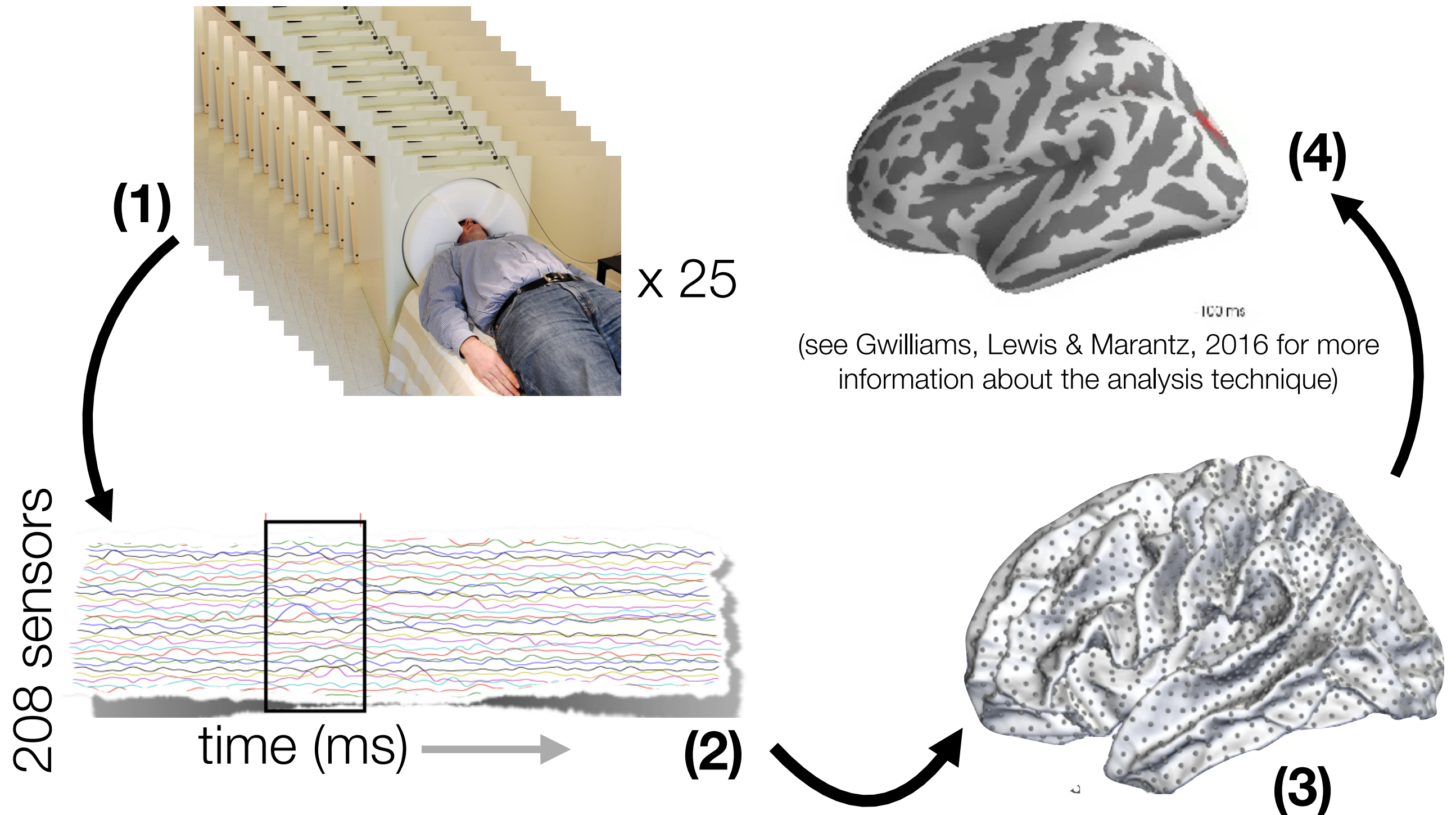


x 25

Procedure & Analysis

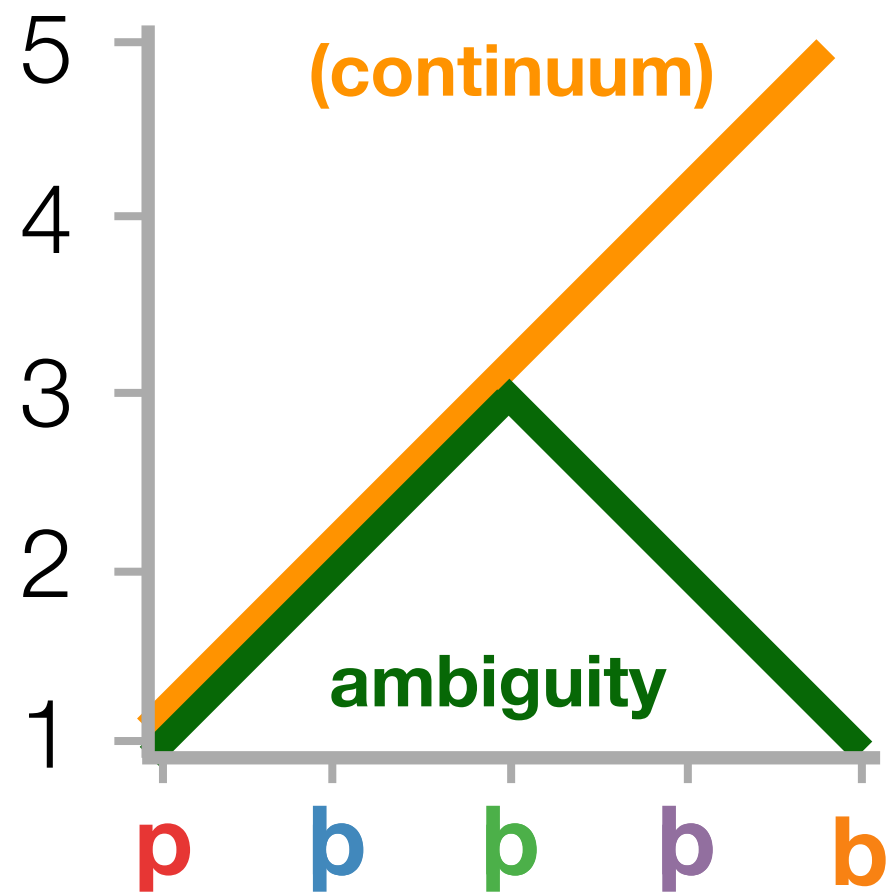


Procedure & Analysis

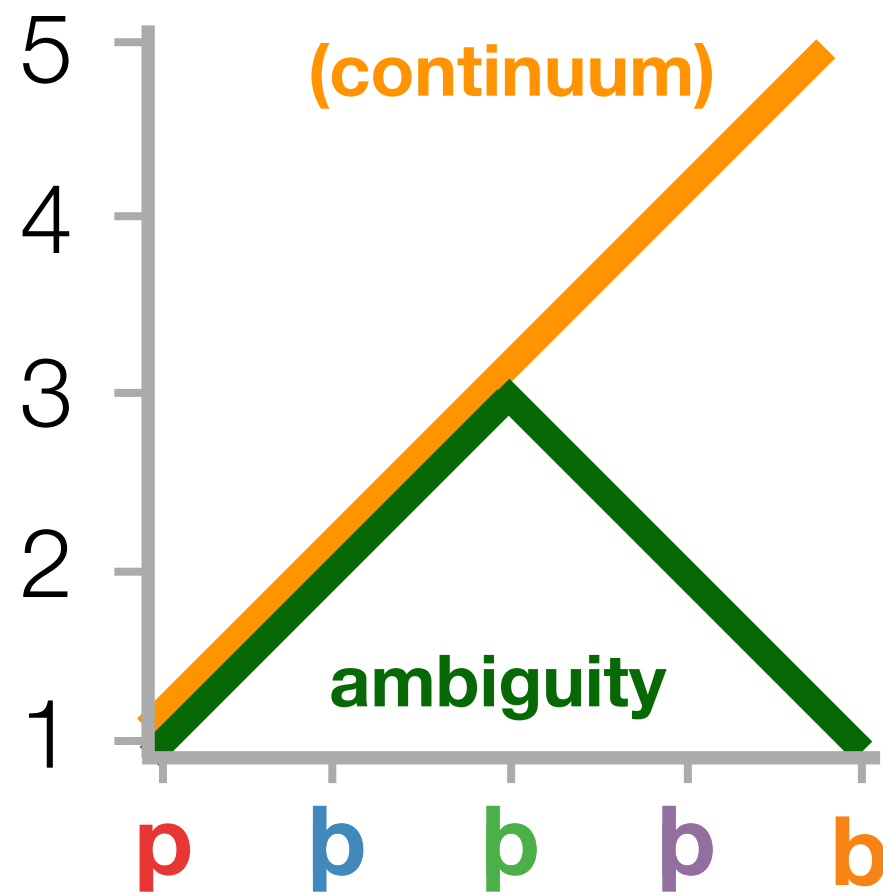


Three Experimental Variables

Three Experimental Variables

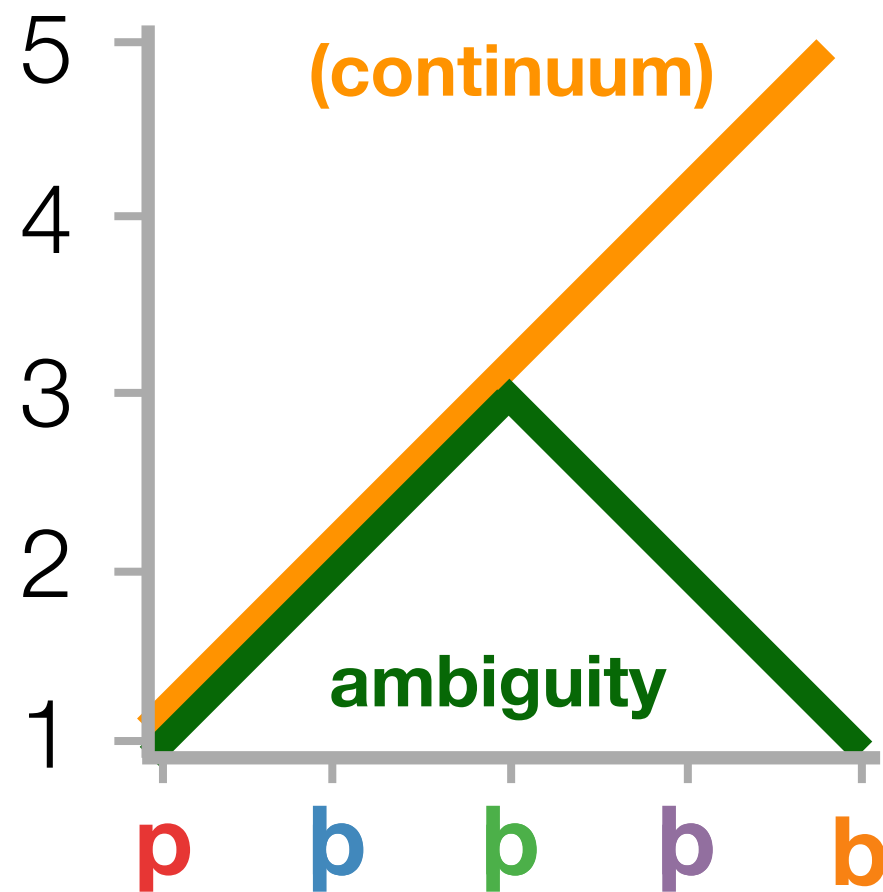


Three Experimental Variables



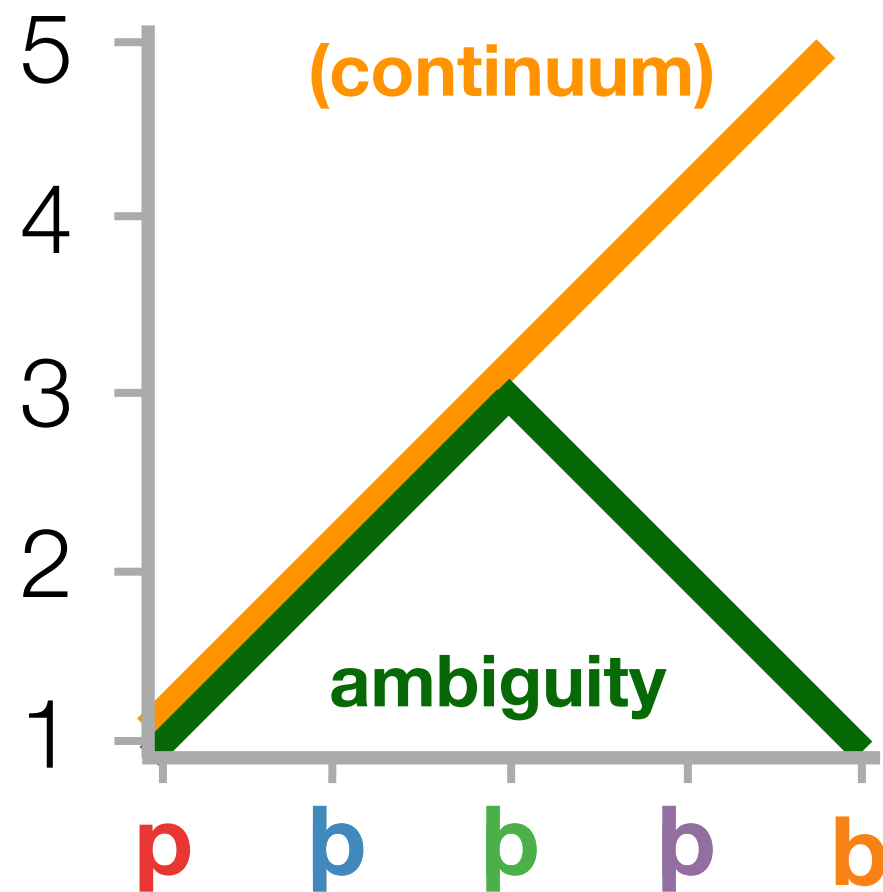
voiceless	voiced	
p p	b b	bilabial
t t	d d	labio-dental
k k	g g	velar

Three Experimental Variables



VOT		
voiceless	voiced	
p p	b b	bilabial
t t	d d	labio-dental
k k	g g	velar

Three Experimental Variables



VOT		
voiceless	voiced	
p p	b b	bilabial
t t	d d	labio-dental
k k	g g	velar

POA

Today's Questions

How does the auditory cortex
respond to phonological ambiguity?

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

Ambiguity at Onset

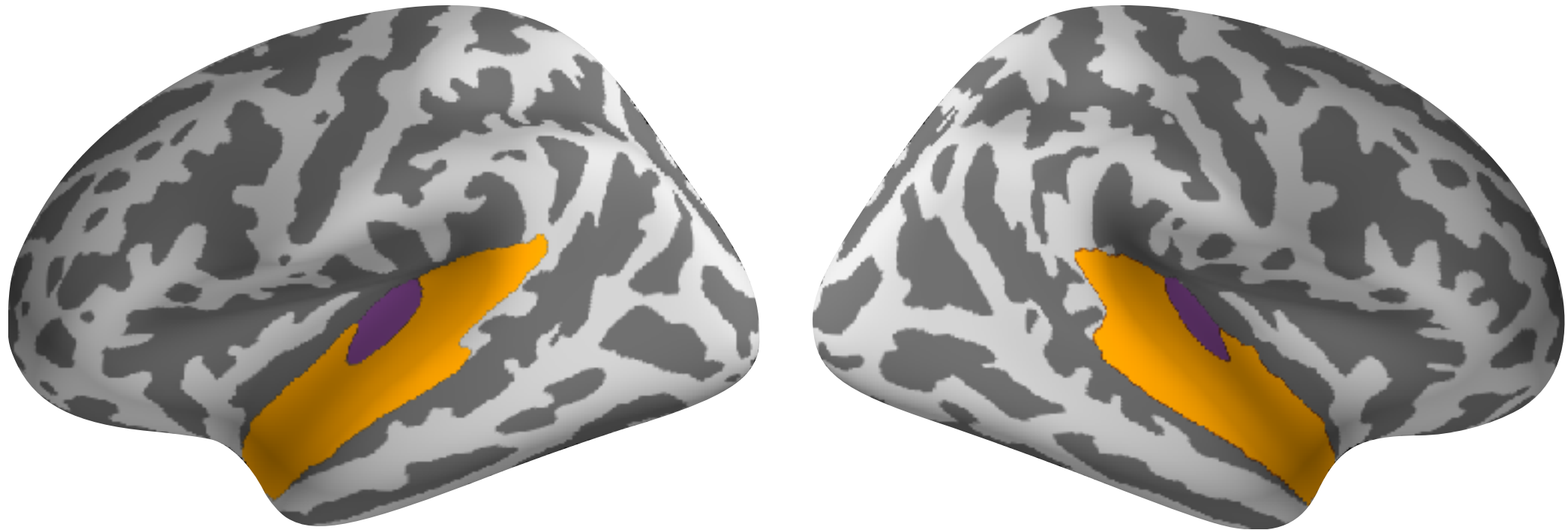
Ambiguity at Onset



Ambiguity at Onset



Ambiguity at Onset



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



Ambiguity at Onset



Ambiguity at Onset

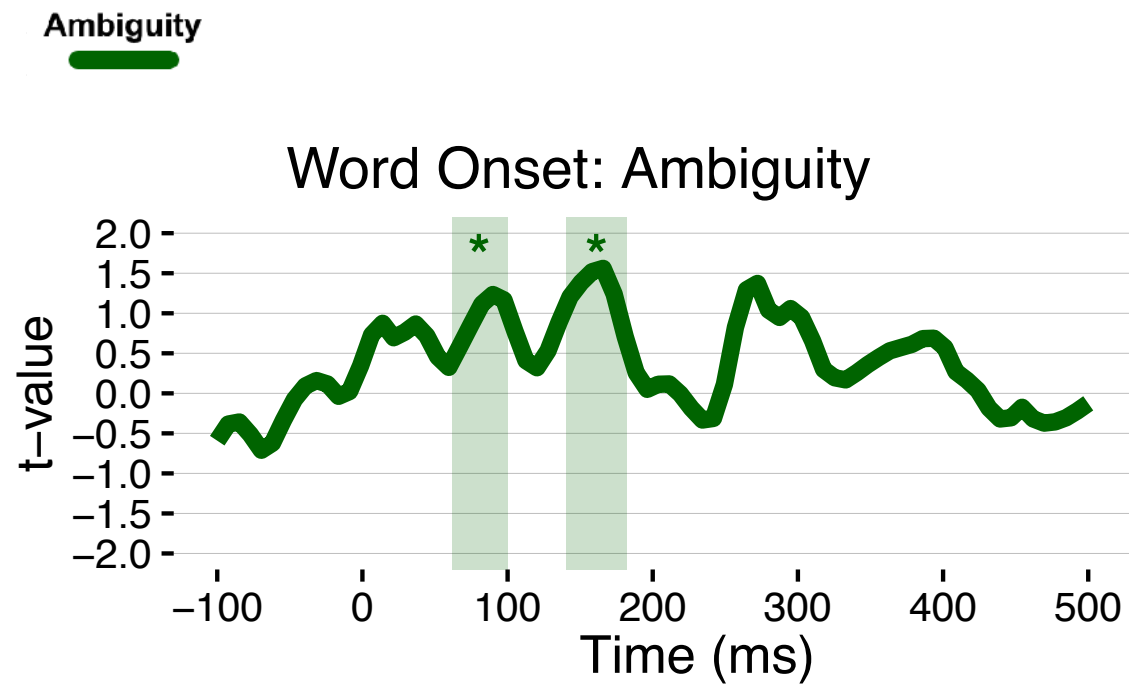


Ambiguity at Onset

Ambiguity

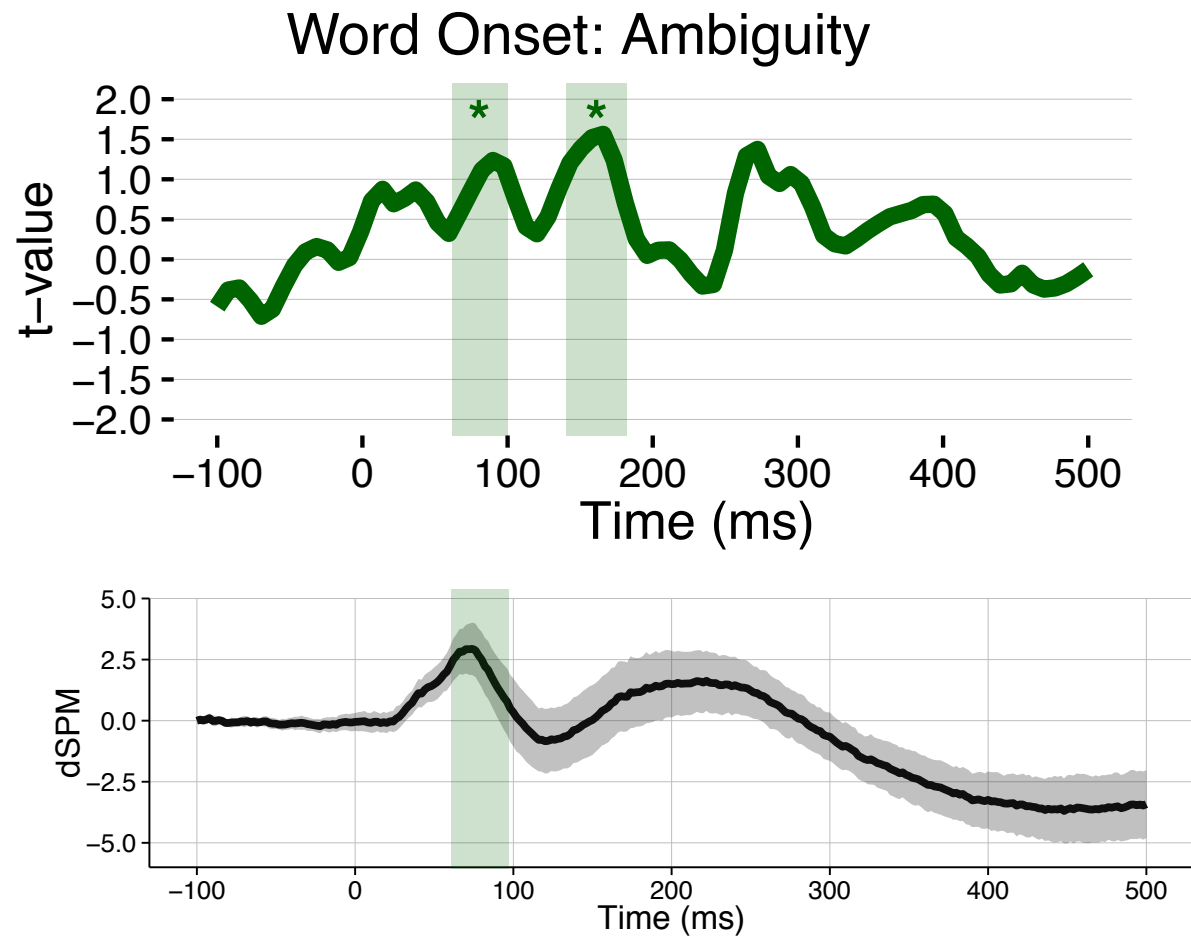


Ambiguity at Onset

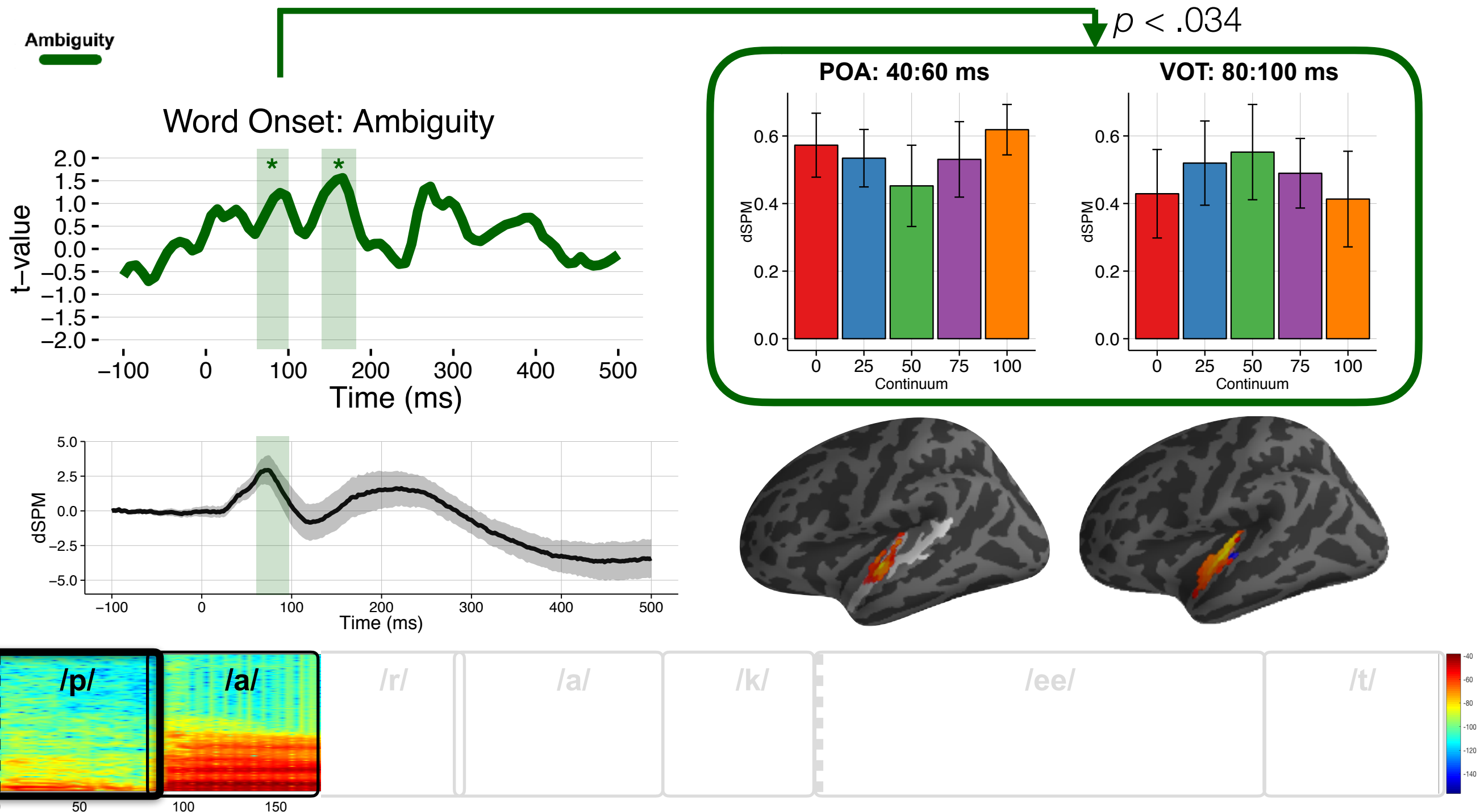


Ambiguity at Onset

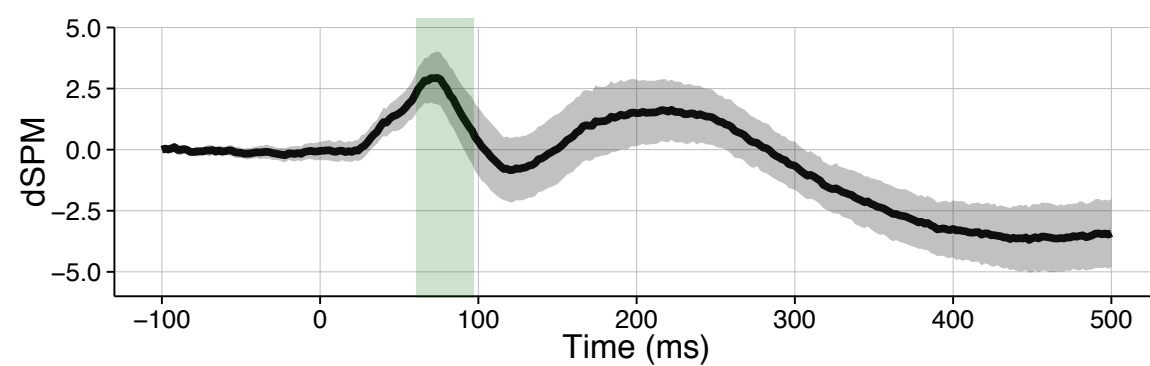
Ambiguity



Ambiguity at Onset



Ambiguity at Onset

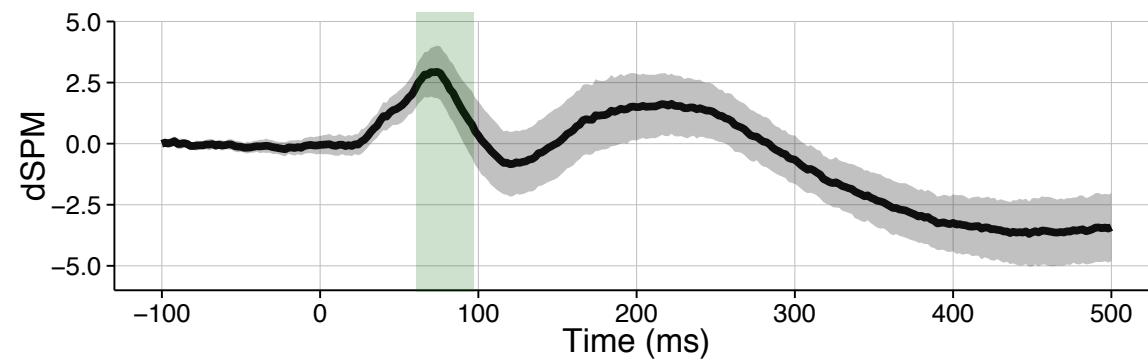


Ambiguity at Onset

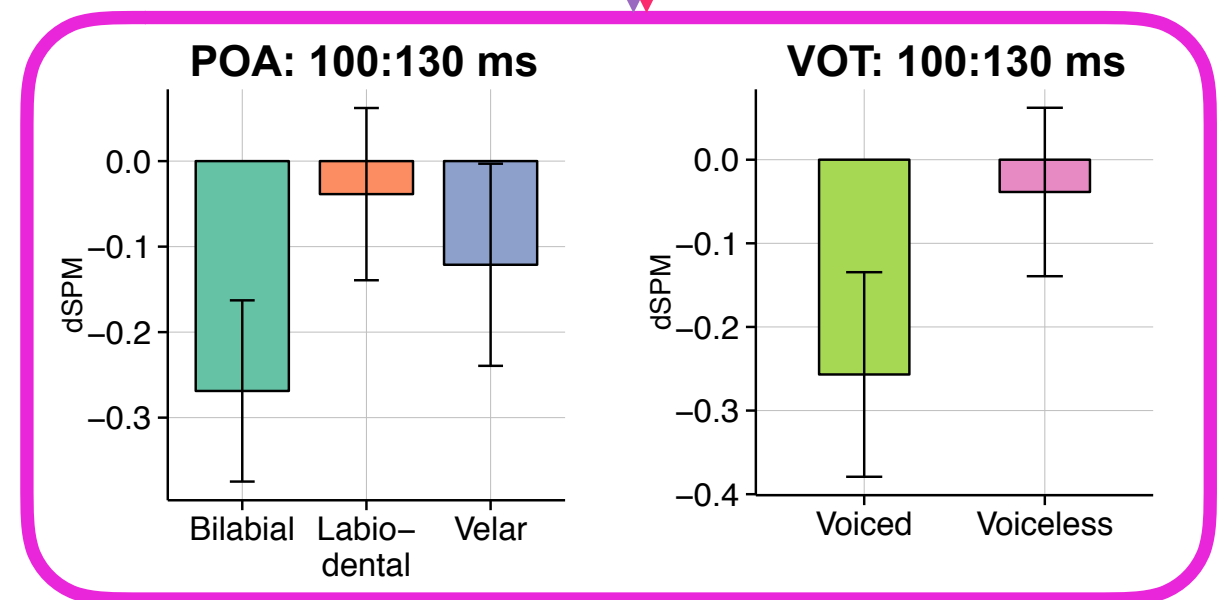
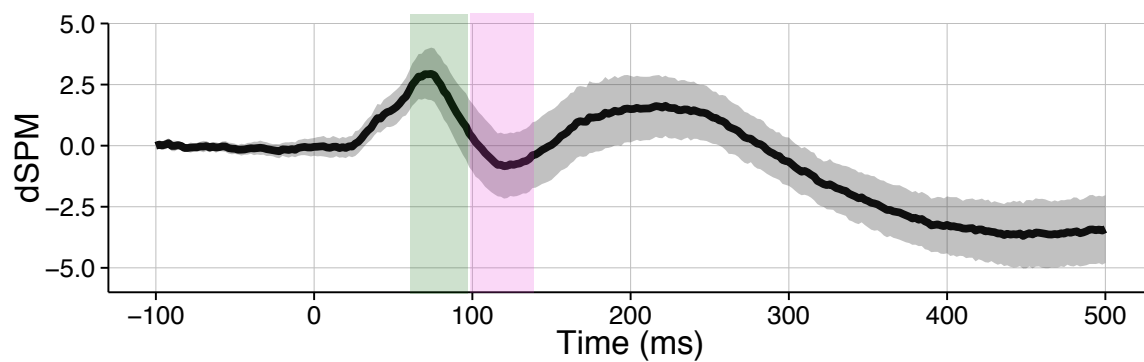
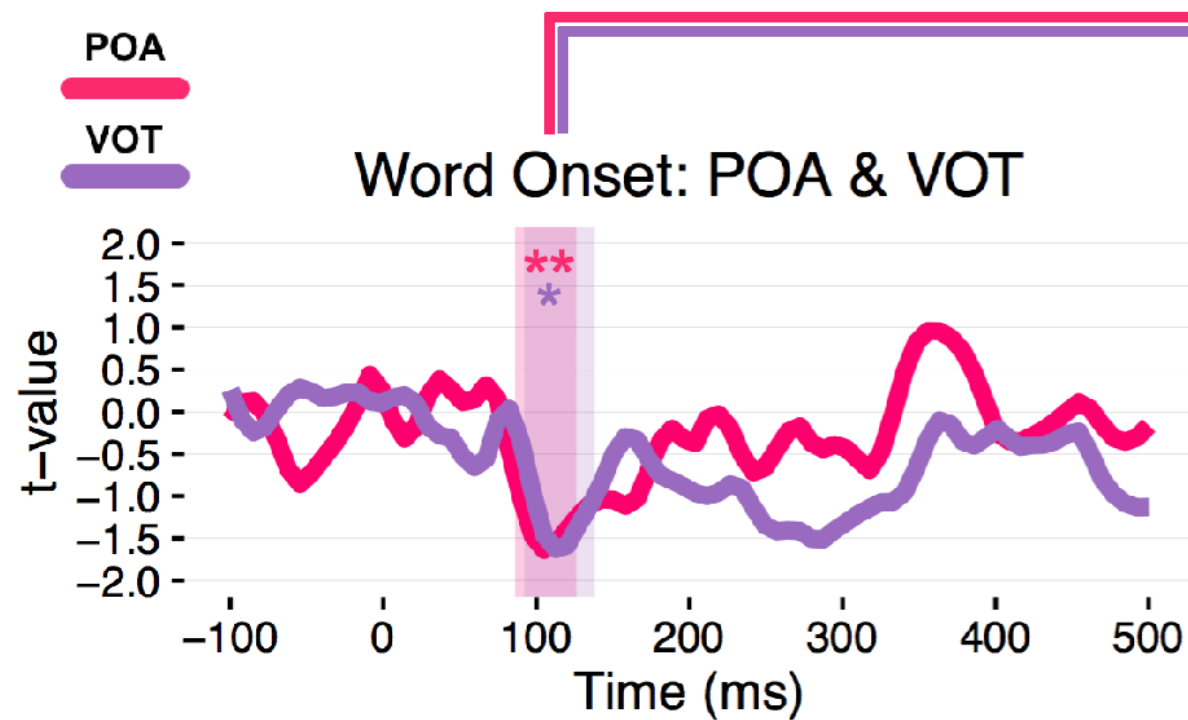
POA



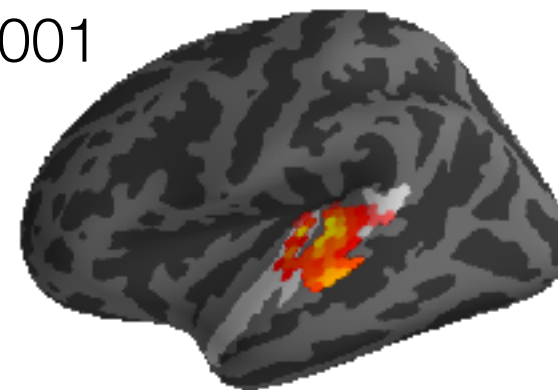
VOT



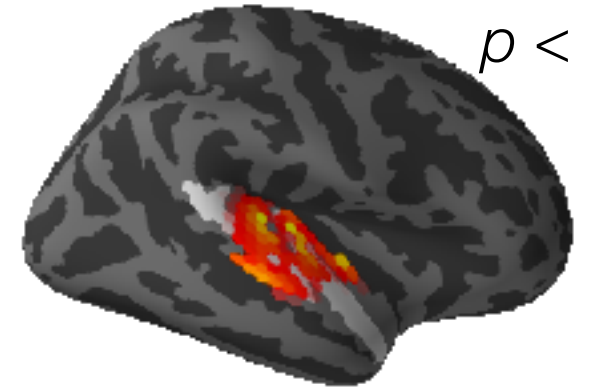
Ambiguity at Onset



$p < .001$



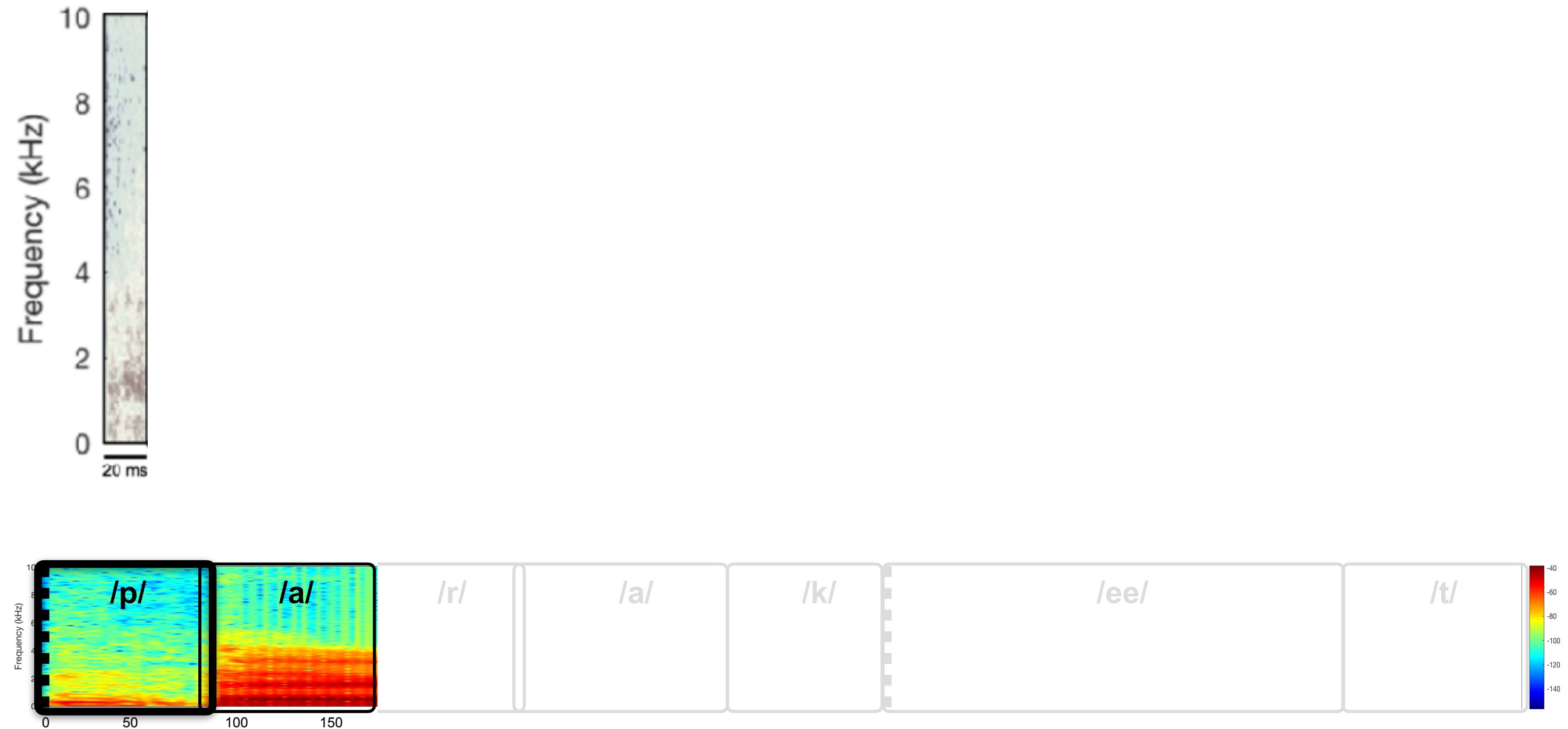
$p < .001$



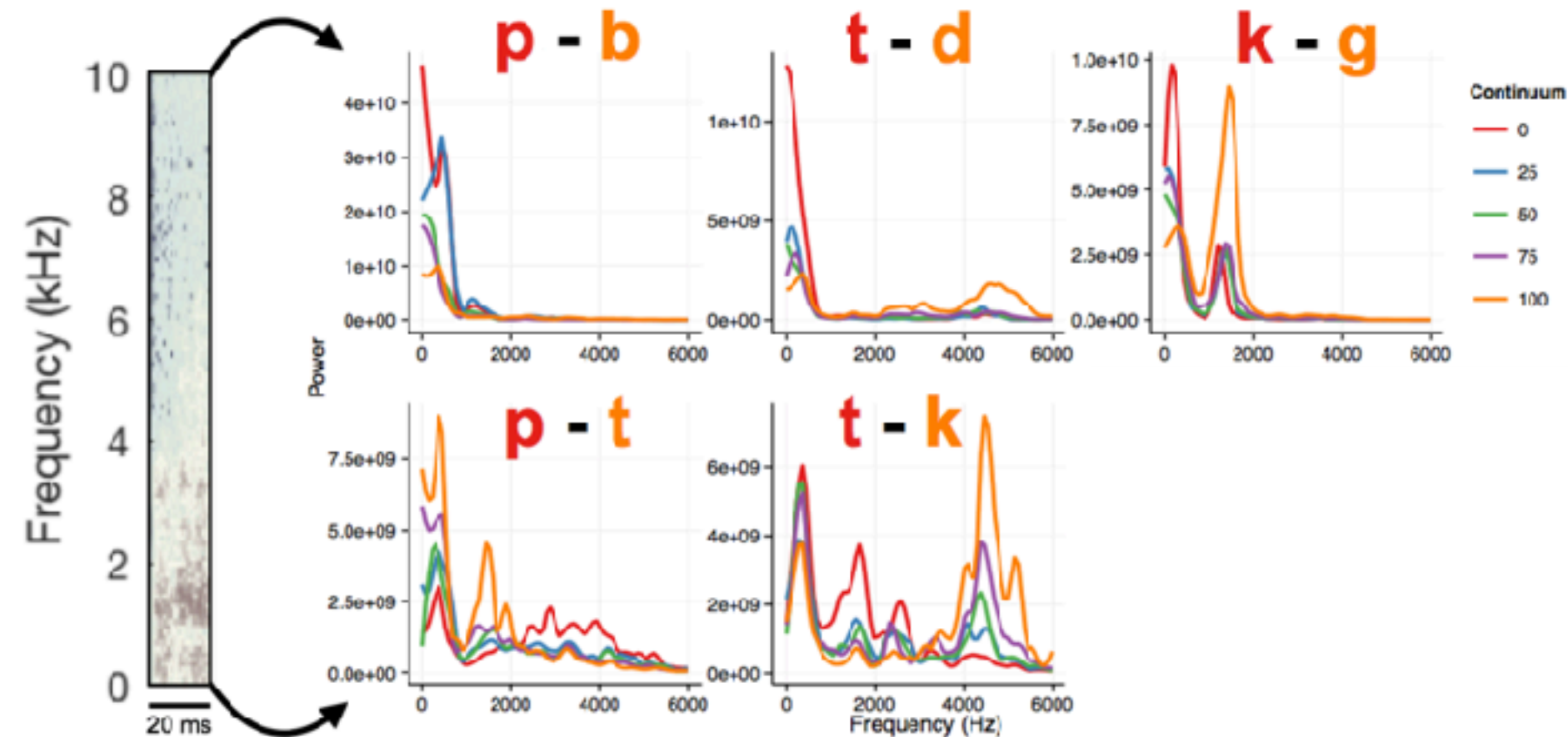
Ambiguity at Onset



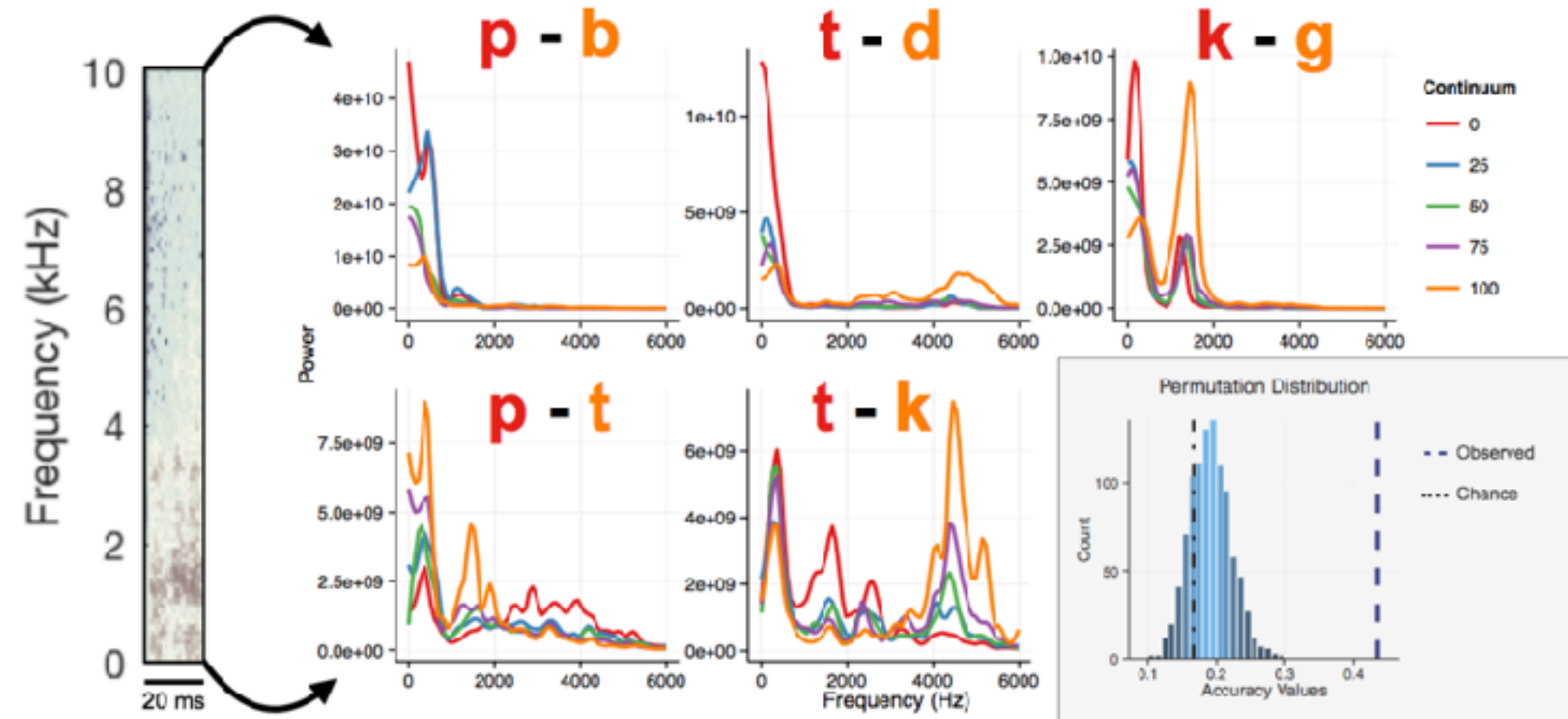
Ambiguity at Onset



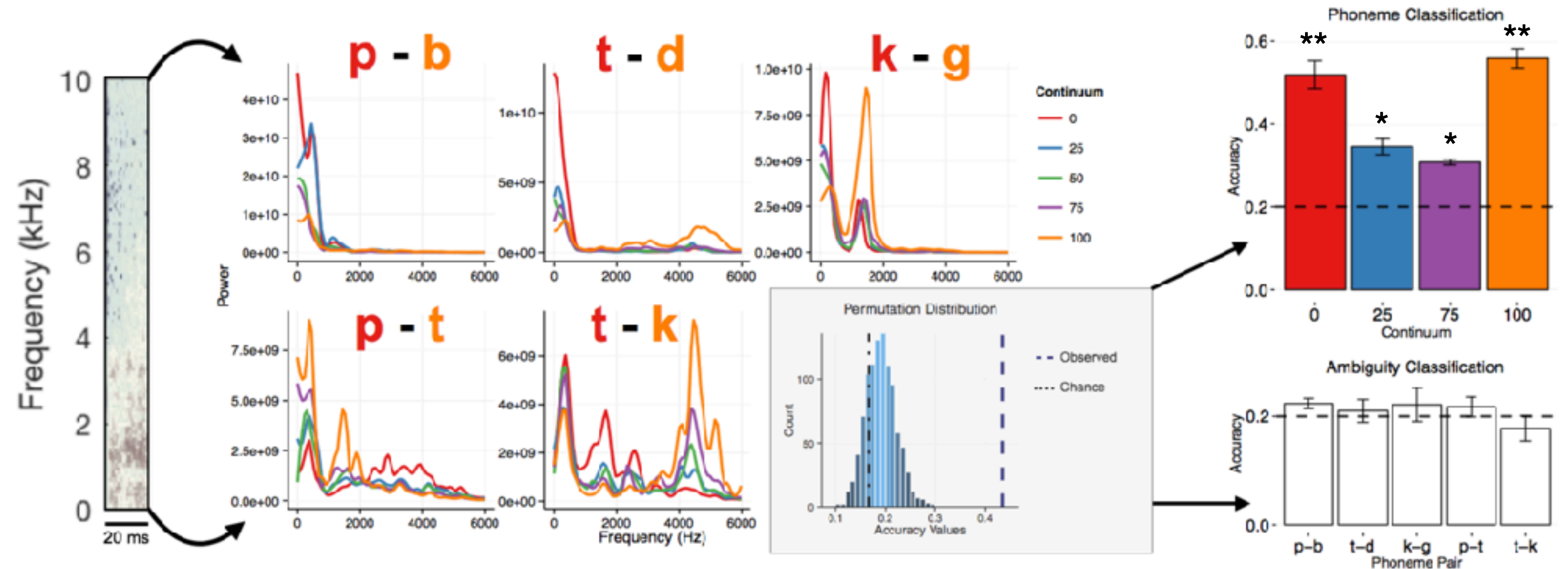
Ambiguity at Onset



Ambiguity at Onset



Ambiguity at Onset



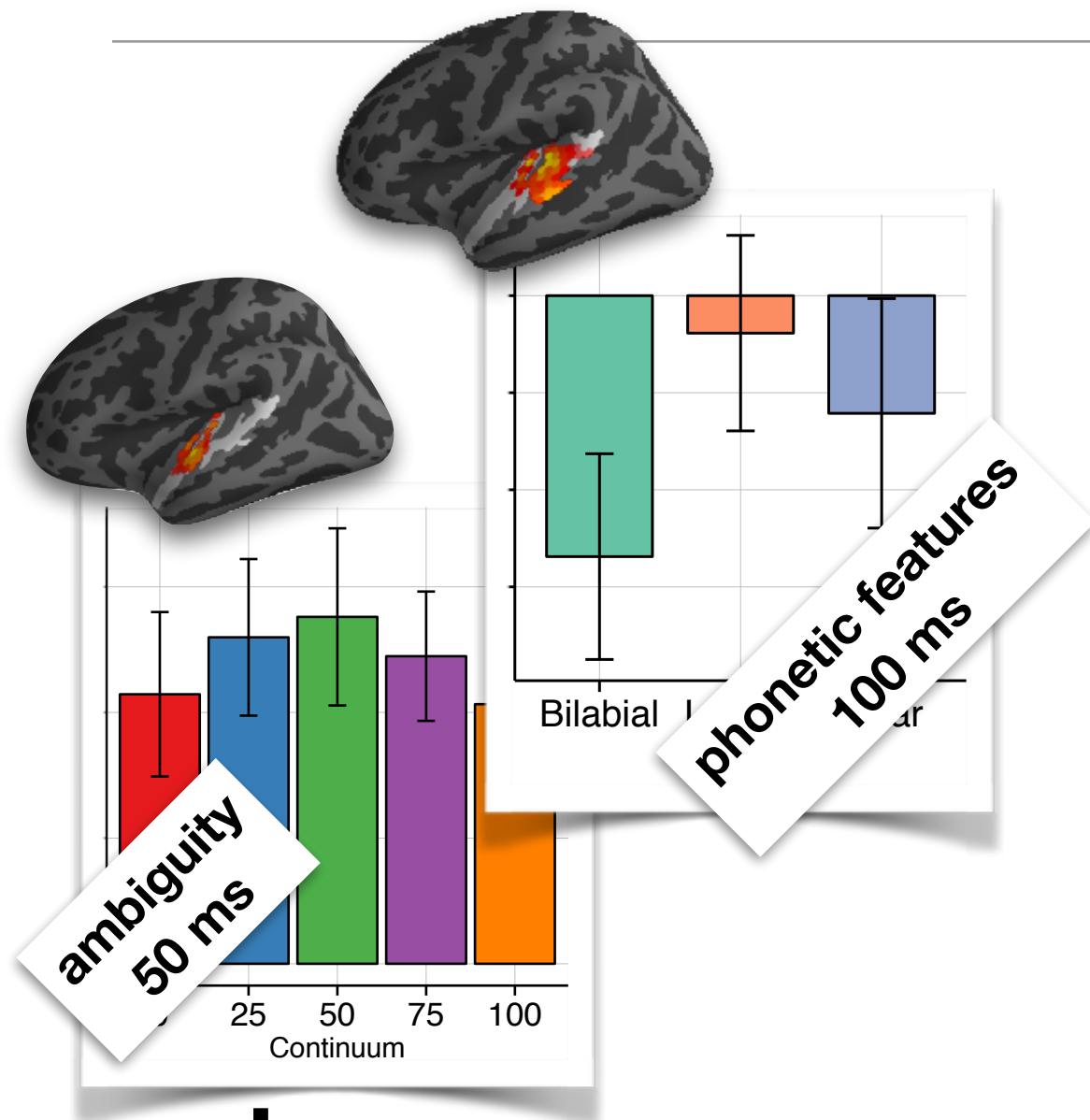
Interim Conclusion

?



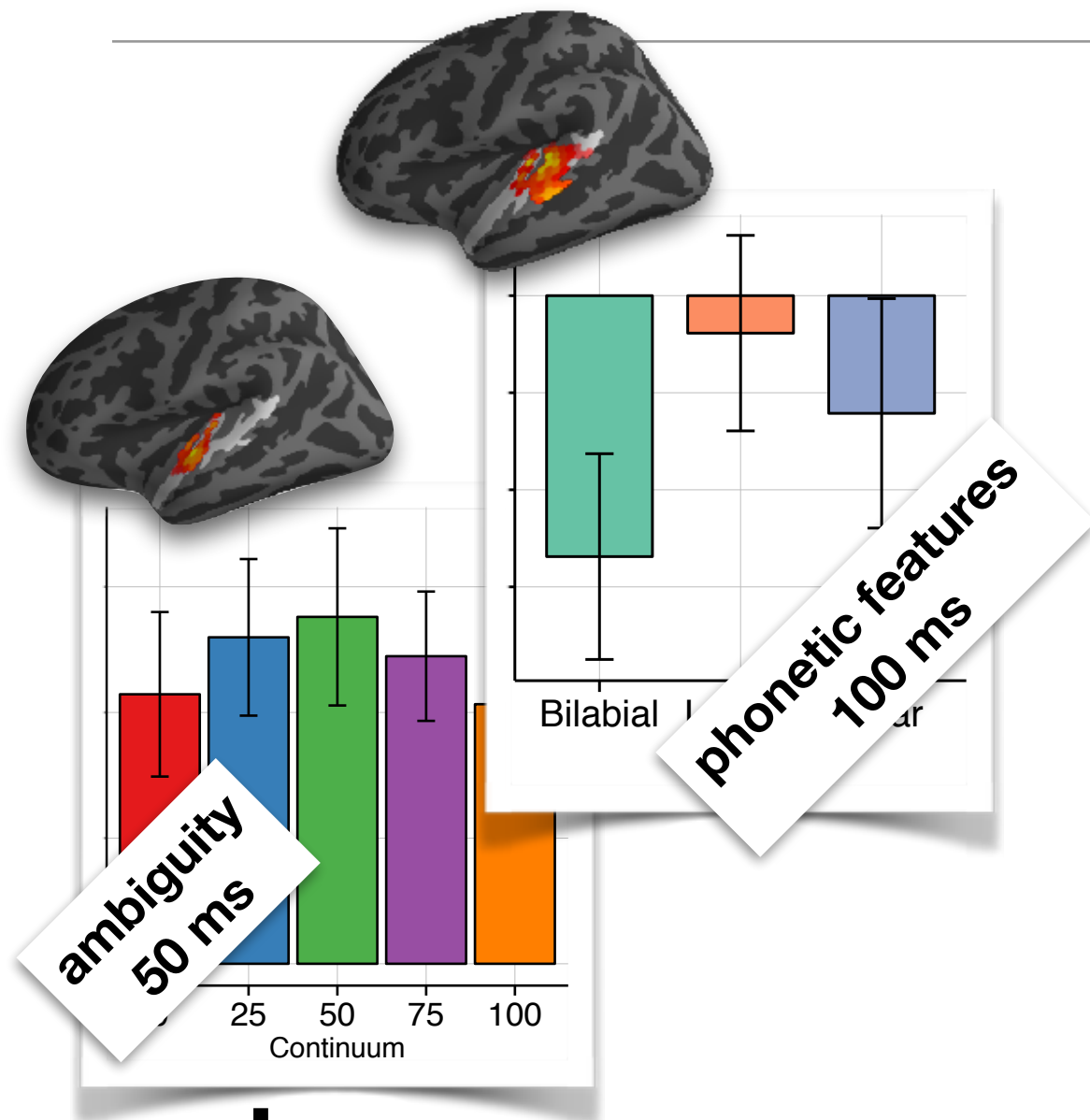
p a r a k e e t

Interim Conclusion



p a r a k e e t

Interim Conclusion



?

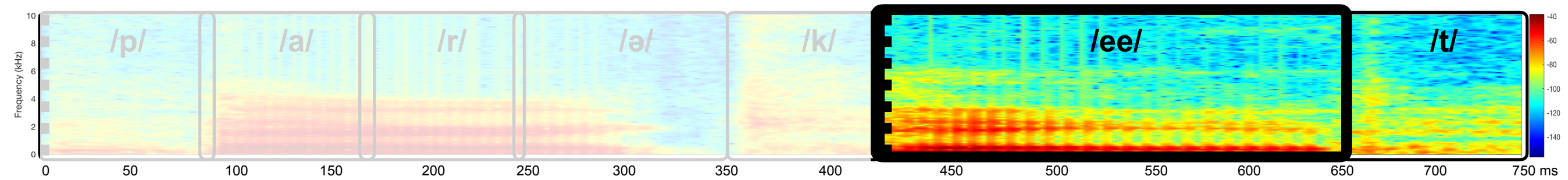


p a r a k e e t

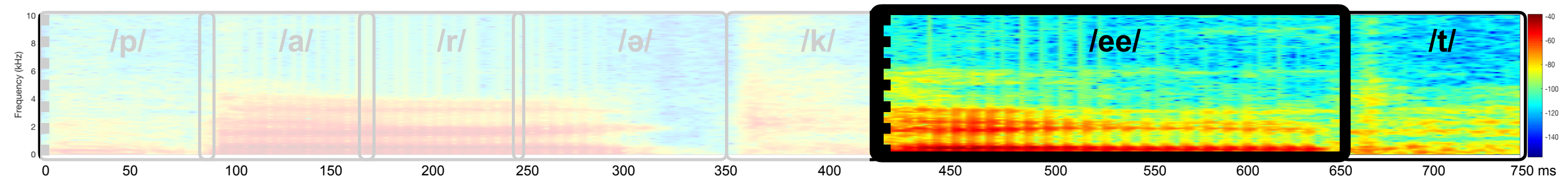
Today's Questions

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

Ambiguity at POD

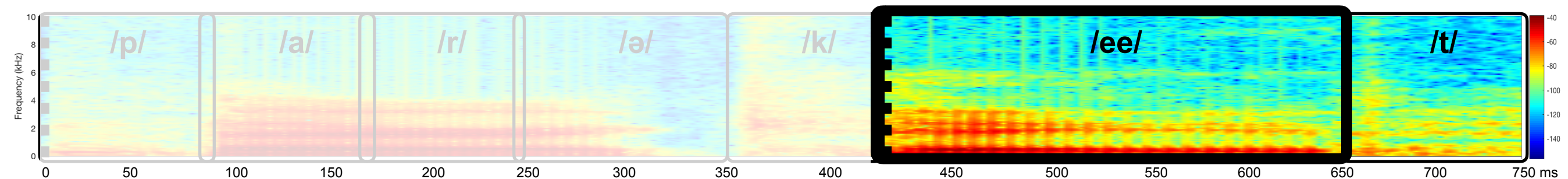


Ambiguity at POD



Ambiguity at POD

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

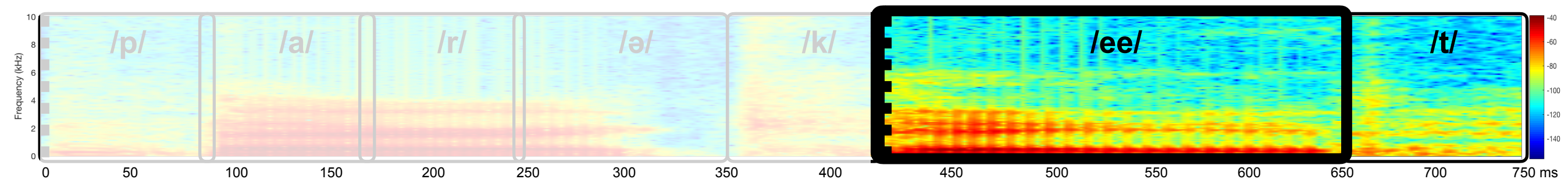


Ambiguity at POD

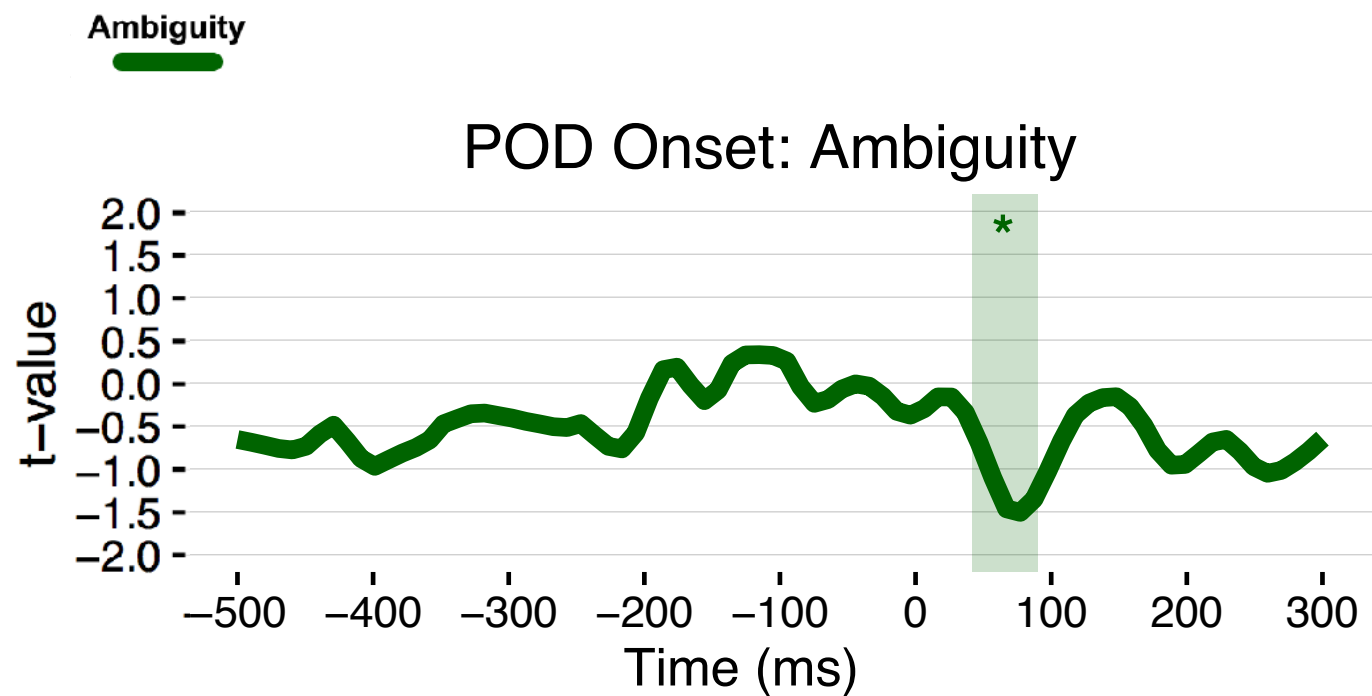
Ambiguity



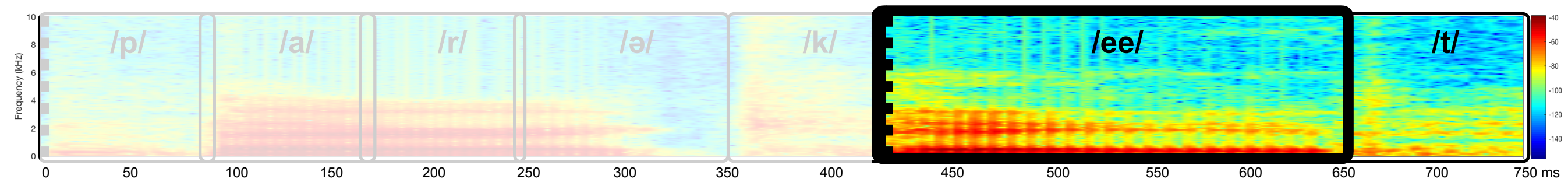
- Time-window: 0-200 ms after POD onset
- Region: **Heschl's gyrus** & **superior temporal gyrus** bilaterally



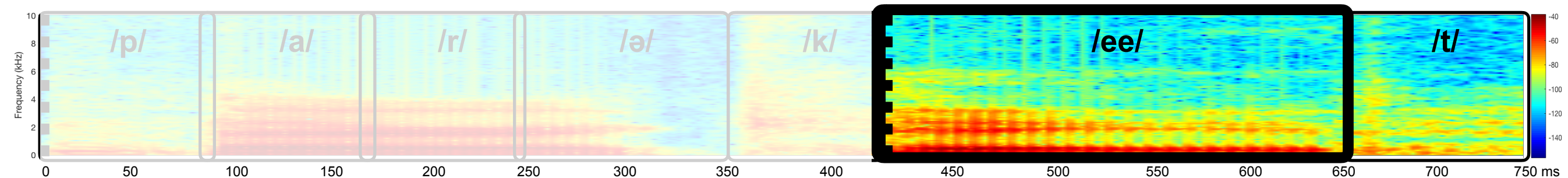
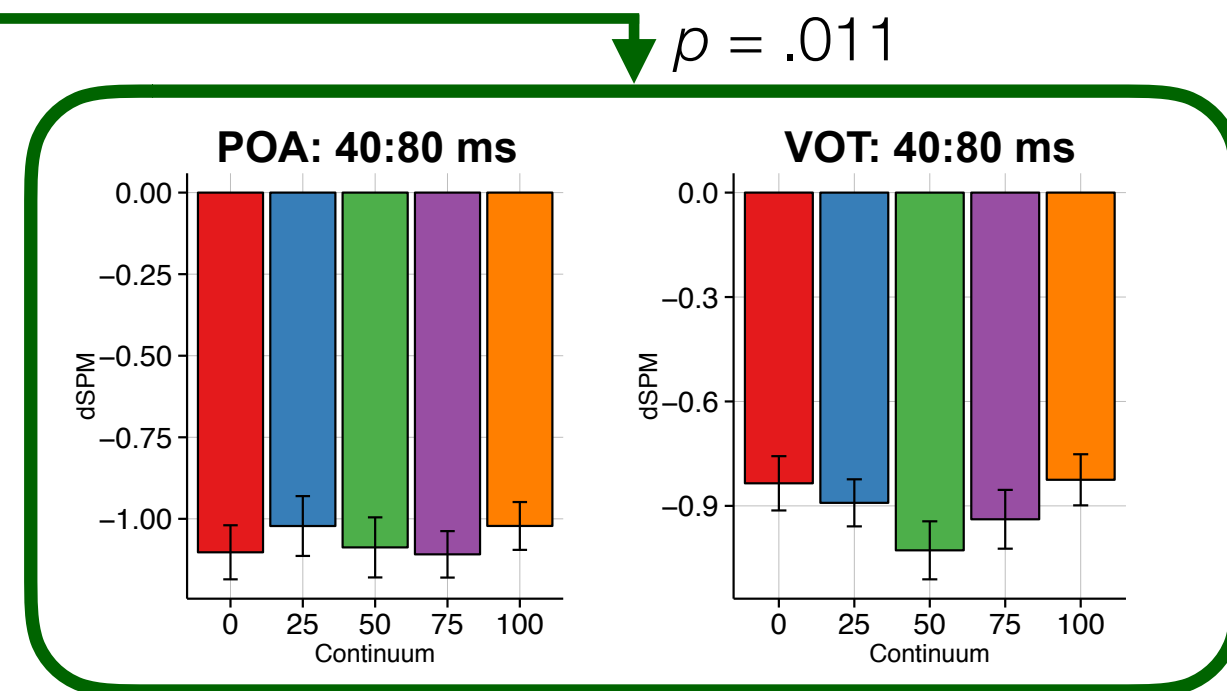
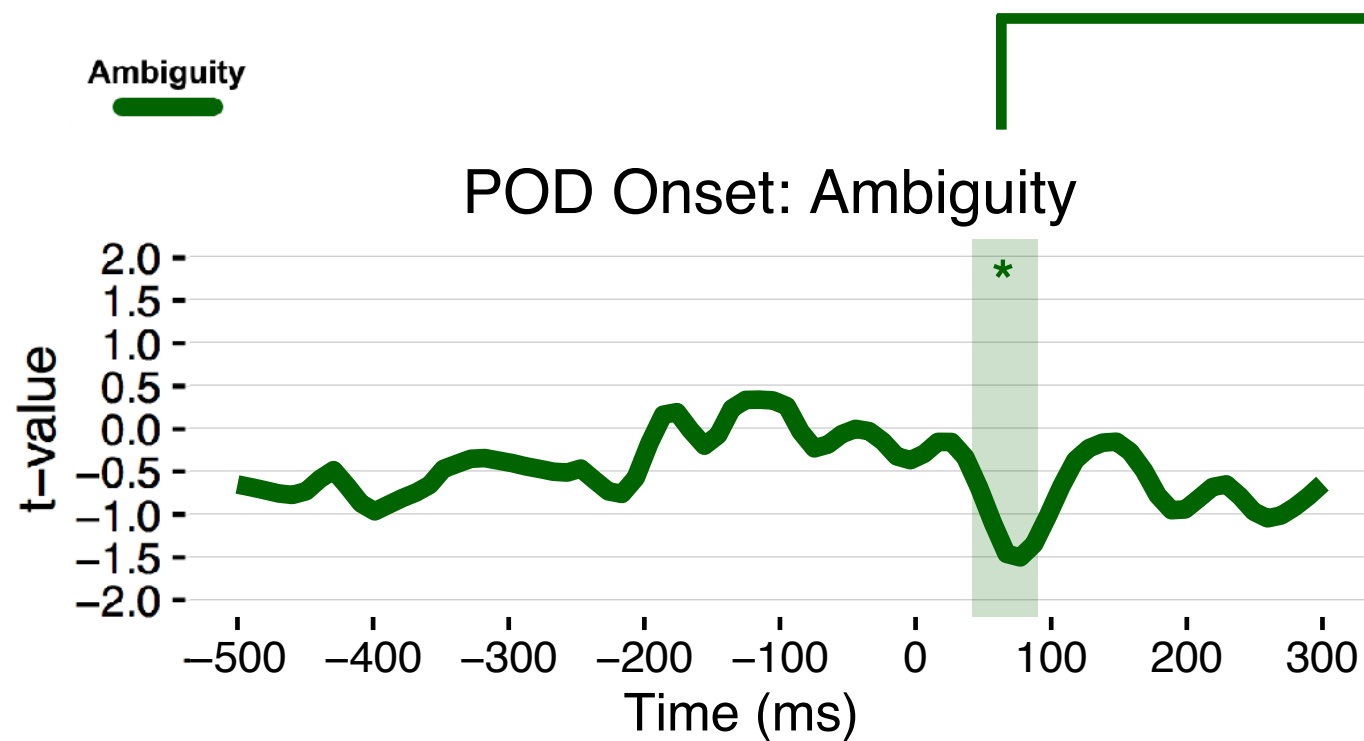
Ambiguity at POD



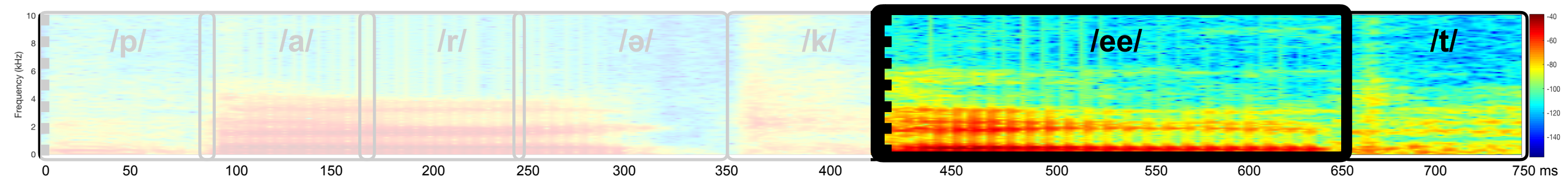
- Time-window: 0-200 ms after POD onset
- Region: **Heschl's gyrus** & **superior temporal gyrus** bilaterally



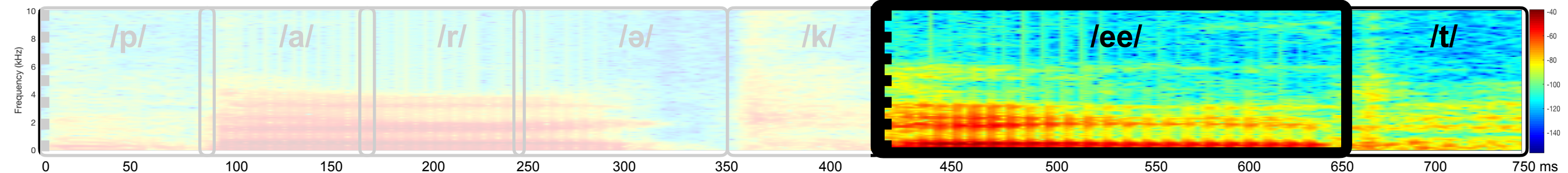
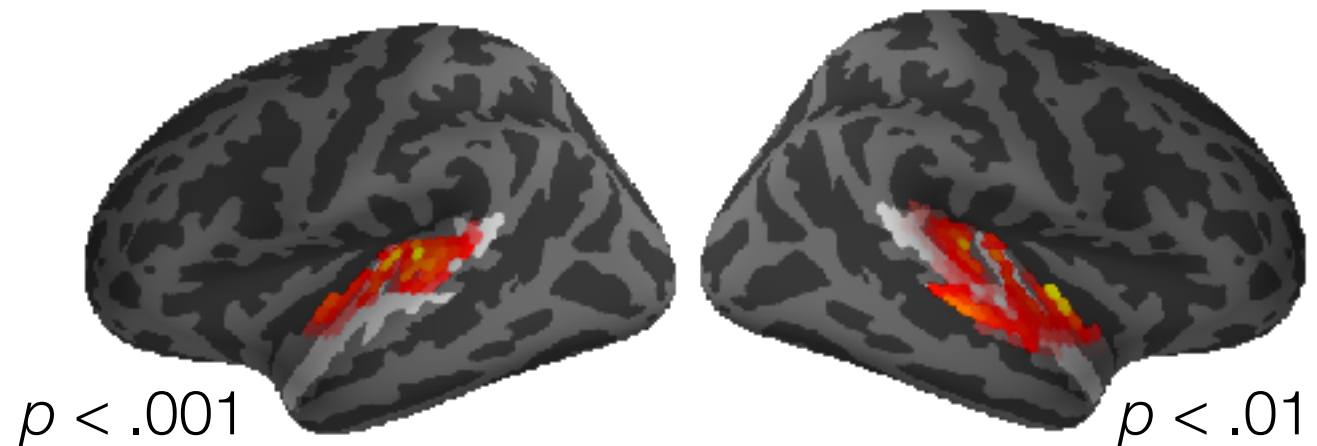
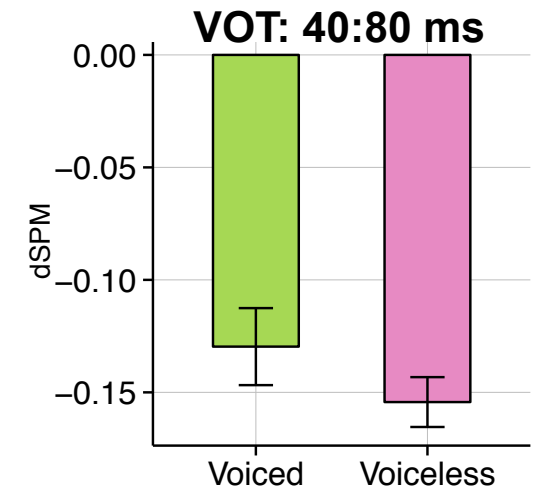
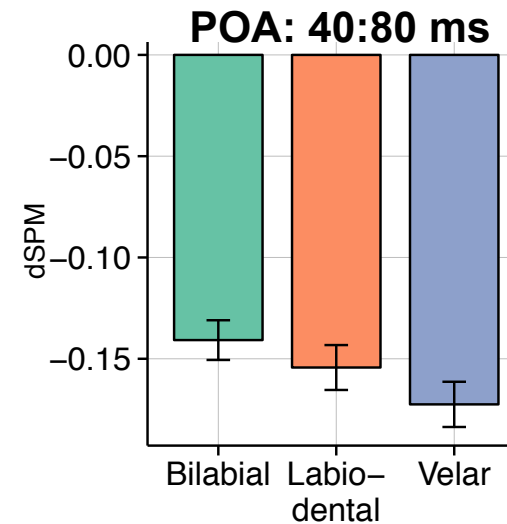
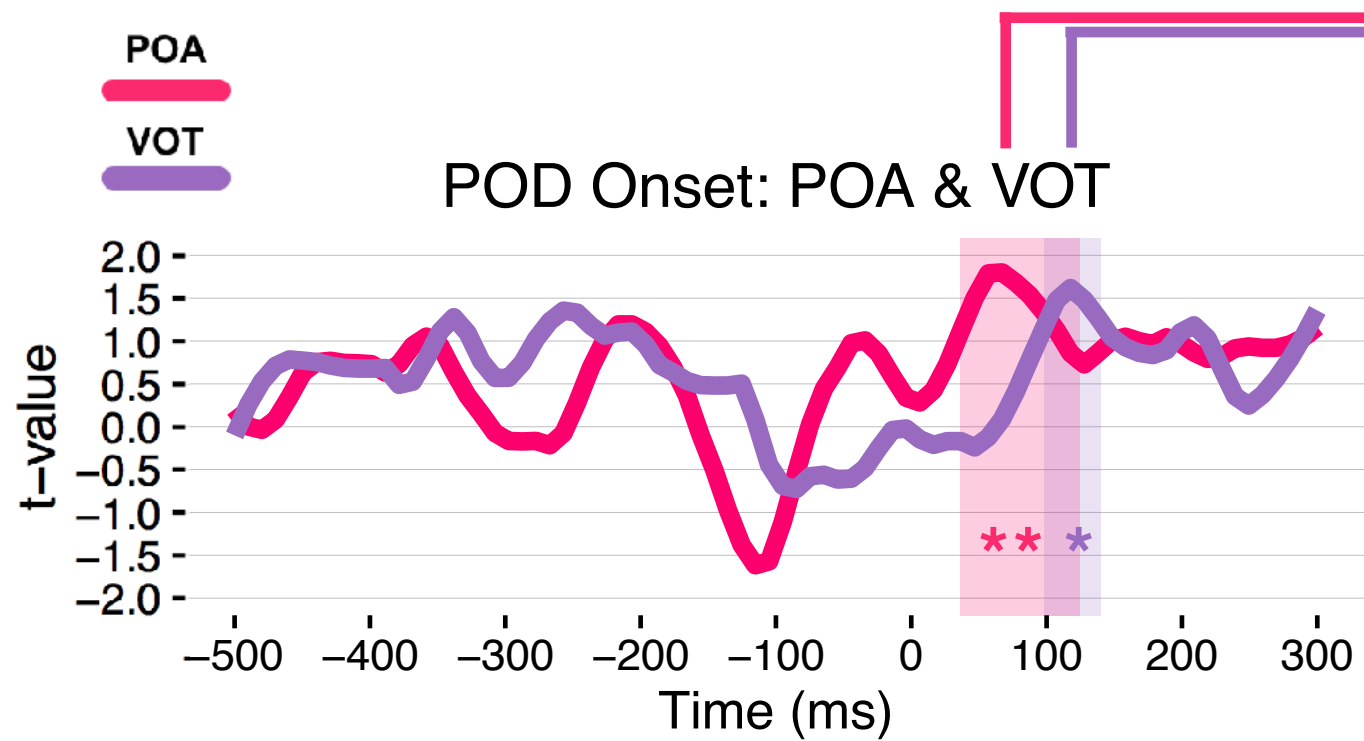
Ambiguity at POD



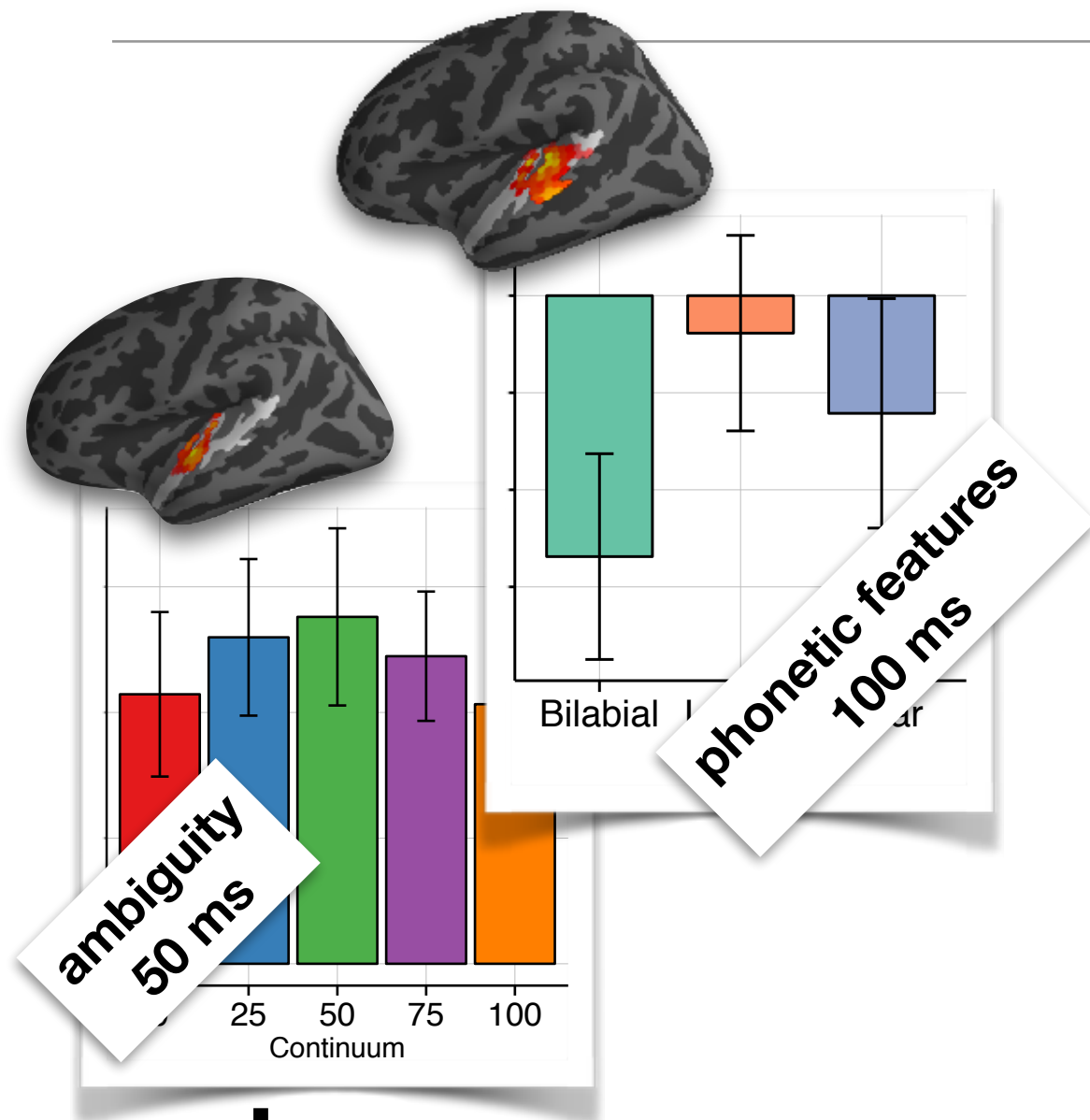
Ambiguity at POD



Ambiguity at POD



Interim Conclusion

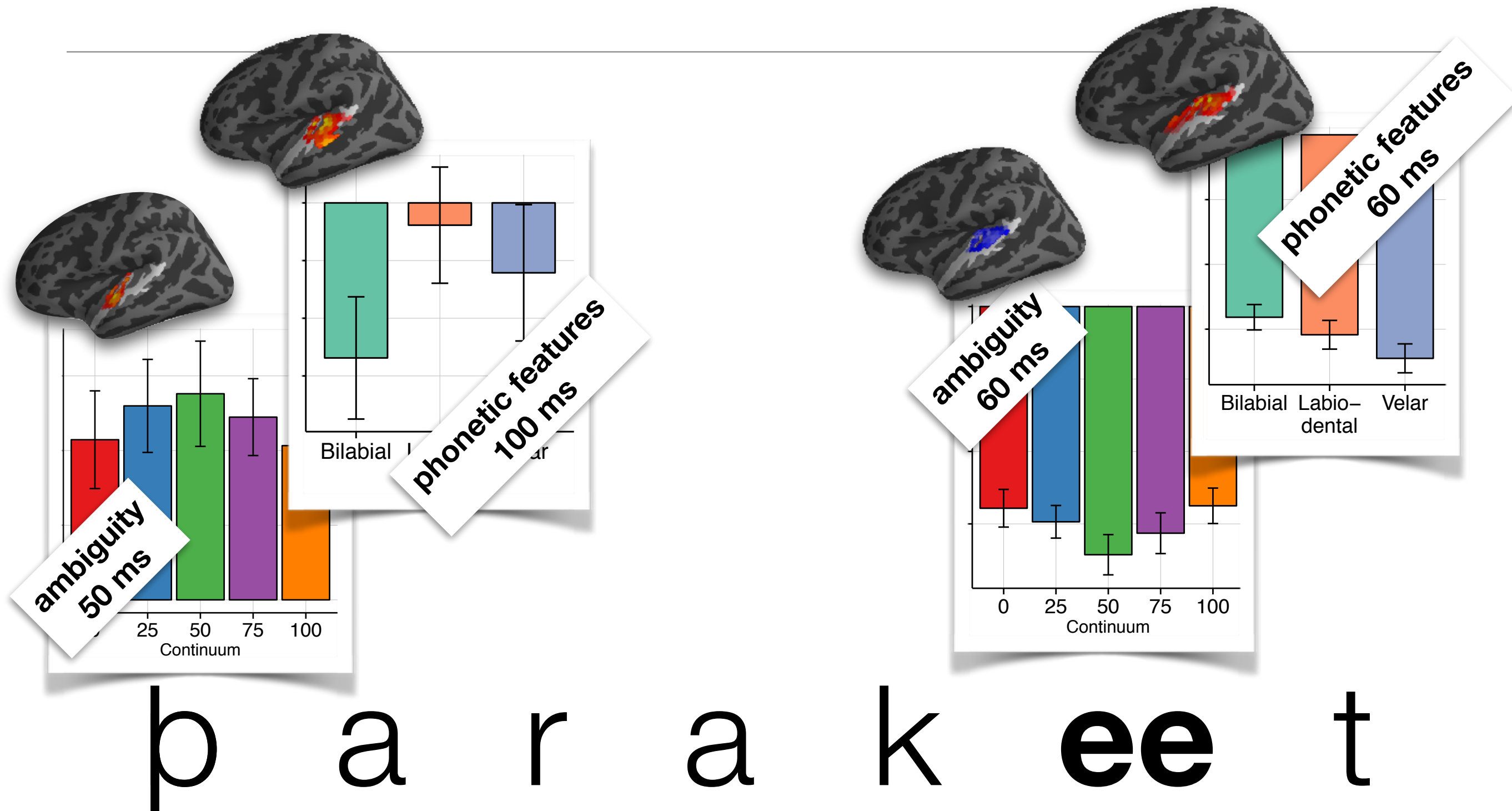


?

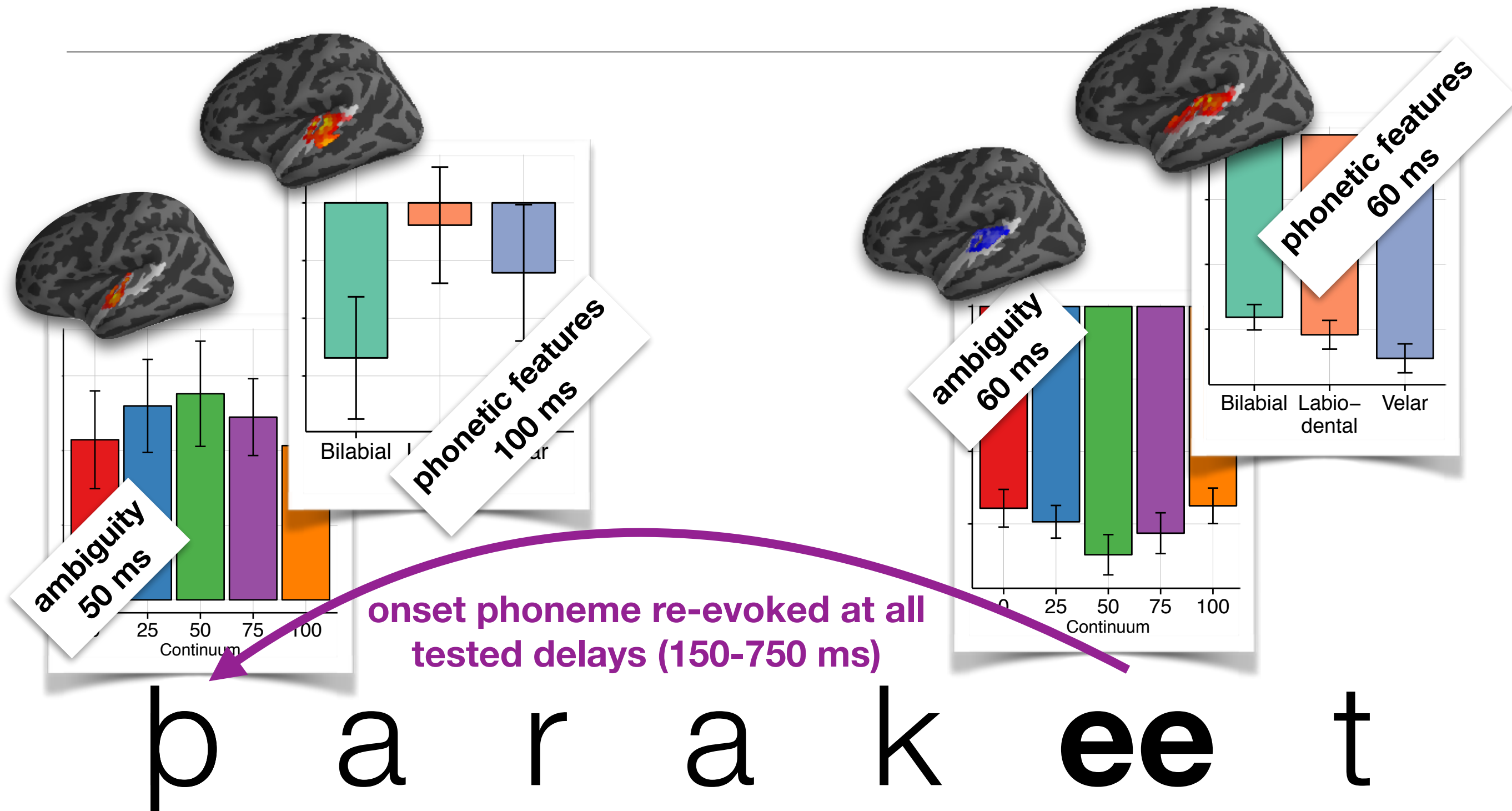


p a r a k e e t

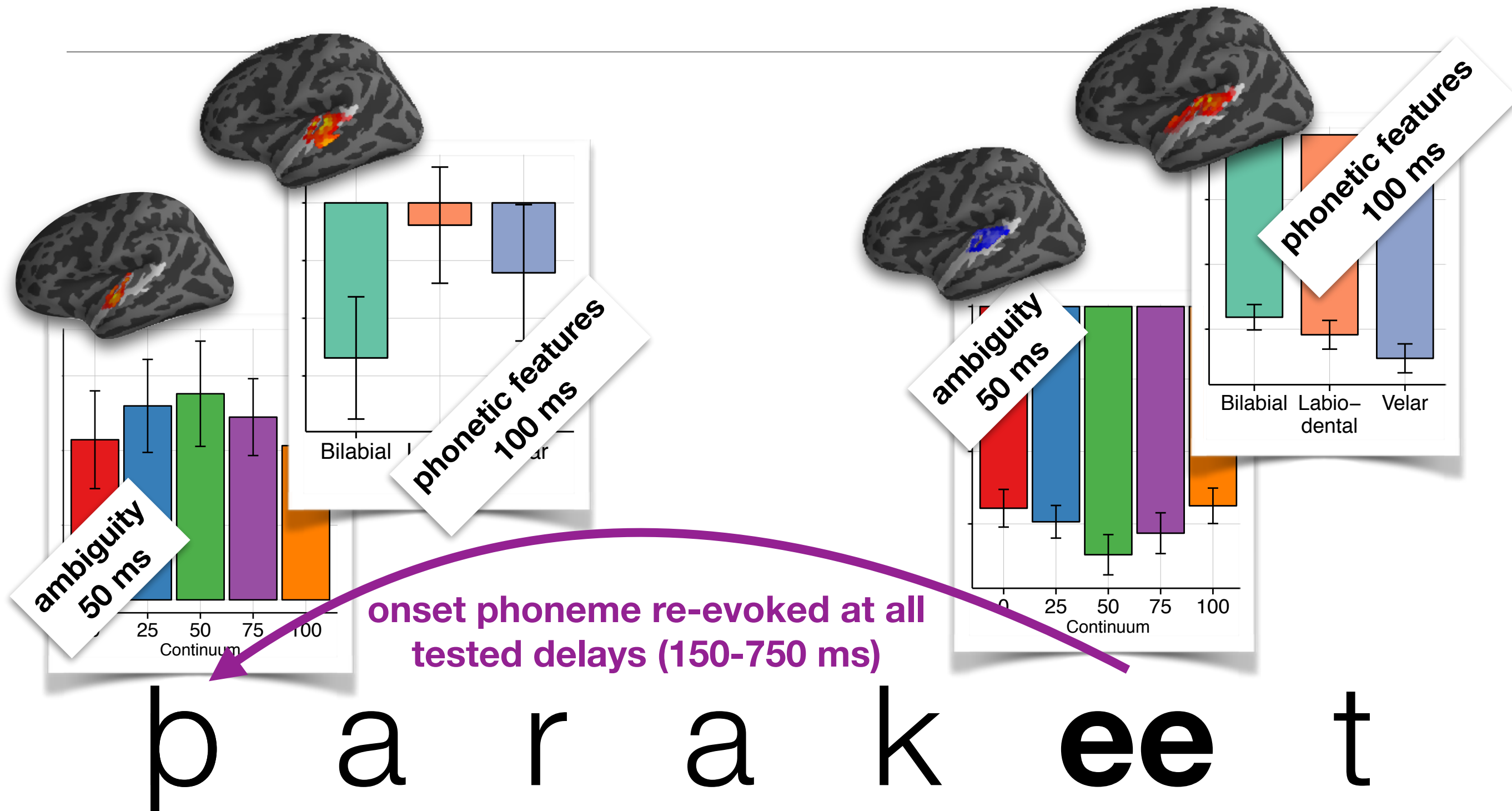
Interim Conclusion



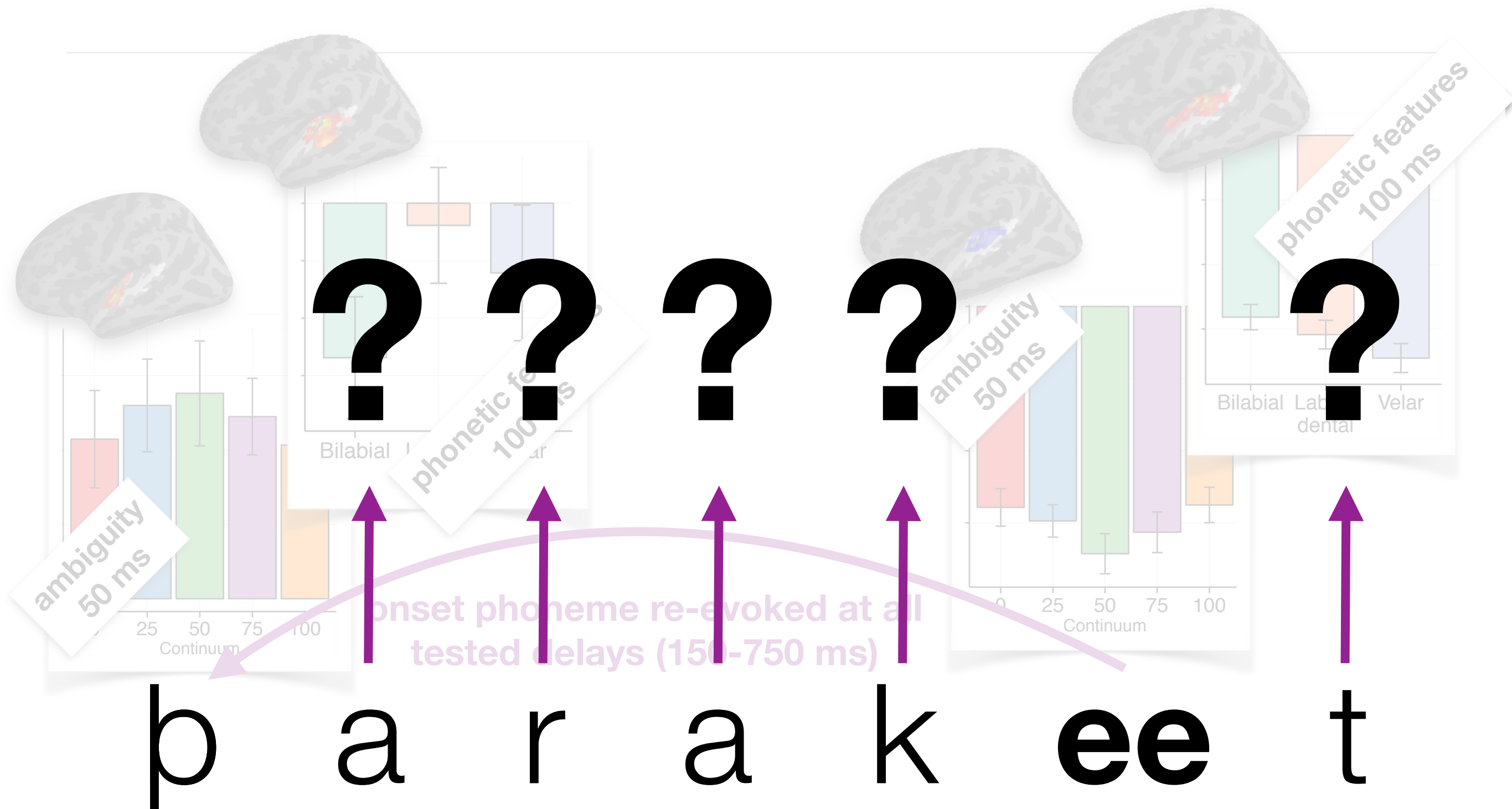
Interim Conclusion



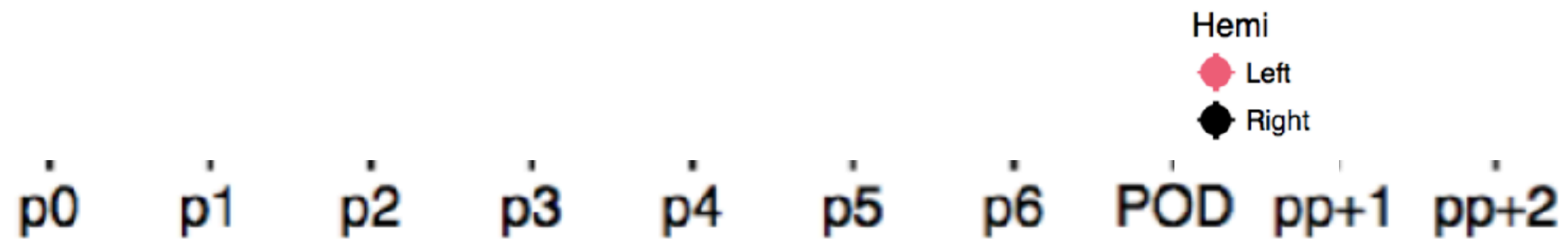
Interim Conclusion



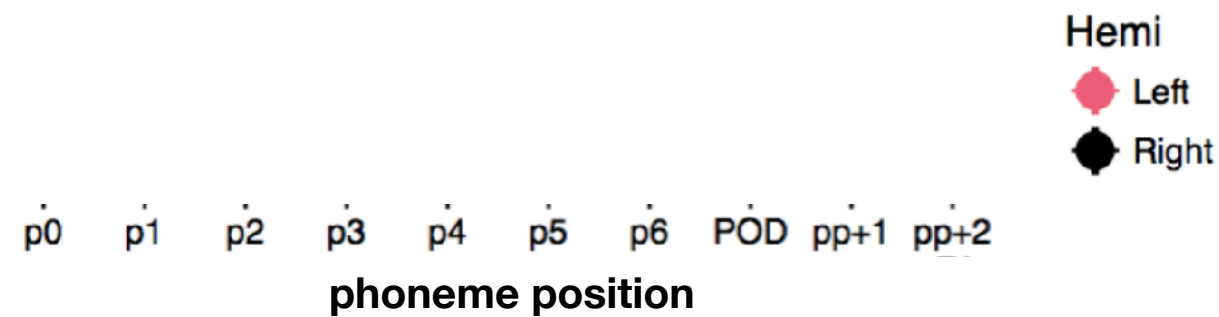
Interim Conclusion



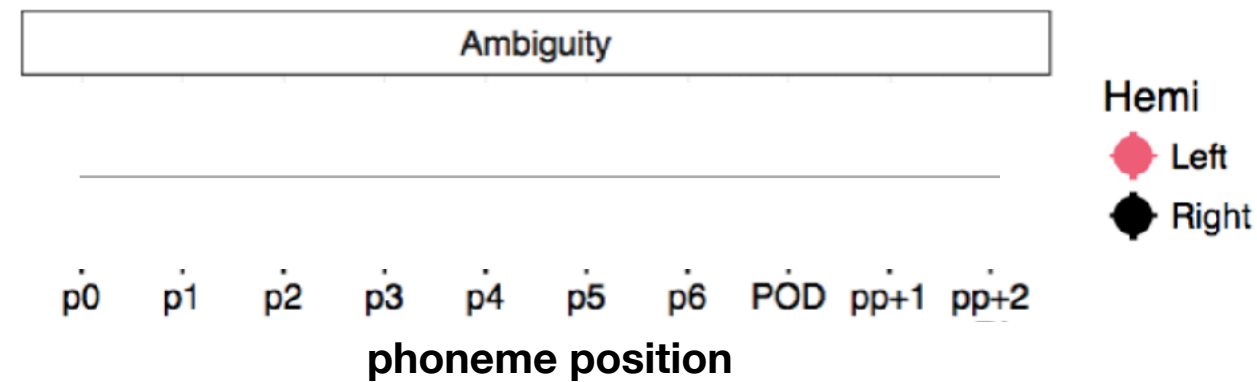
Reactivation in Intermediate Positions



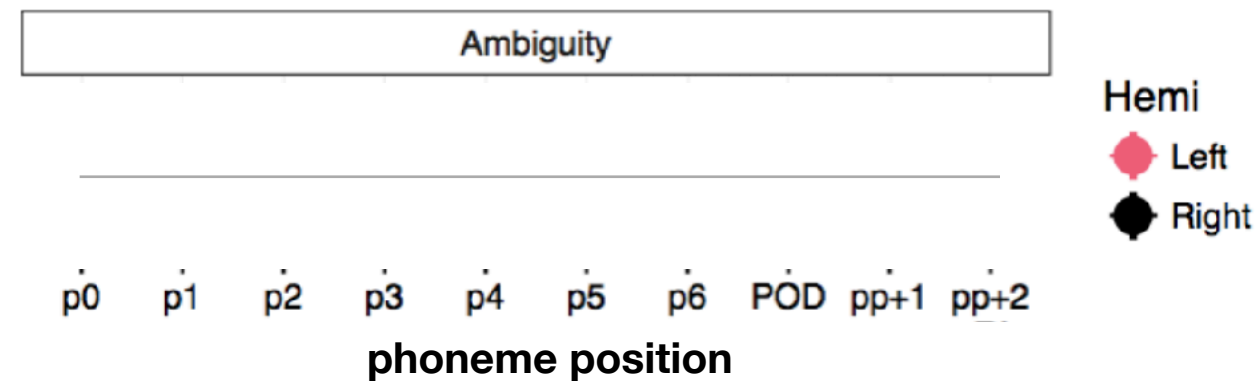
Reactivation in Intermediate Positions



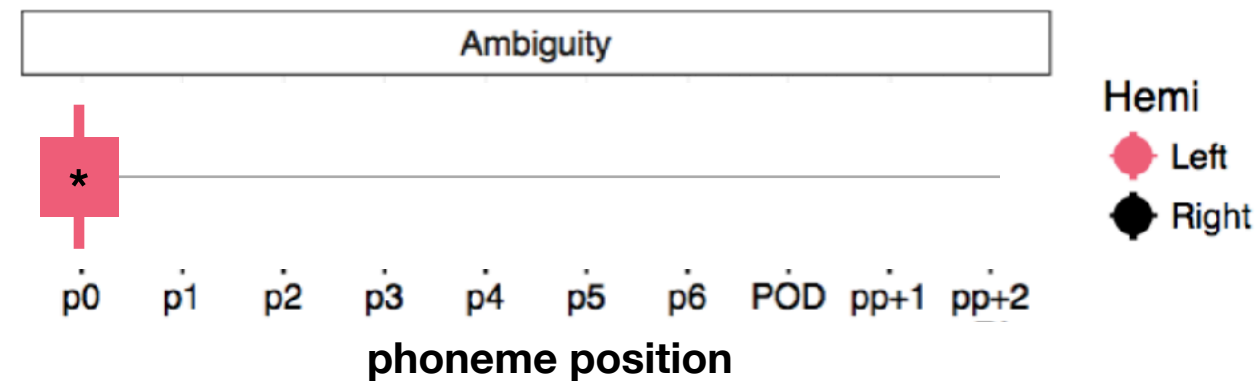
Reactivation in Intermediate Positions



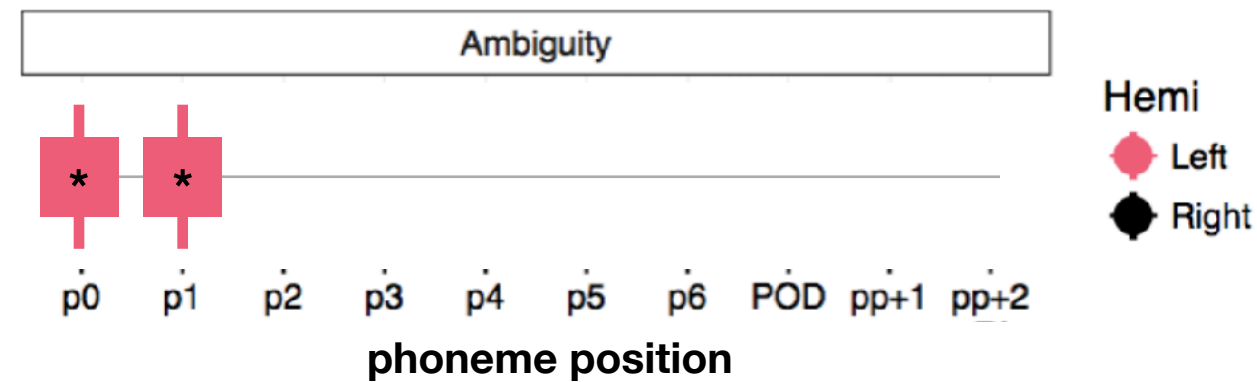
Reactivation in Intermediate Positions



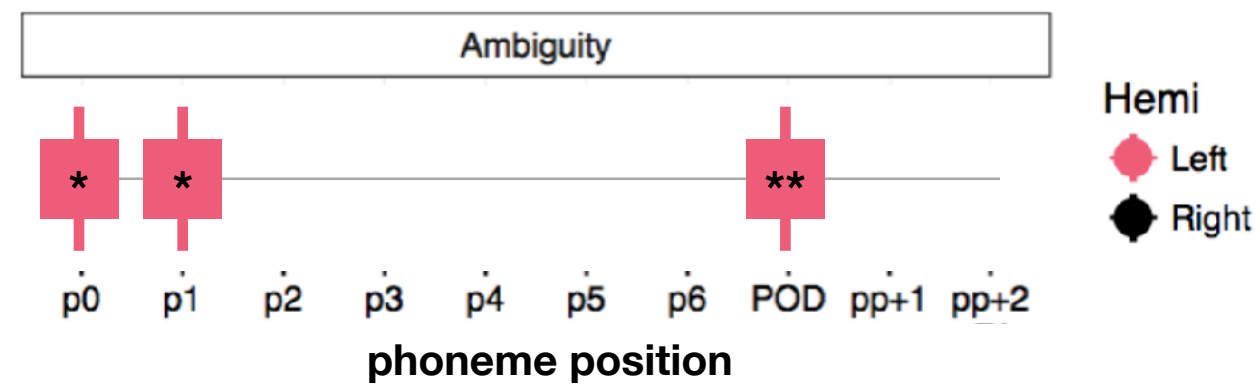
Reactivation in Intermediate Positions



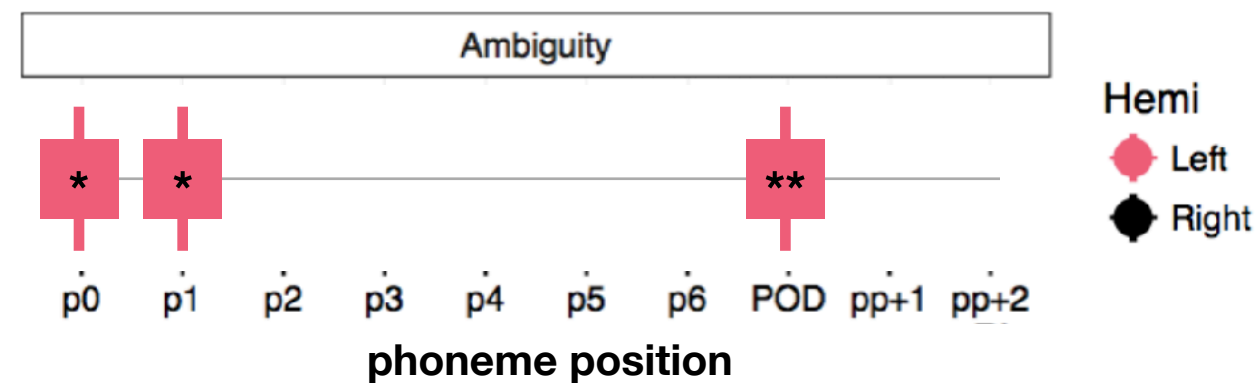
Reactivation in Intermediate Positions



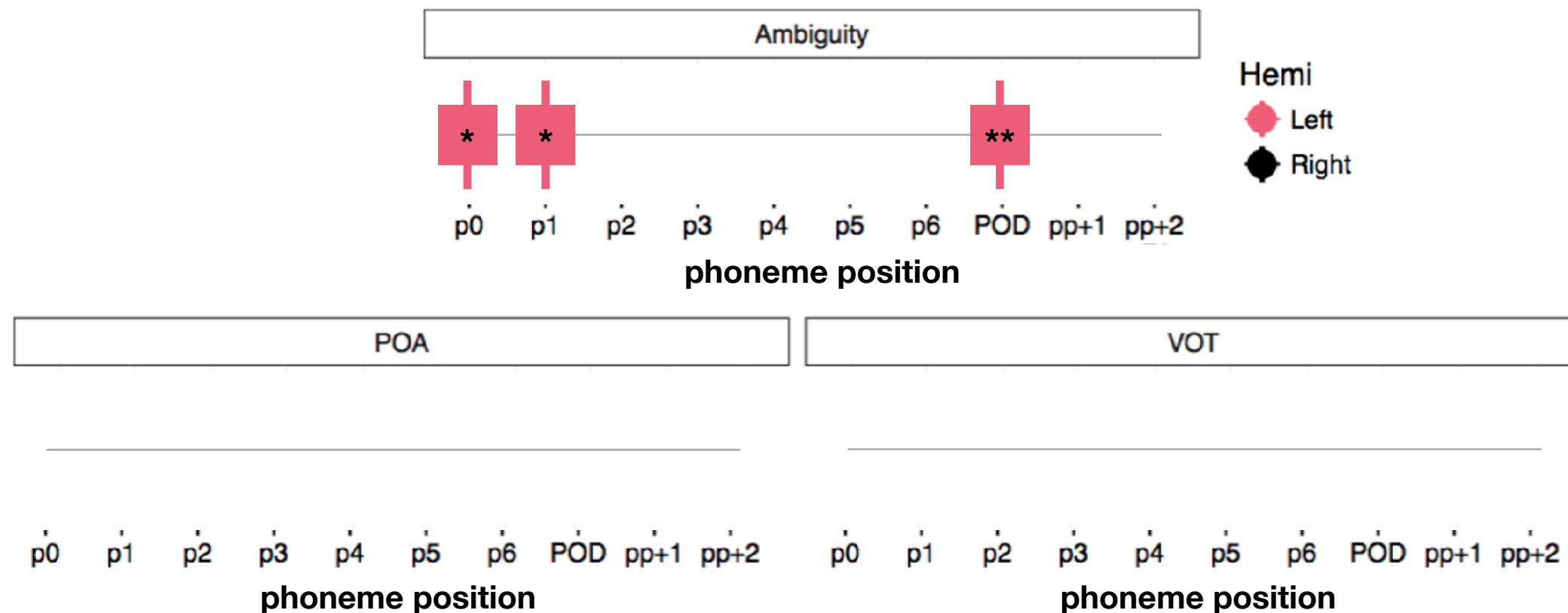
Reactivation in Intermediate Positions



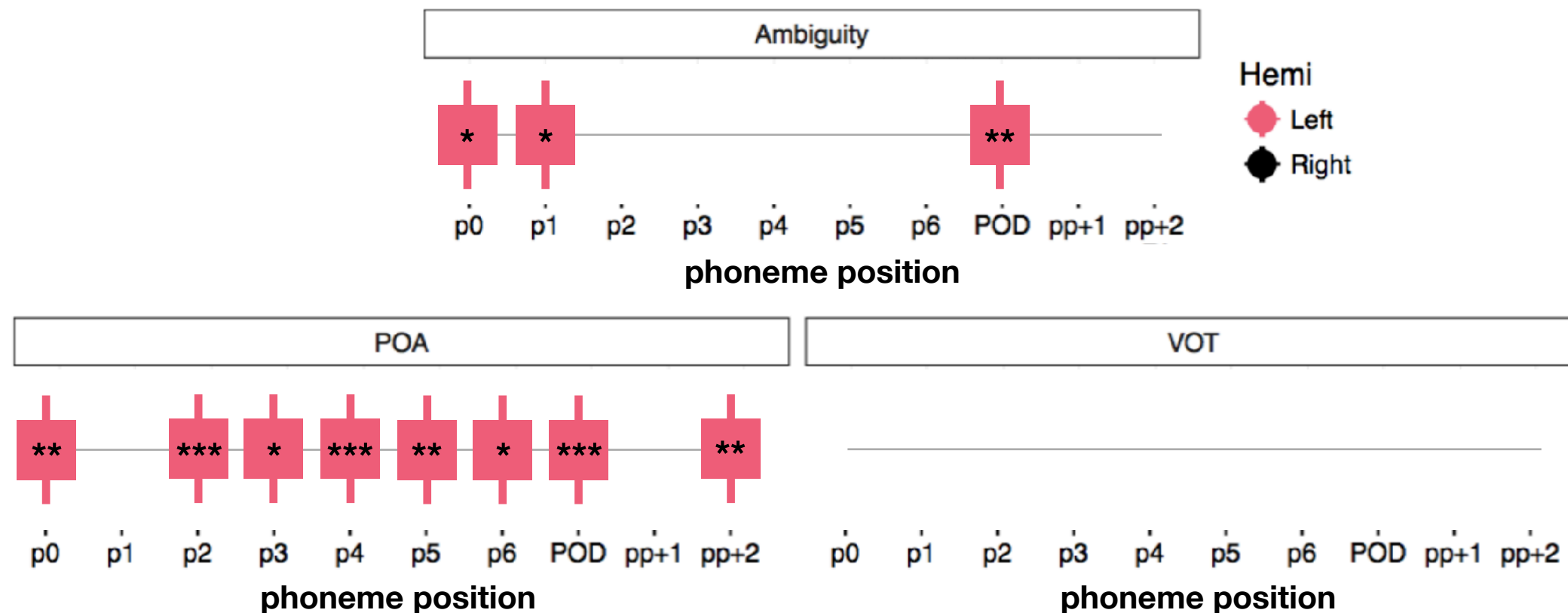
Reactivation in Intermediate Positions



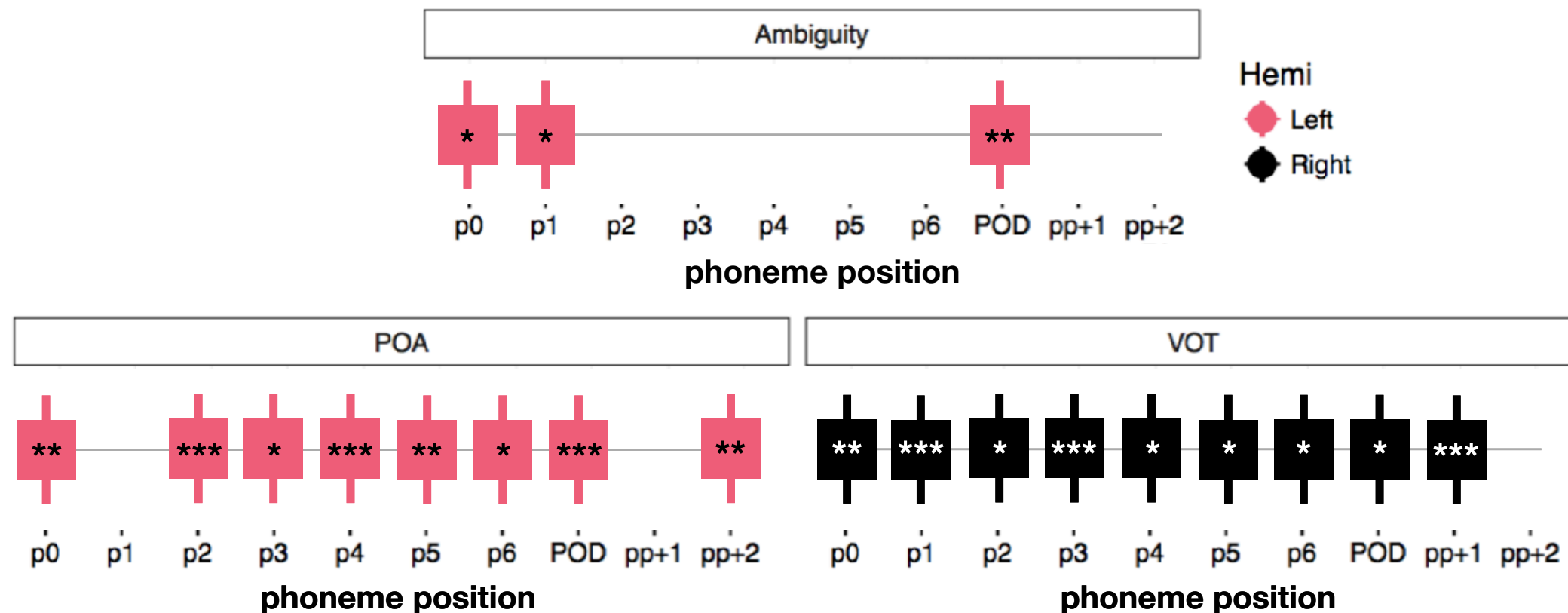
Reactivation in Intermediate Positions



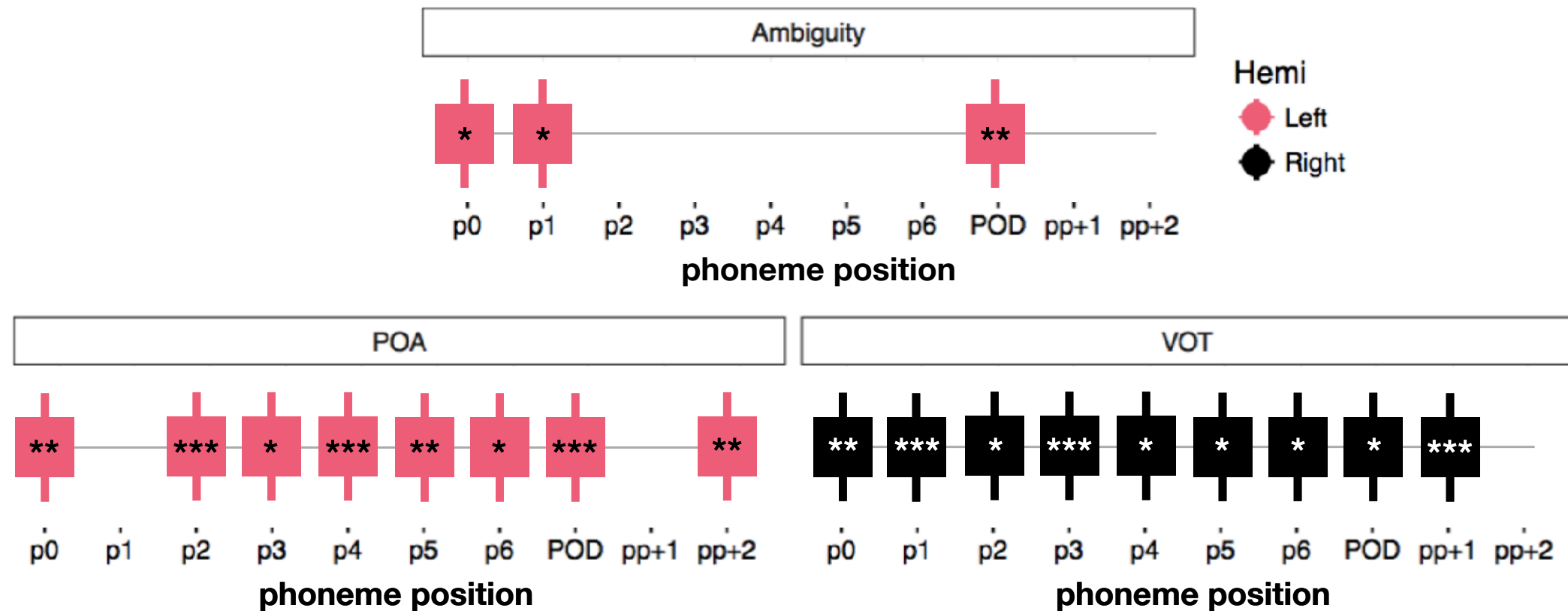
Reactivation in Intermediate Positions



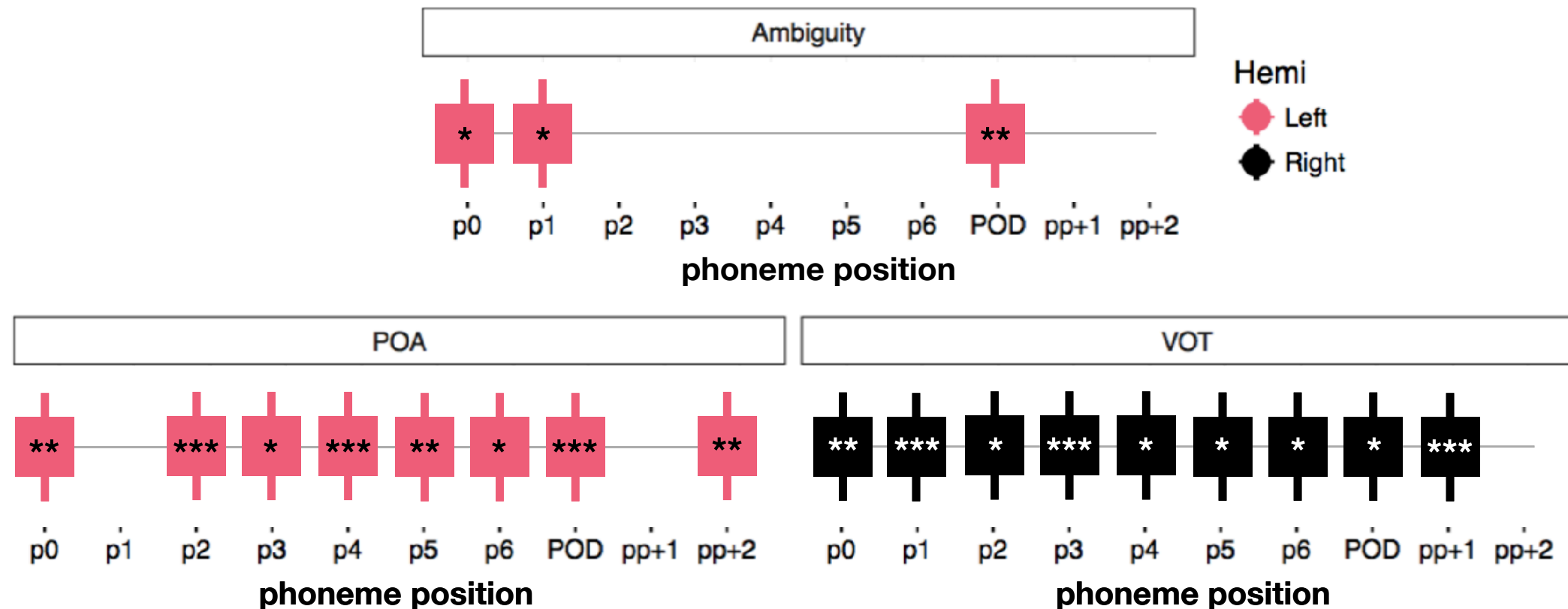
Reactivation in Intermediate Positions



Reactivation in Intermediate Positions

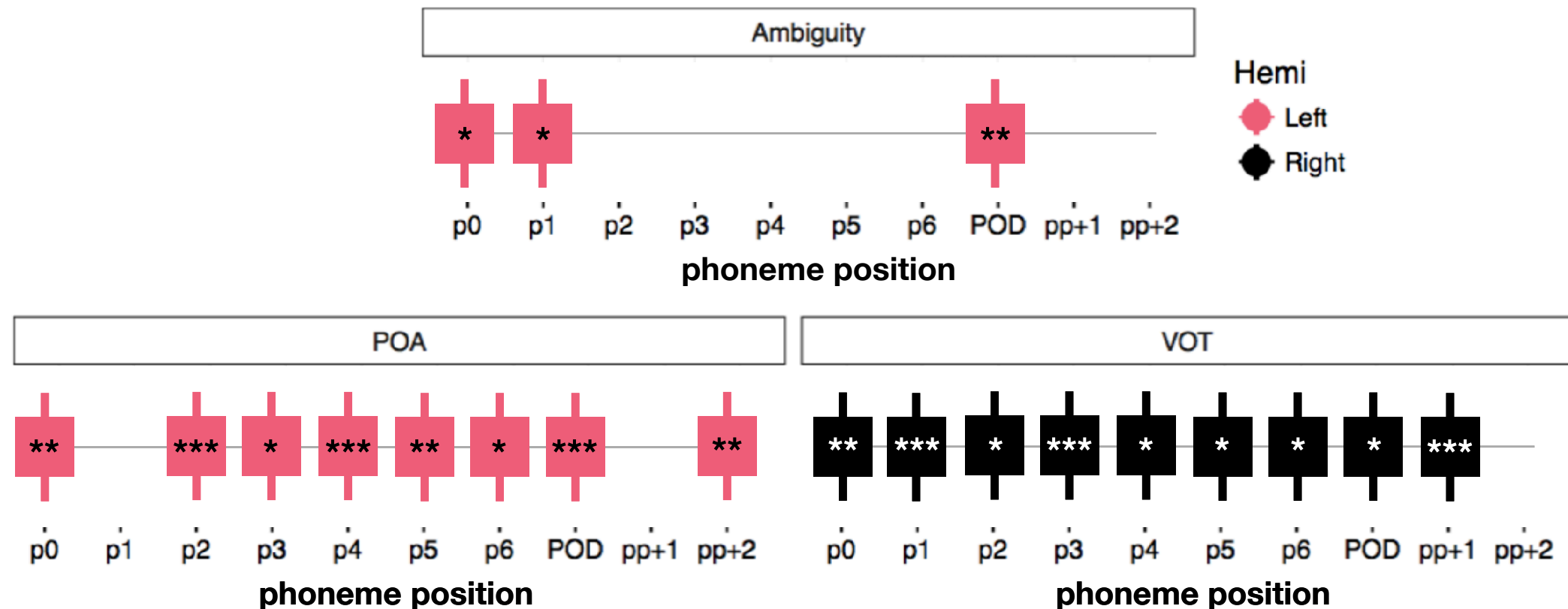


Reactivation in Intermediate Positions



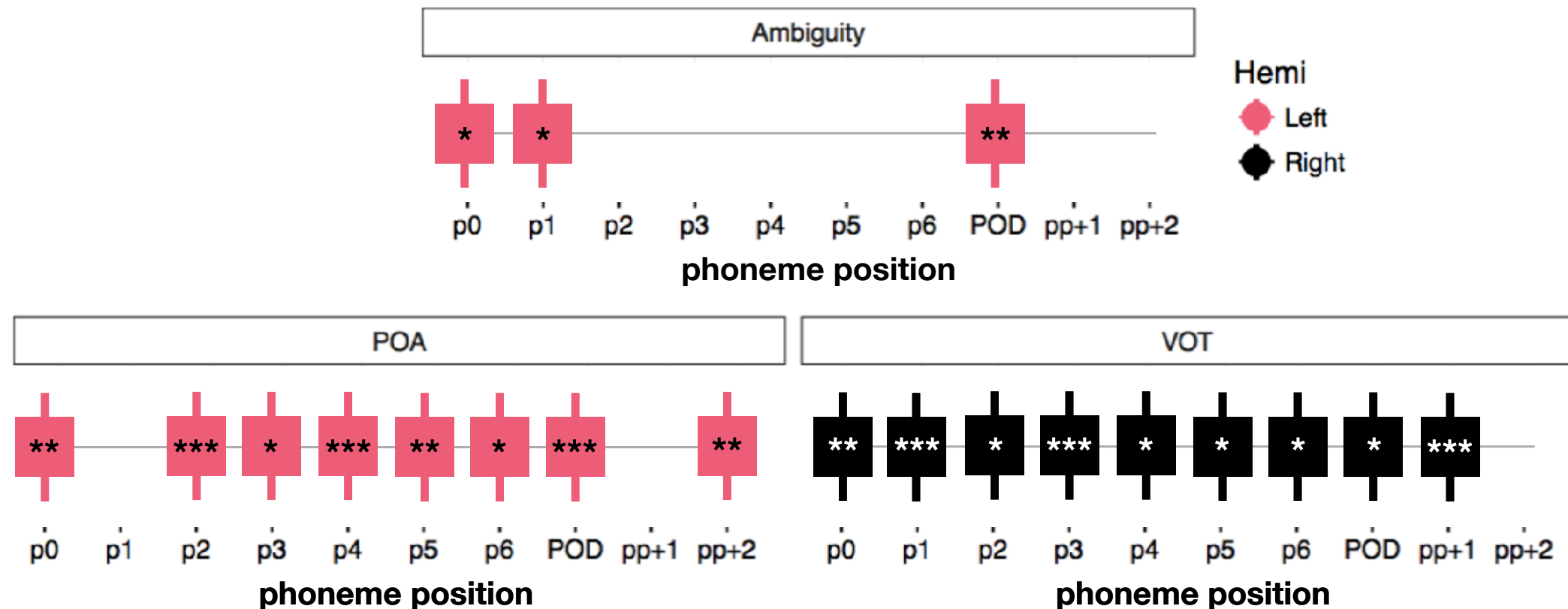
- Information is re-evoked in auditory cortex

Reactivation in Intermediate Positions



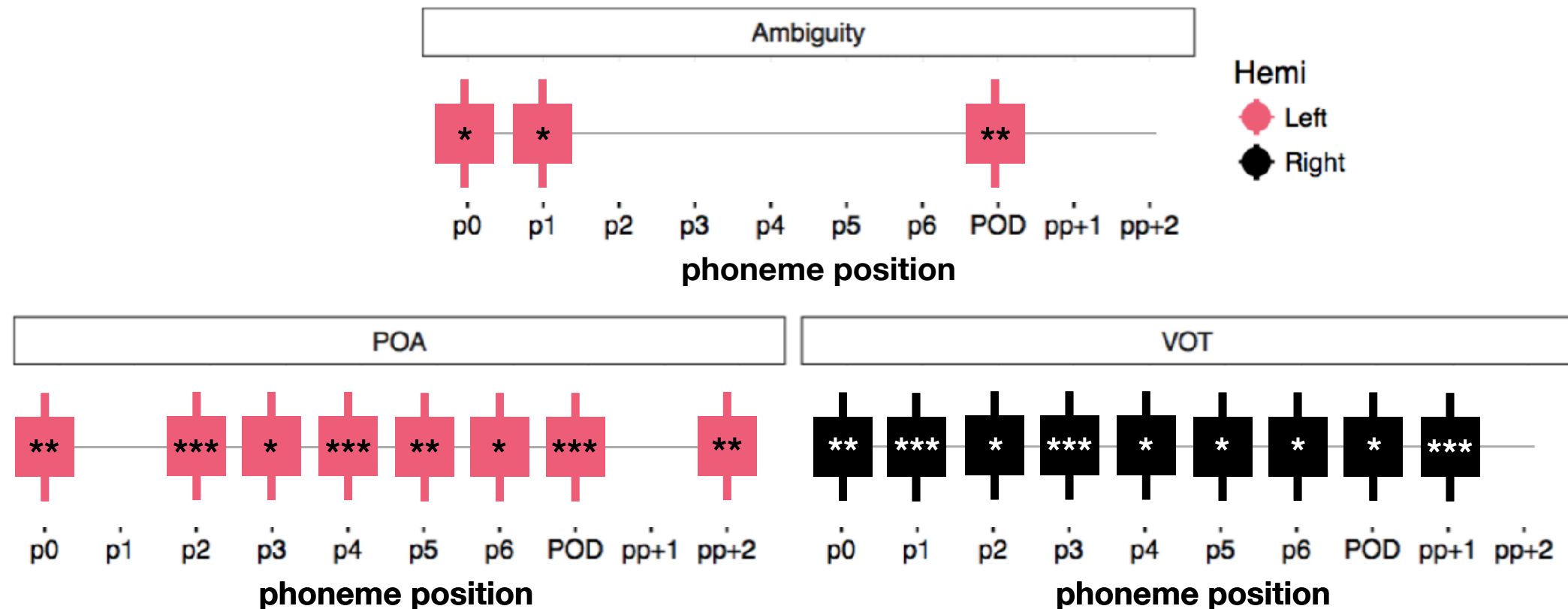
- Information is re-evoked in auditory cortex
- Specifically time-locked to the onset of subsequent phonemes

Reactivation in Intermediate Positions



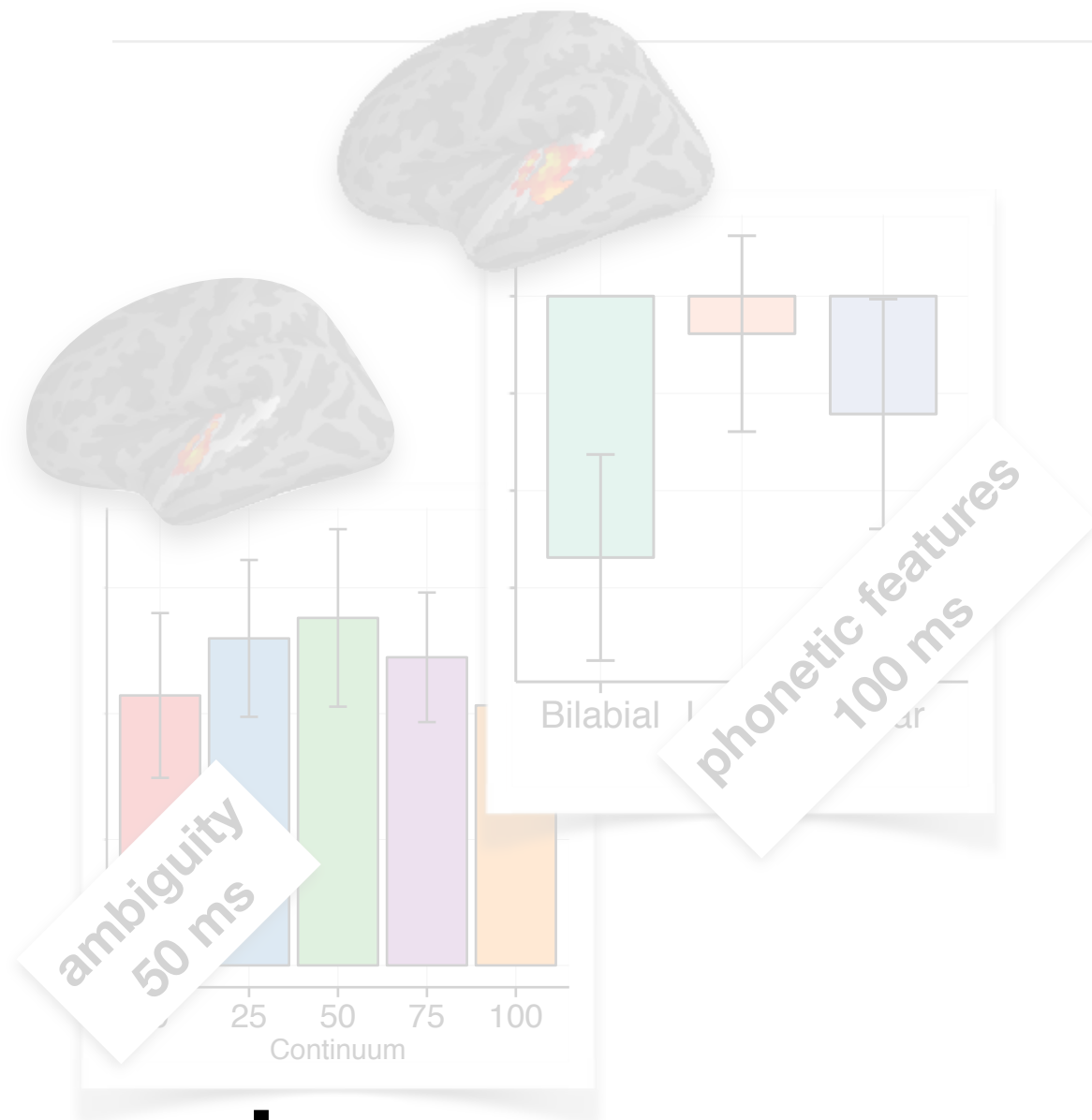
- Information is re-evoked in auditory cortex
- Specifically time-locked to the onset of subsequent phonemes
 - Not driven by residual information in the acoustic signal

Reactivation in Intermediate Positions



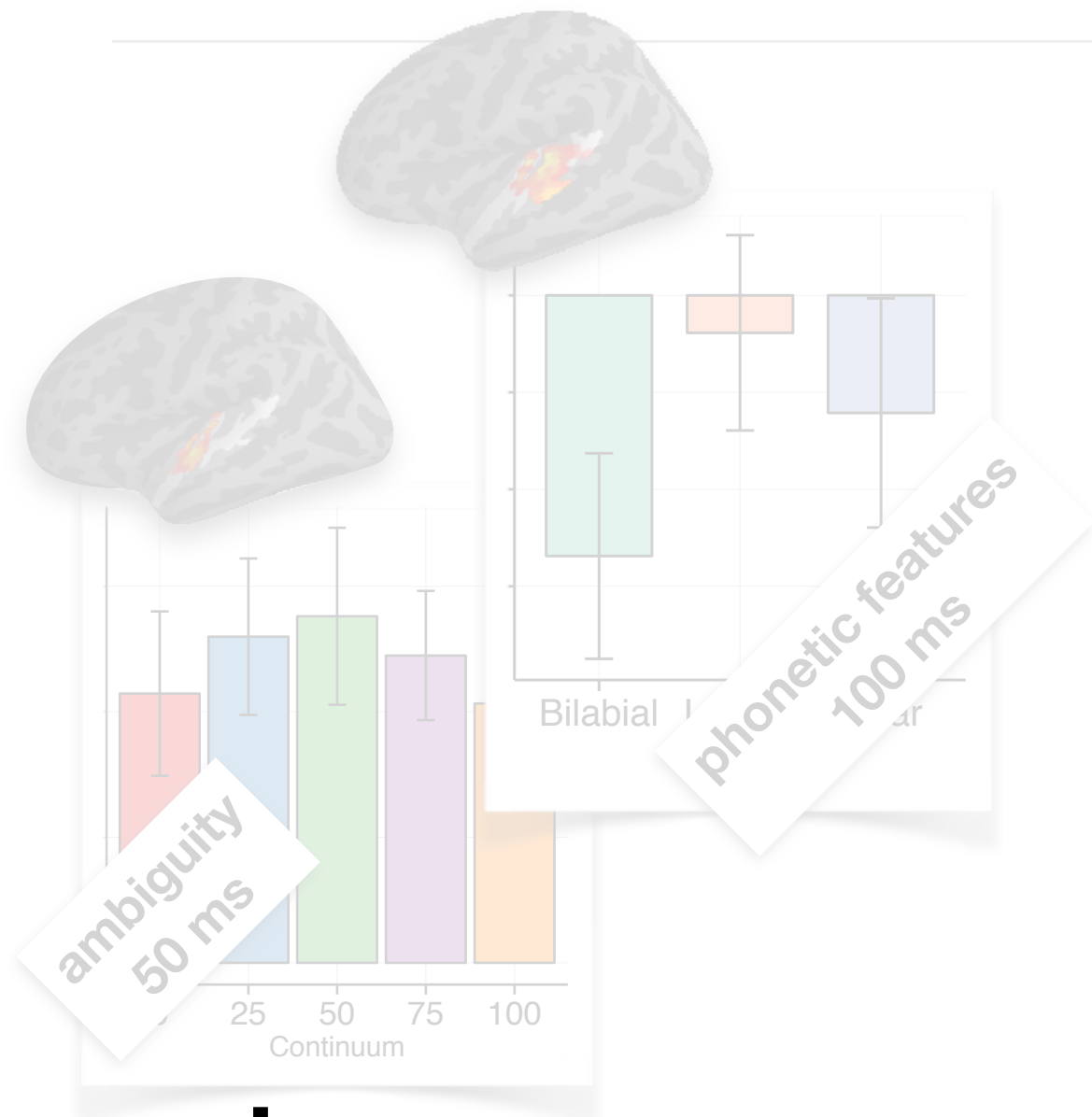
- Information is re-evoked in auditory cortex
- Specifically time-locked to the onset of subsequent phonemes
 - Not driven by residual information in the acoustic signal
- Not specific to the ambiguous tokens — general to language processing

Interim Conclusion



p a r a k e e t

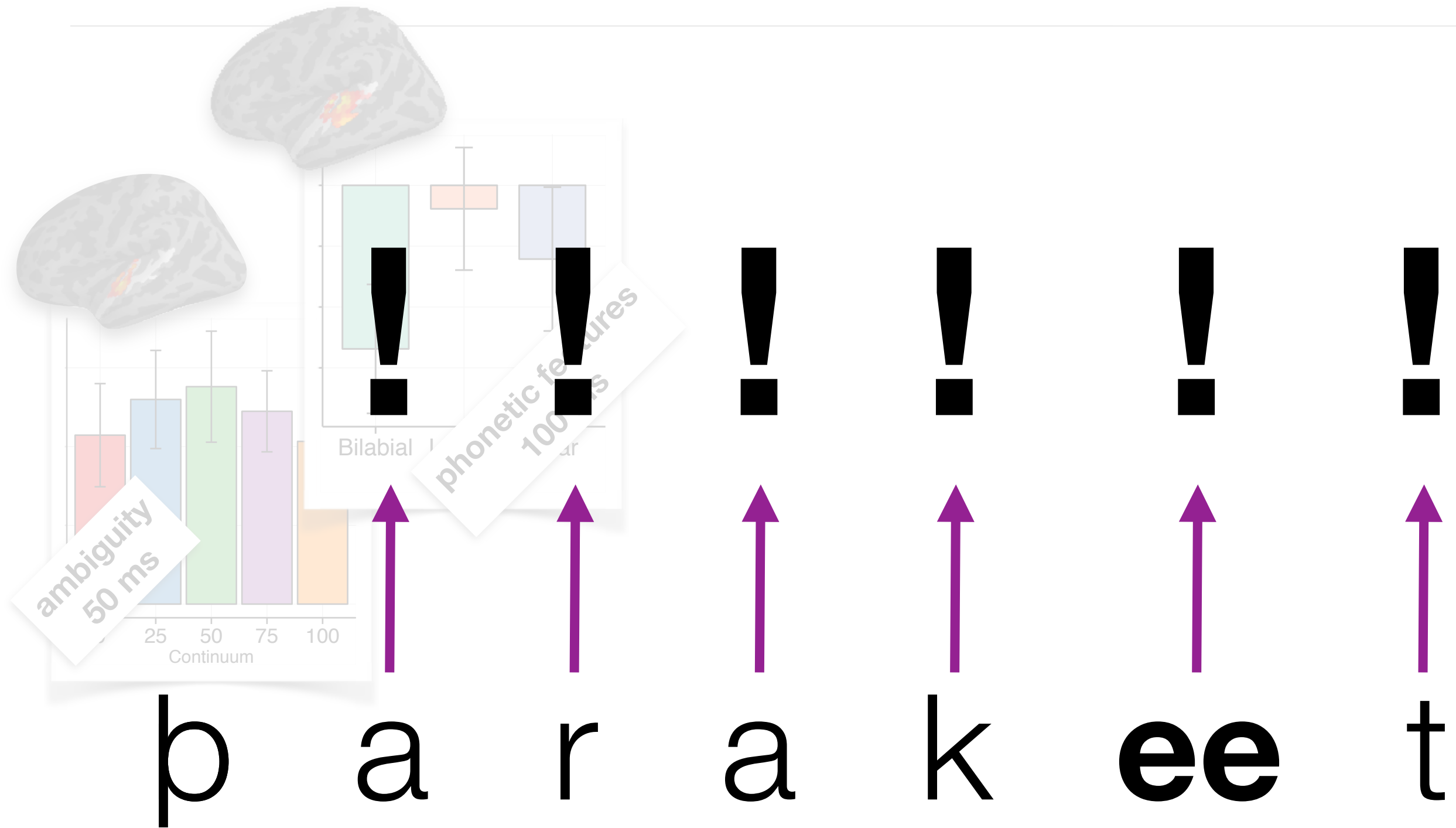
Interim Conclusion



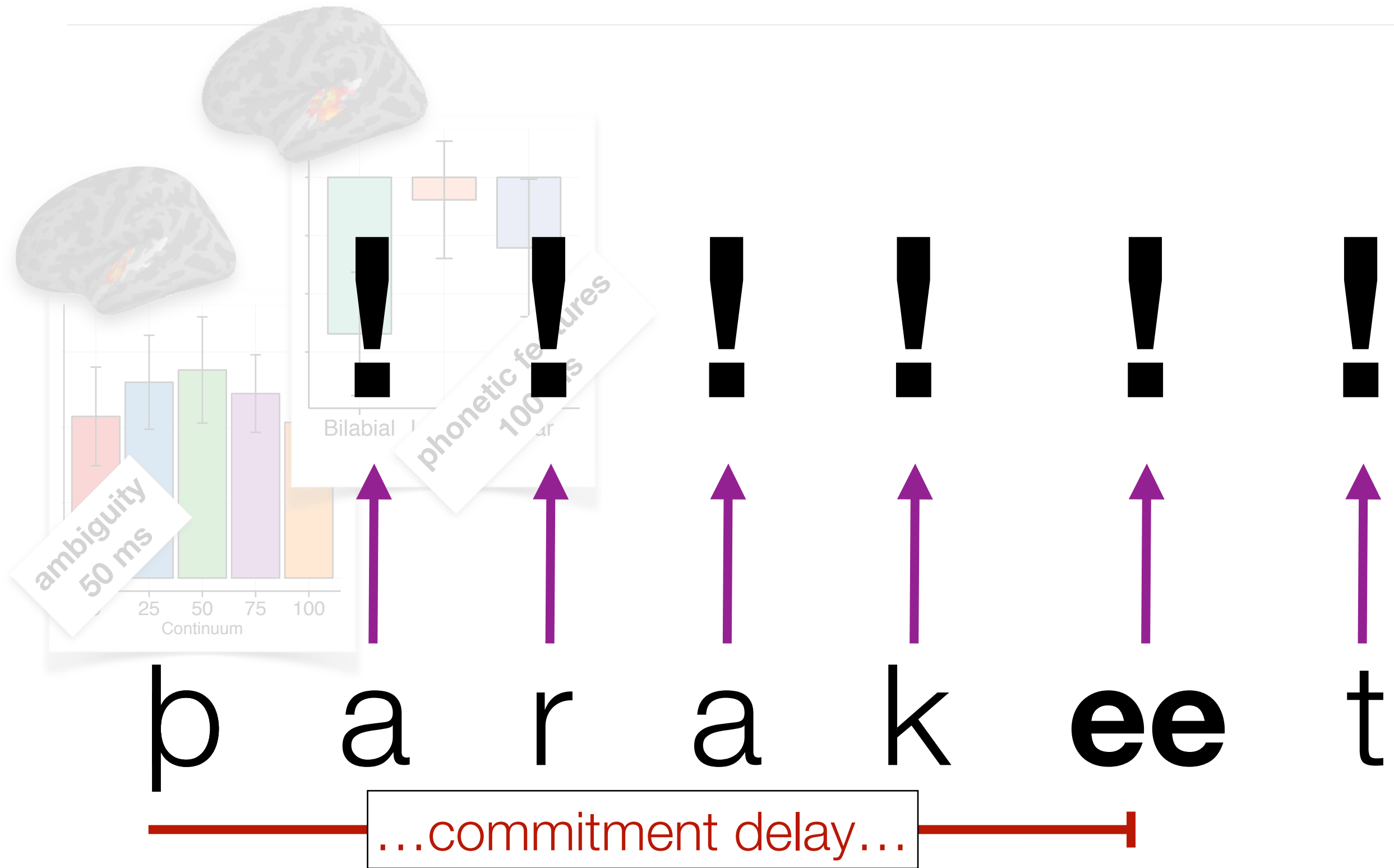
p a r a k e e t

! ↑

Interim Conclusion



Interim Conclusion



Today's Questions

How long can the system delay
phonological commitment?

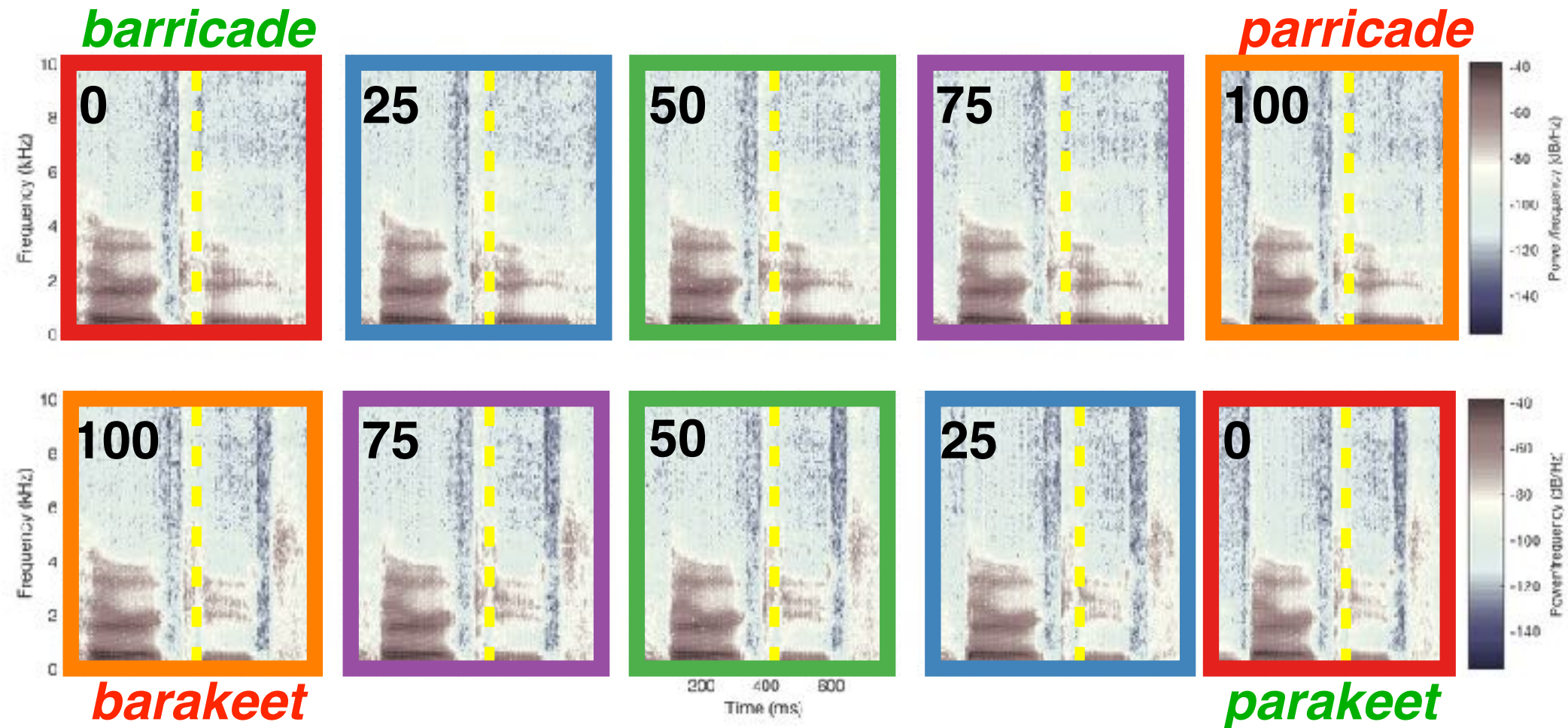
Today's Questions

How long can the system delay
phonological commitment?

Psycholinguistic investigations into this question:

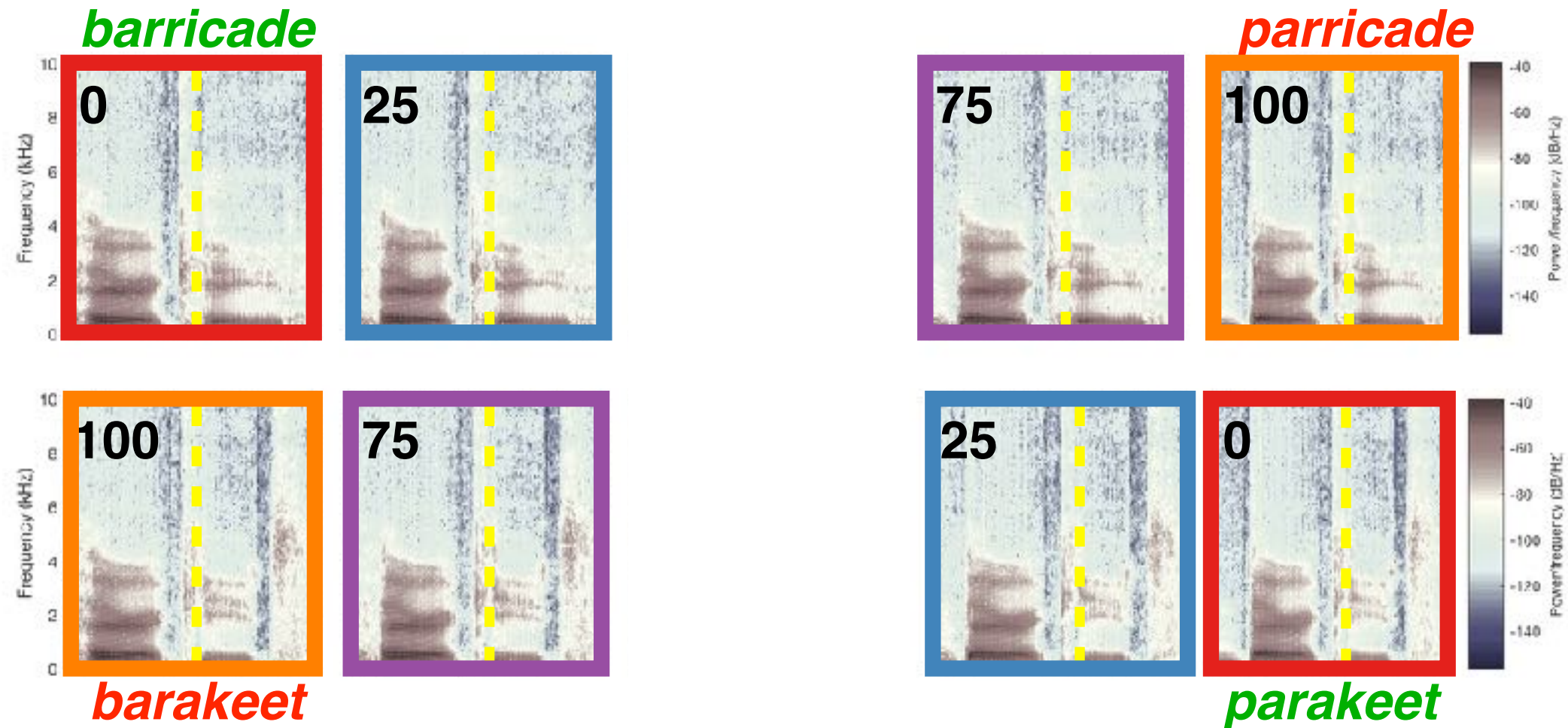
Connine et al. 1991; Samuel 1991; McMurray et al. 2009; Szostak and Pitt 2013

Example Continuum Pair



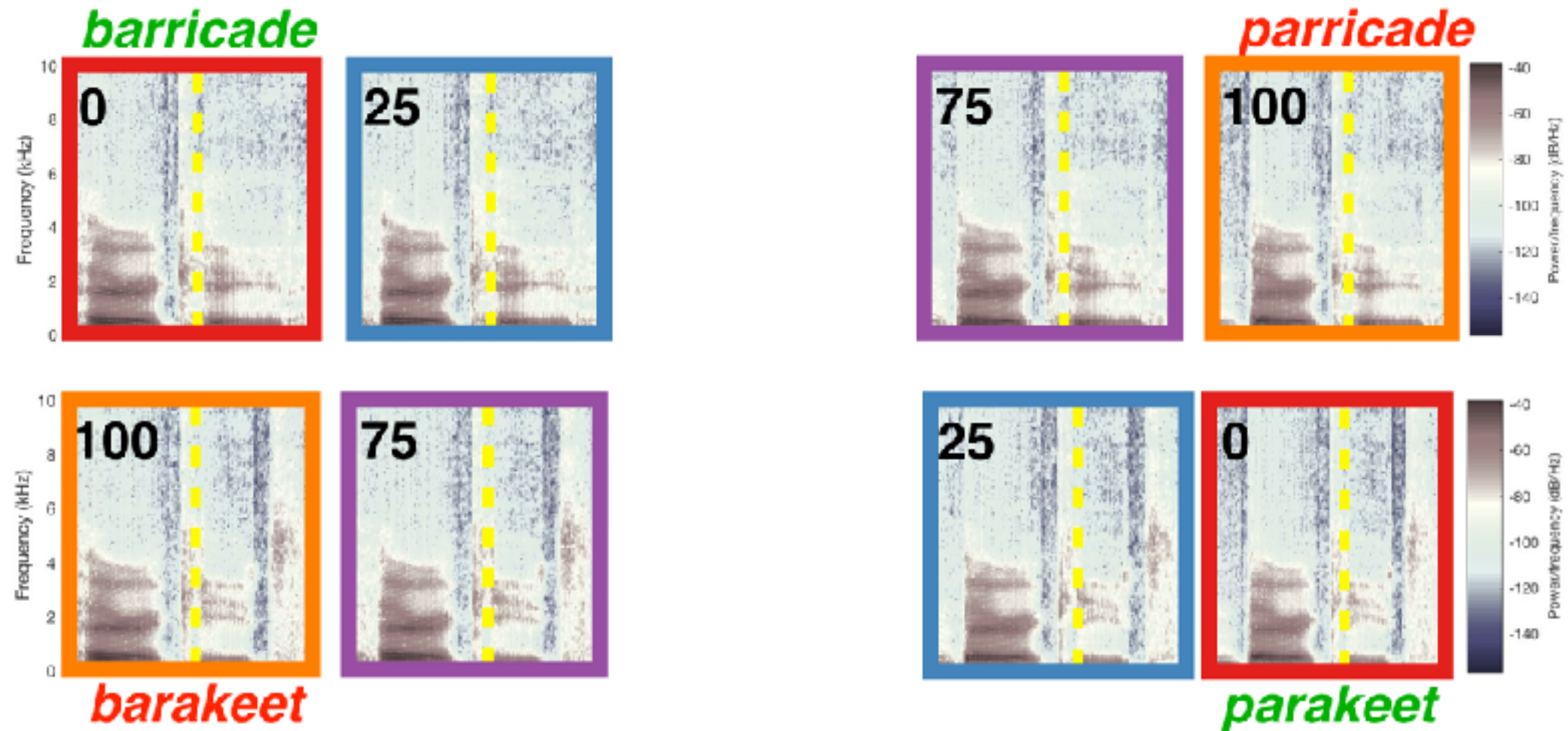
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



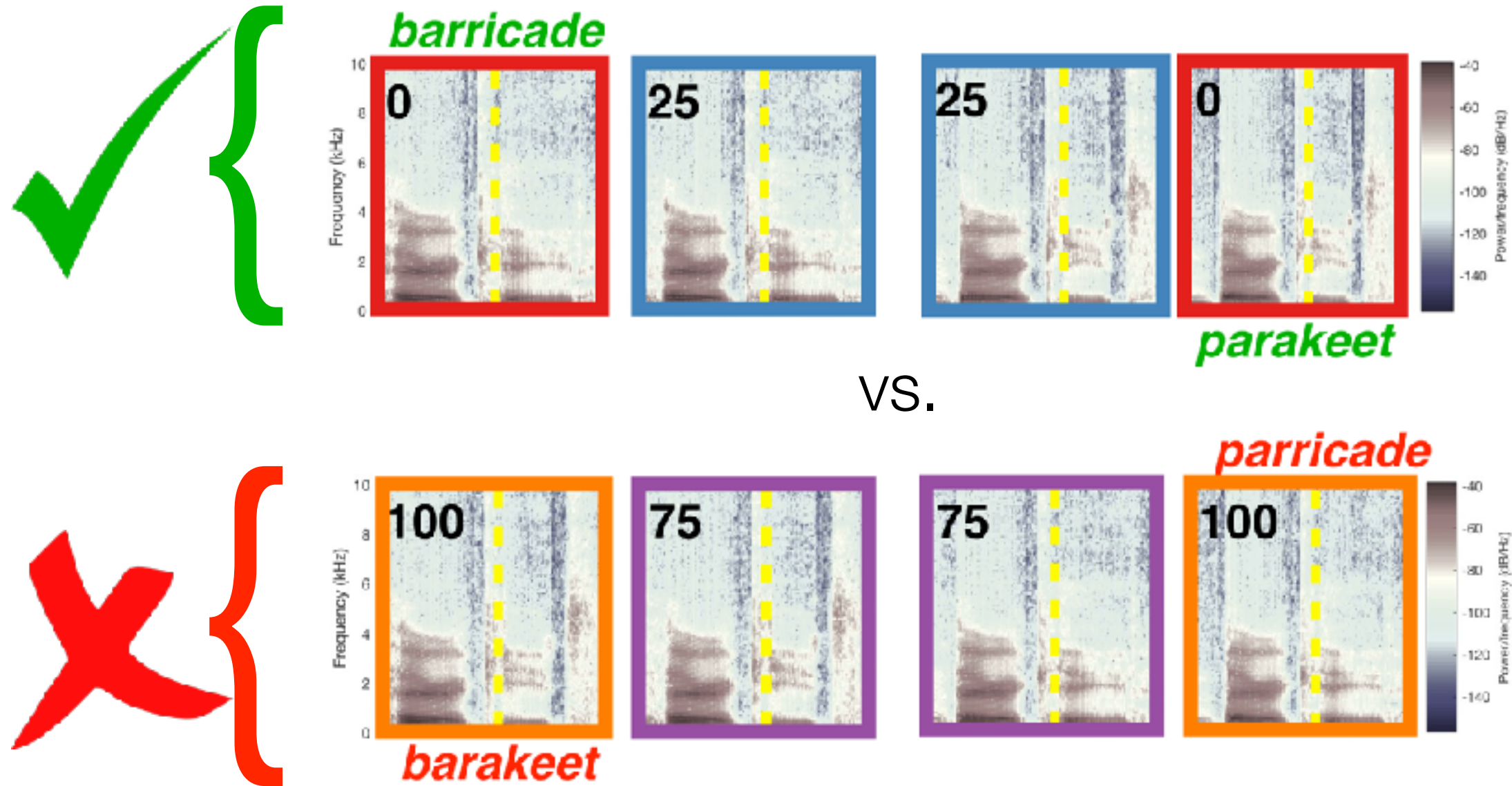
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



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

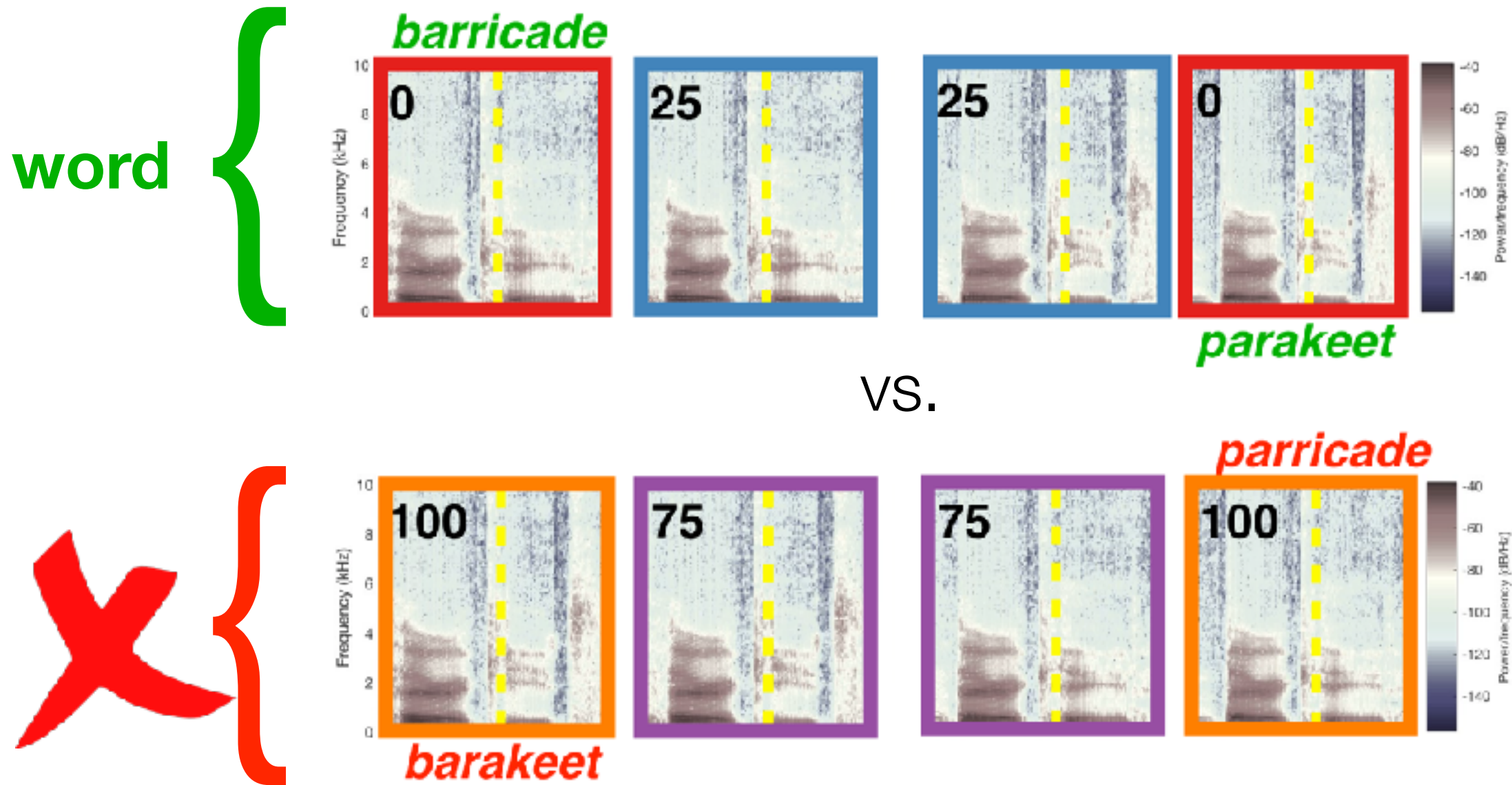
Example Continuum Pair



VS.

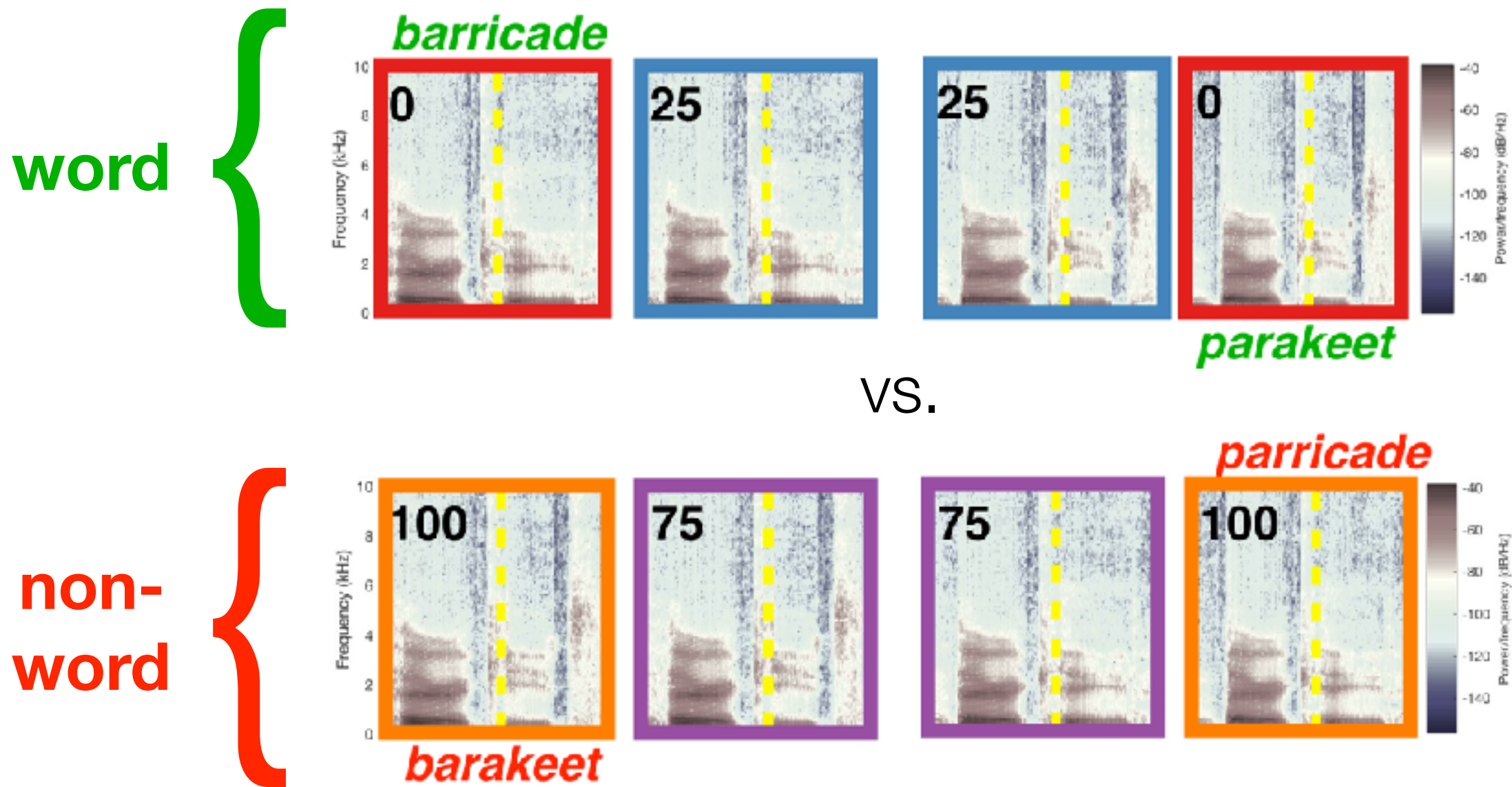
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



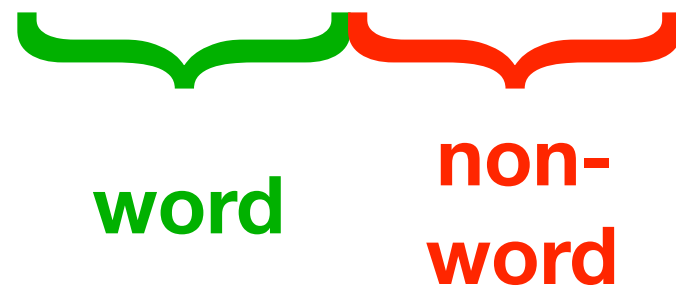
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



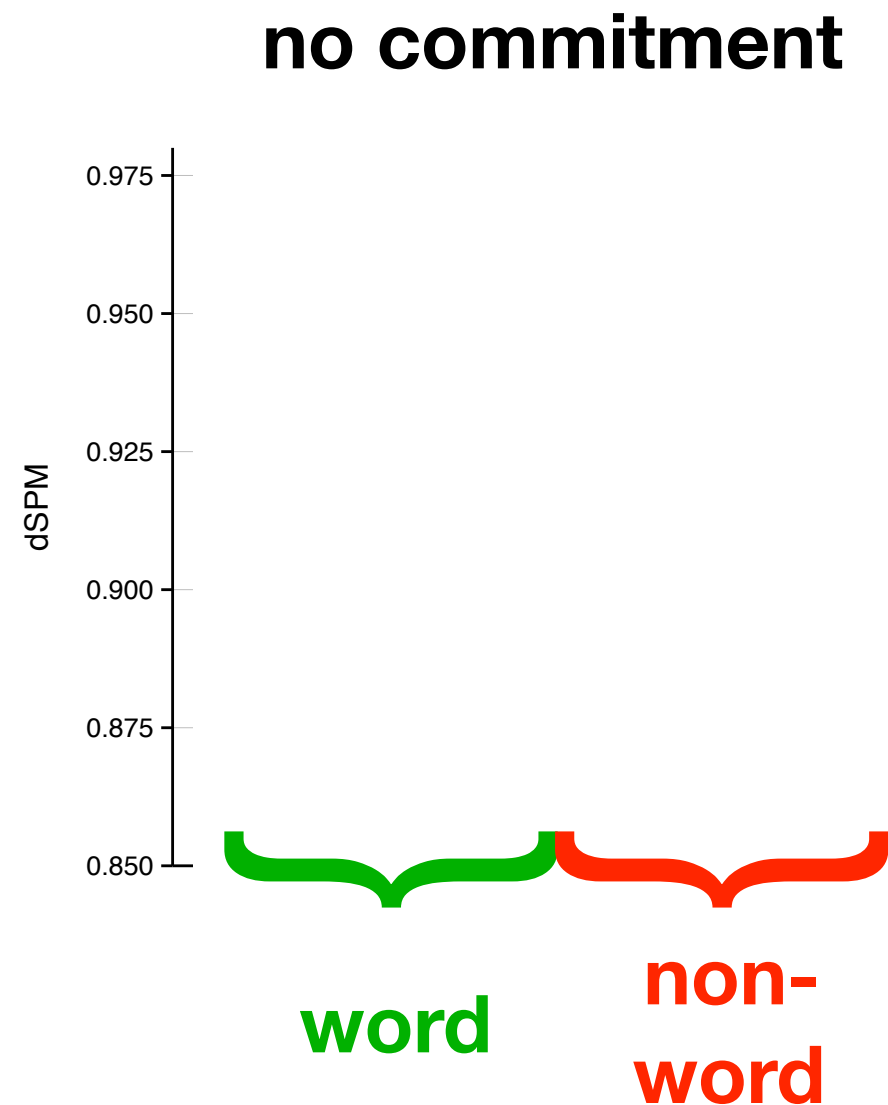
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



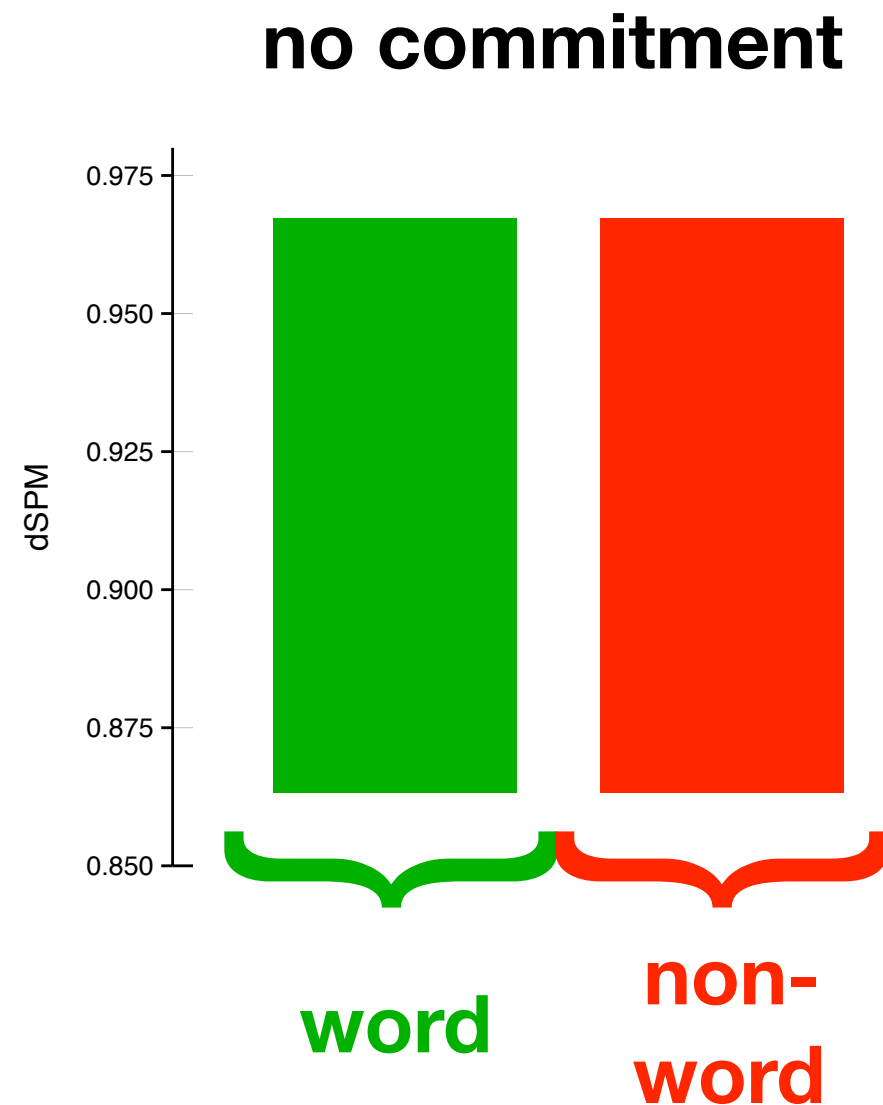
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



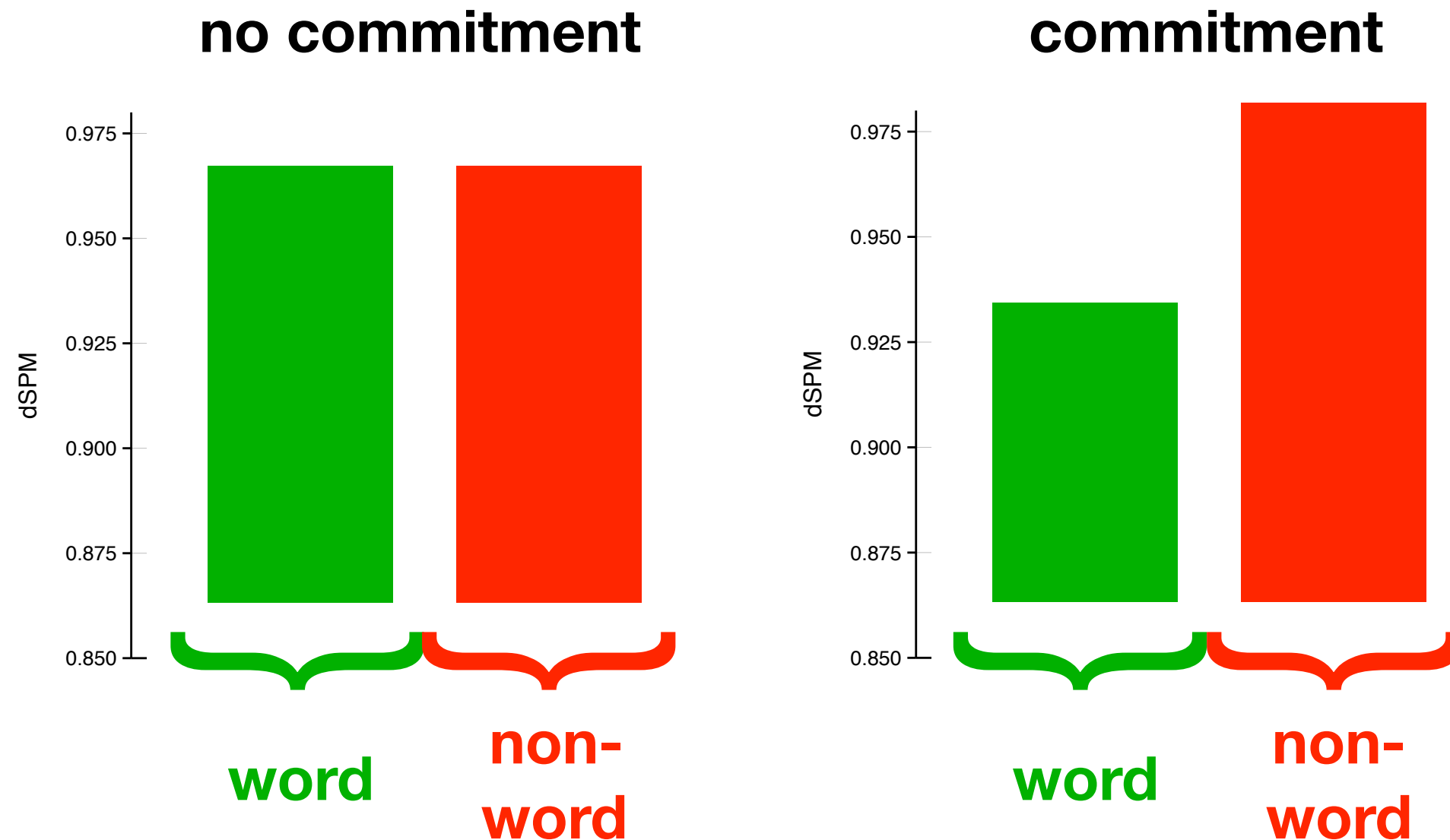
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Example Continuum Pair



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

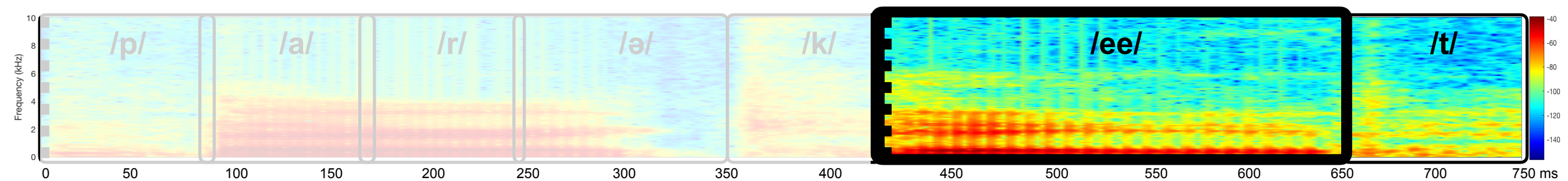
Example Continuum Pair



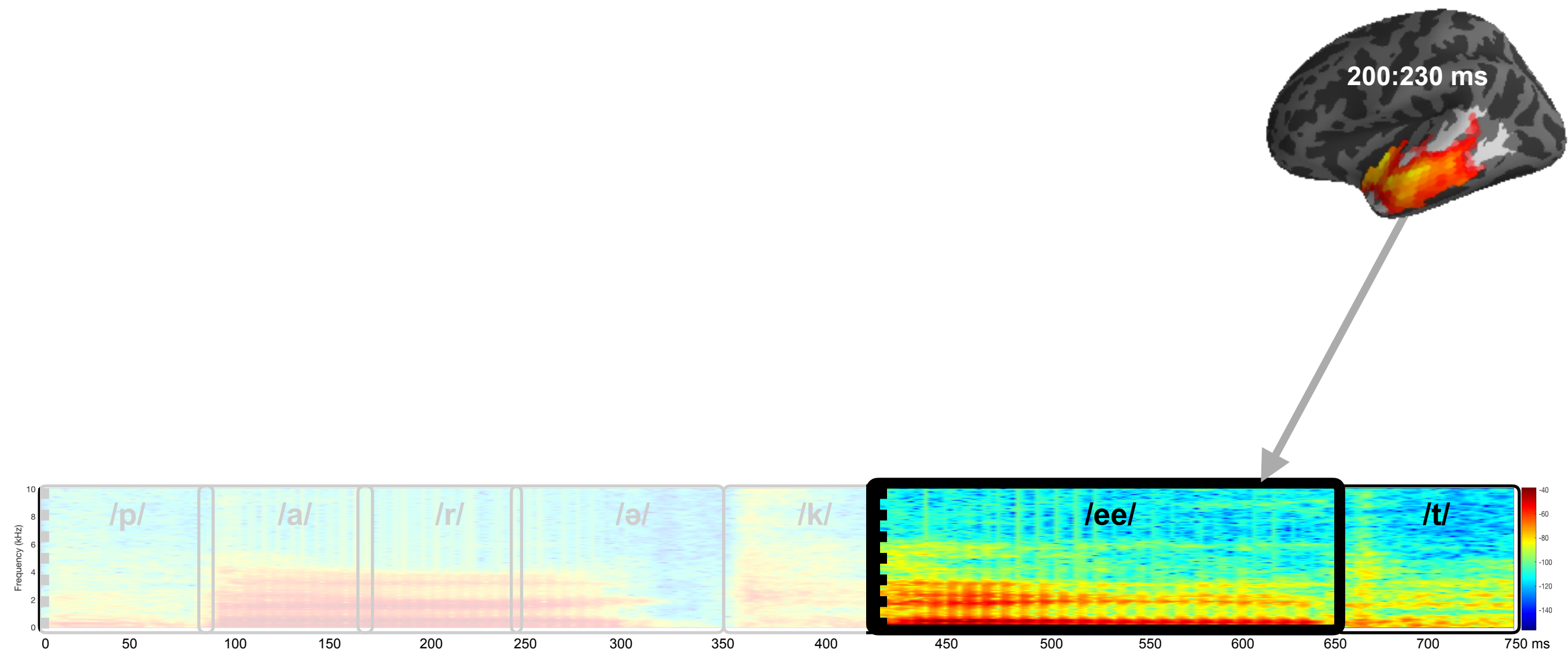
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms

Commitment Before POD

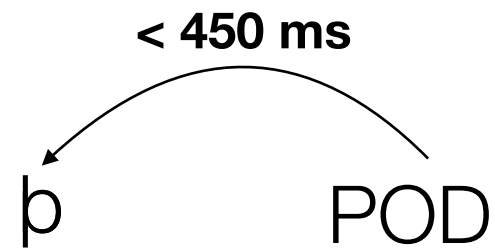
Commitment Before POD



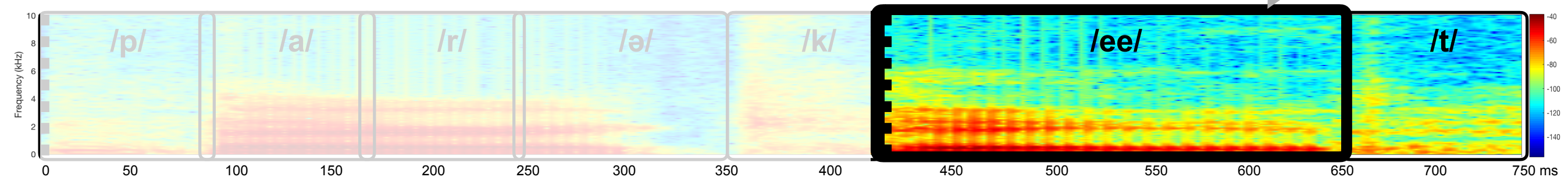
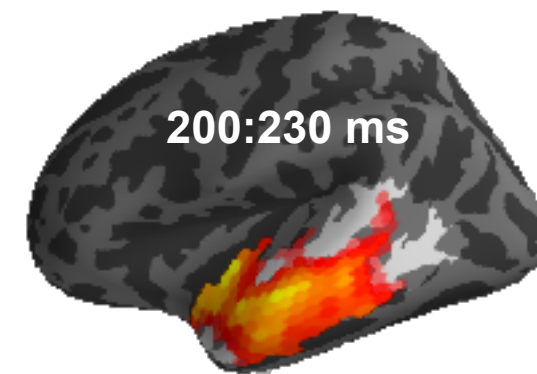
Commitment Before POD



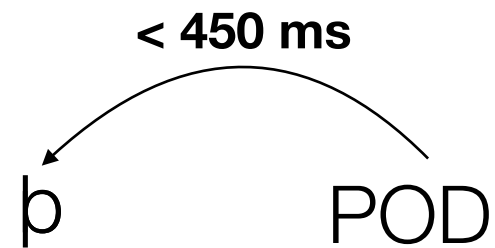
Commitment Before POD



Early: POD earlier than 450 ms after word onset

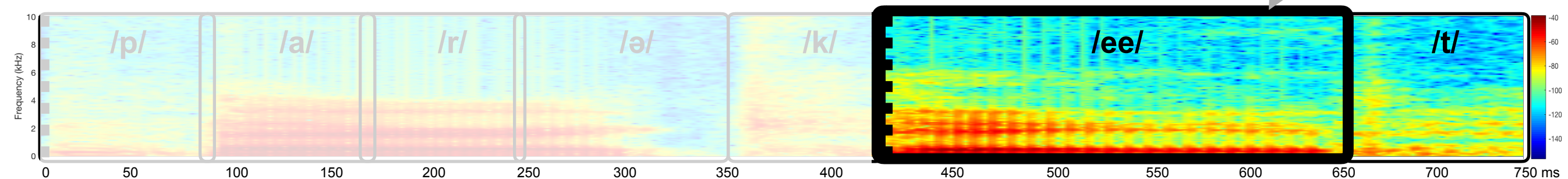
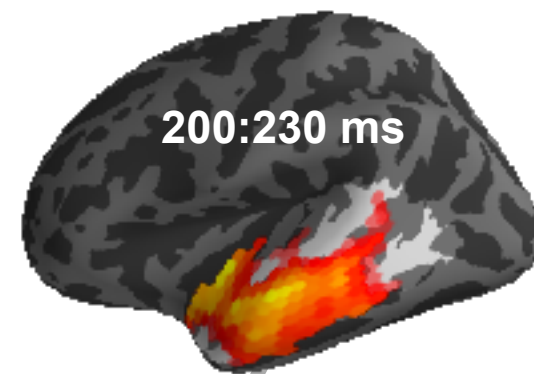


Commitment Before POD



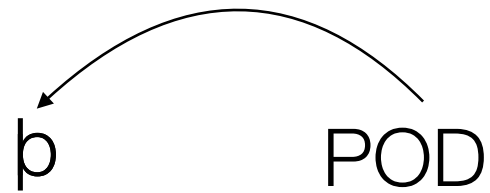
Early: POD earlier than 450 ms after word onset

Late: POD later than 450 ms after word onset



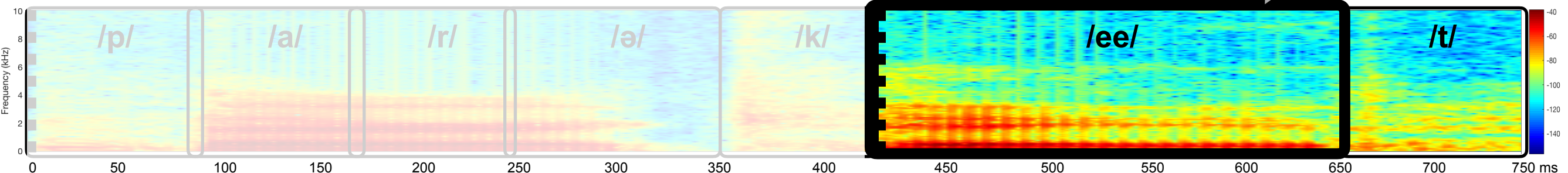
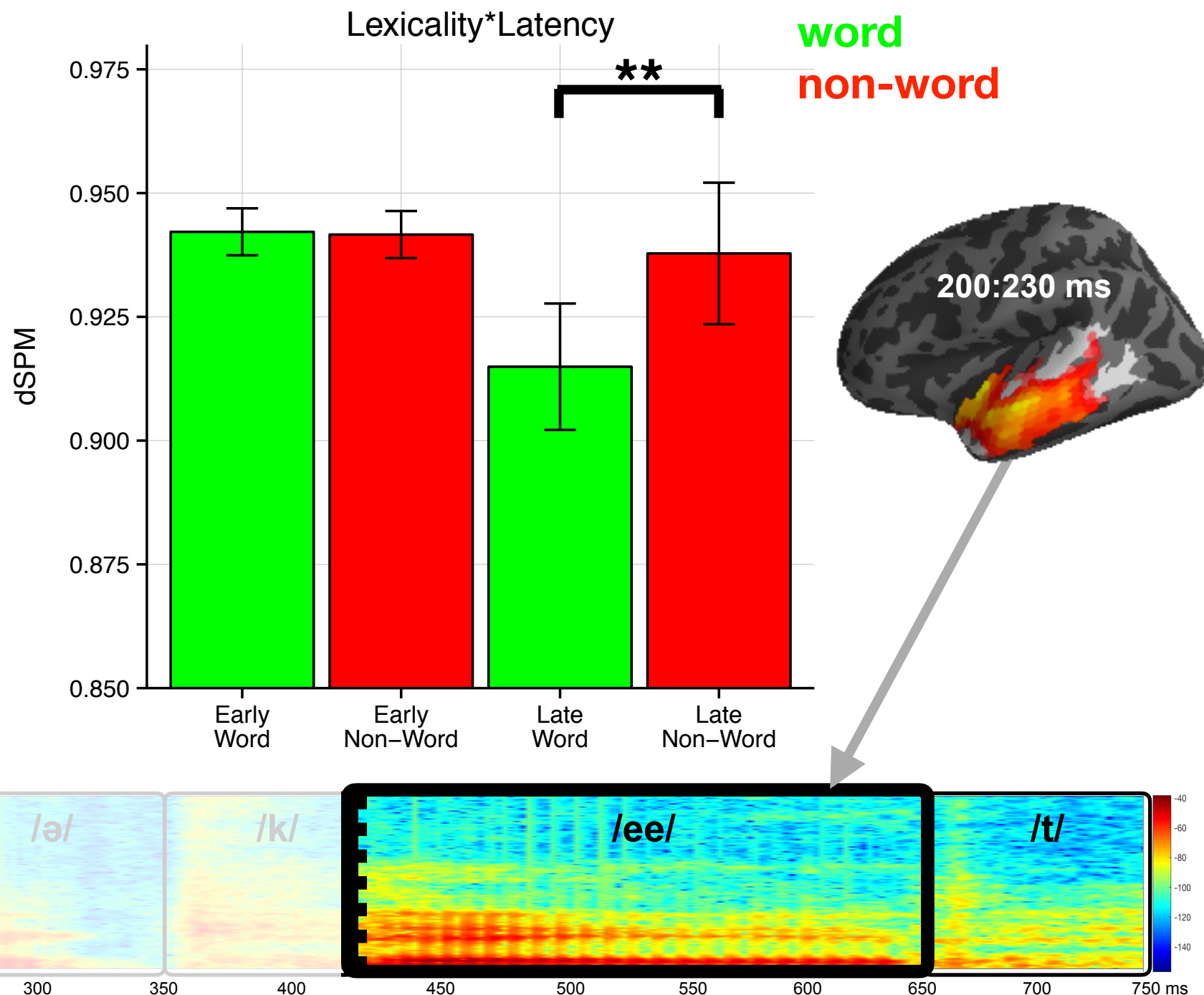
Commitment Before POD

< 450 ms



Early: POD earlier than 450 ms after word onset

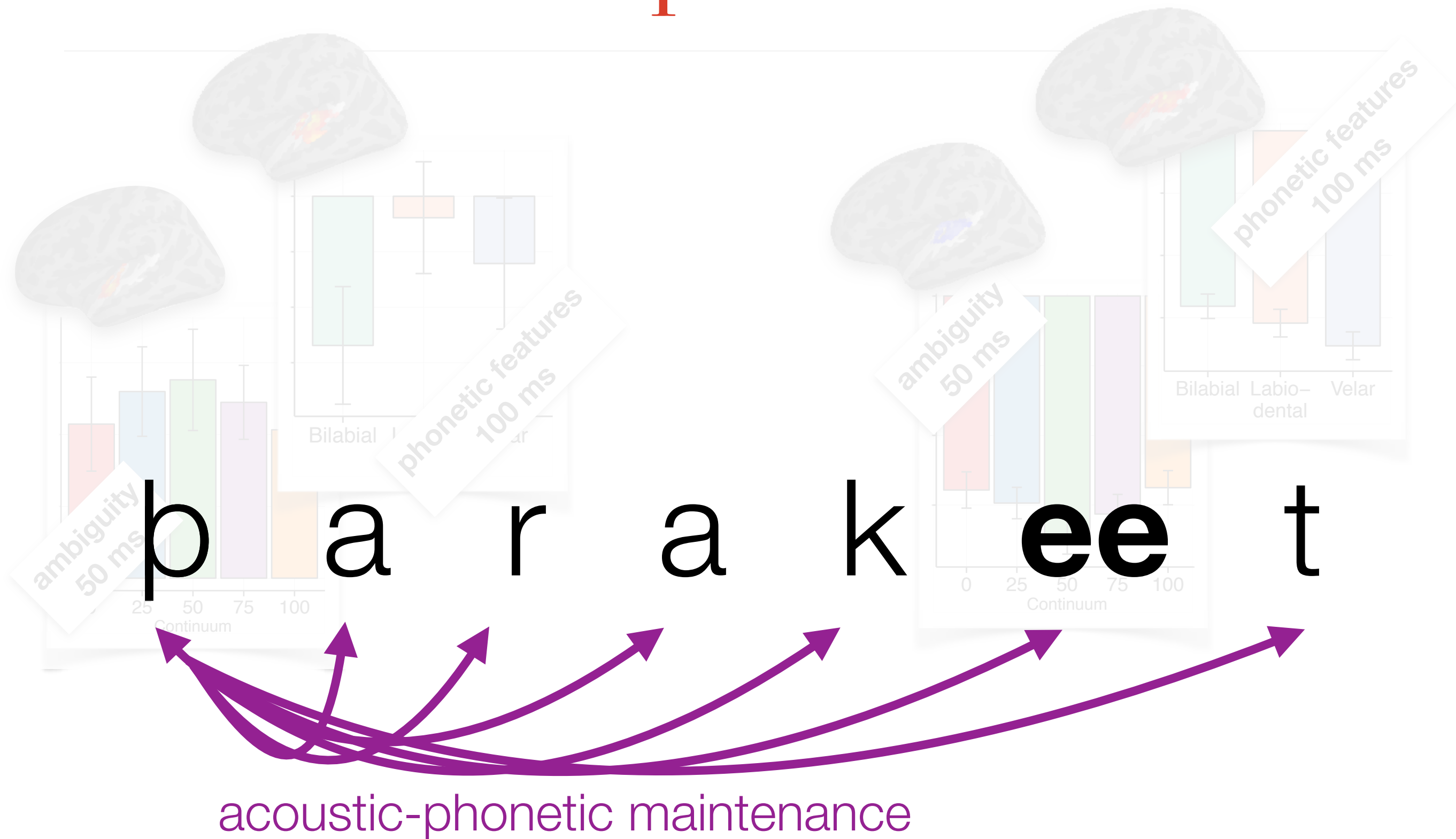
Late: POD later than 450 ms after word onset



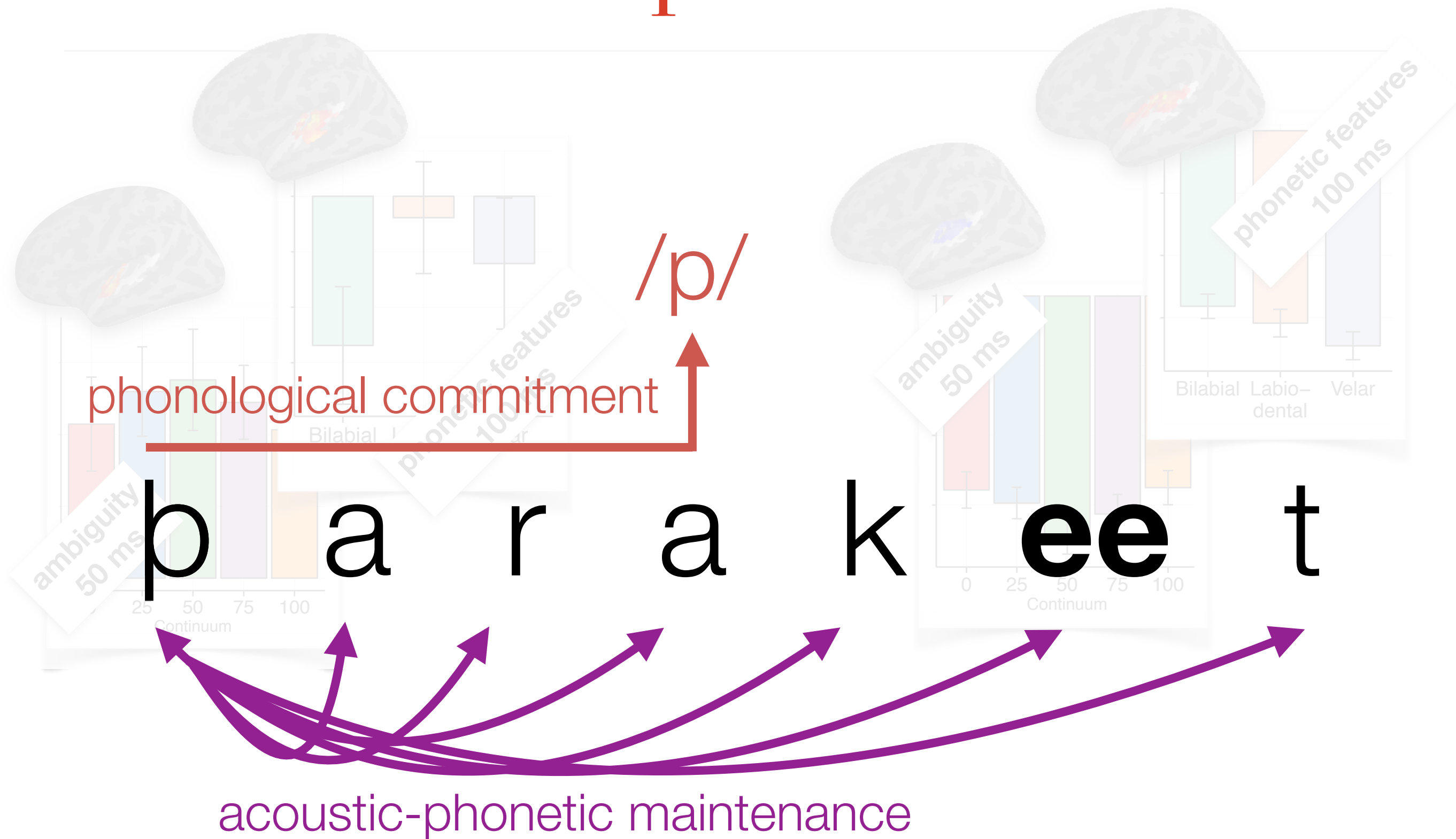
Interpretation



Interpretation



Interpretation



Interpretation

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

phonological commitment

/p/

b a r a k e e t

acoustic-phonetic maintenance

Interpretation

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

lexical access



phonological commitment

/p/

b a r a k e e t

acoustic-phonetic maintenance

Interpretation

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

lexical access



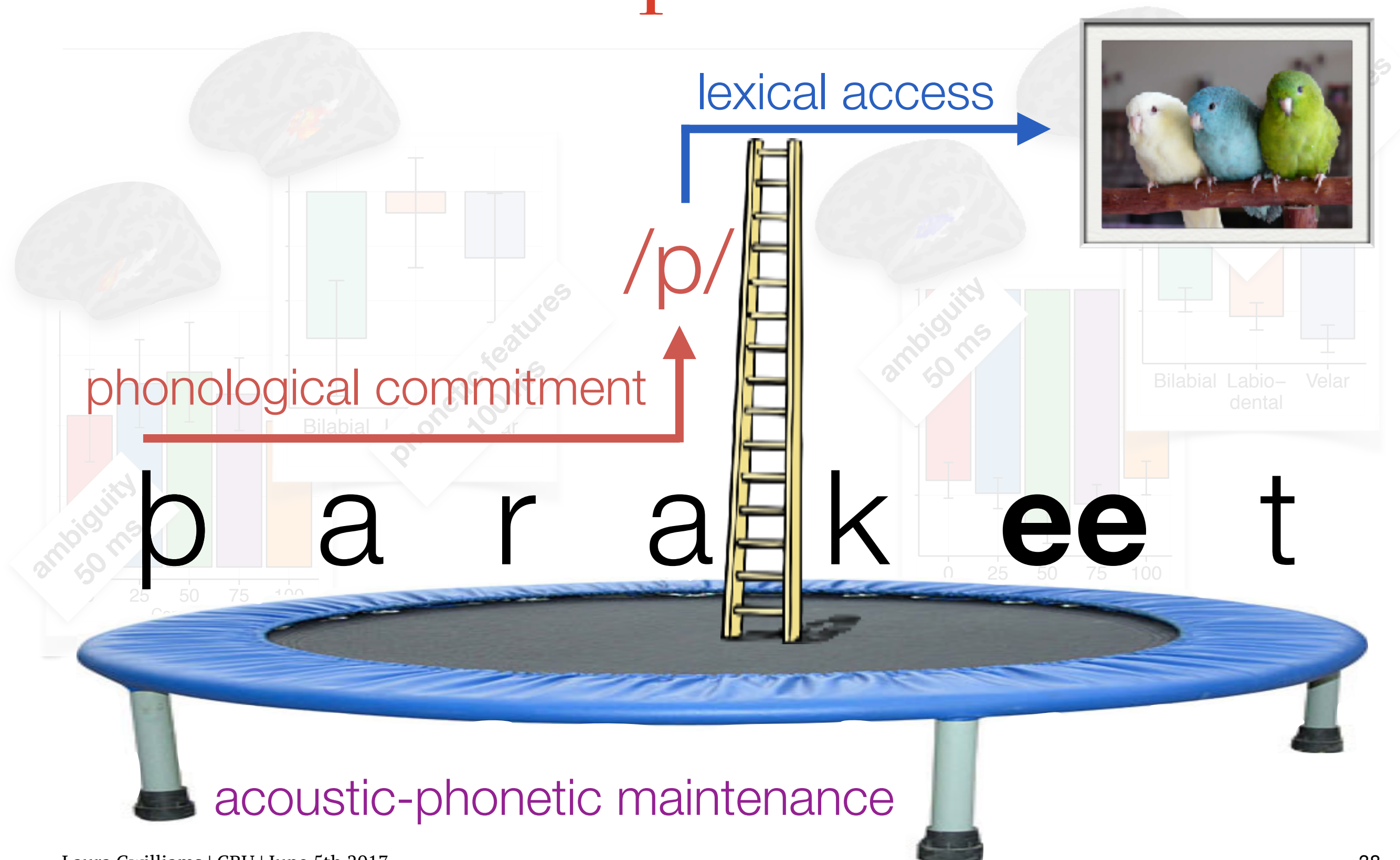
phonological commitment

/p/

b a r a k e e t

acoustic-phonetic maintenance

Interpretation



Interpretation

Processing is not purely feedforward, or feed “up”: TRACE model: McClelland and Elman, 1986; McMurray et al. 2009. cf. MERGE: Norris et al. 2000

lexical access



phonological commitment

/p/

b a r a k e e t

acoustic-phonetic maintenance

Interpretation

lexical access

phonological commitment

/p/

b a r a k e e t

acoustic-phonetic maintenance

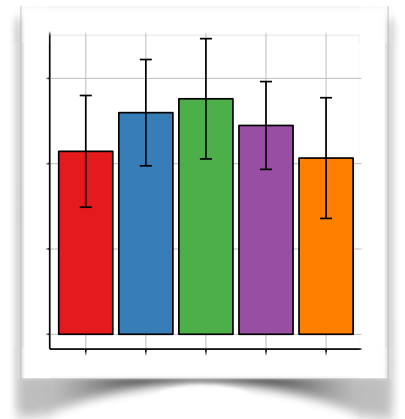


Analogies to other sources of ambiguity, such as homophone resolution: Twilley and Dixon, 2000; Rodd et al., 2010, Rodd, 2017

Conclusion Part 1

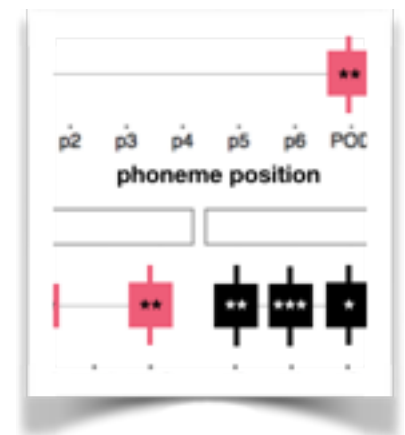
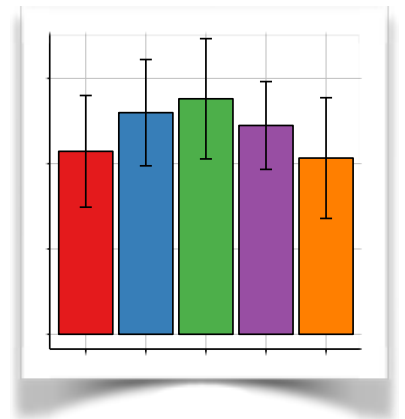
Conclusion Part 1

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



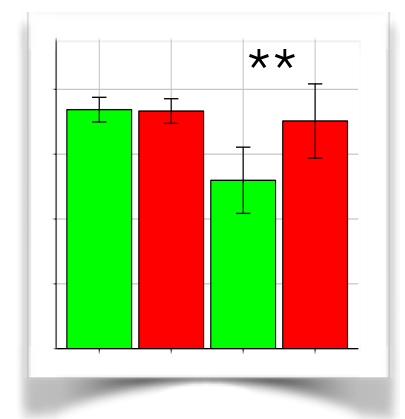
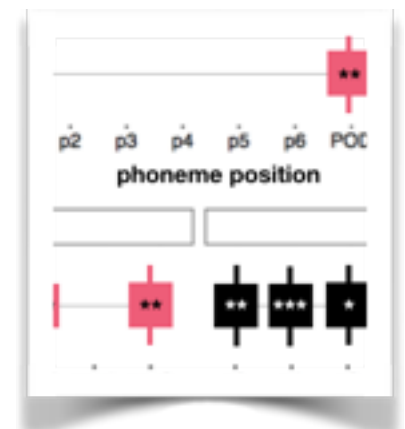
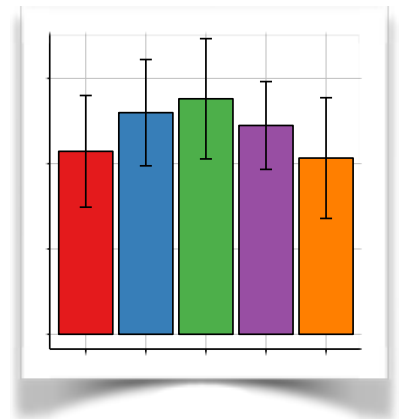
Conclusion Part 1

- 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



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

Research Question

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

Collaborator

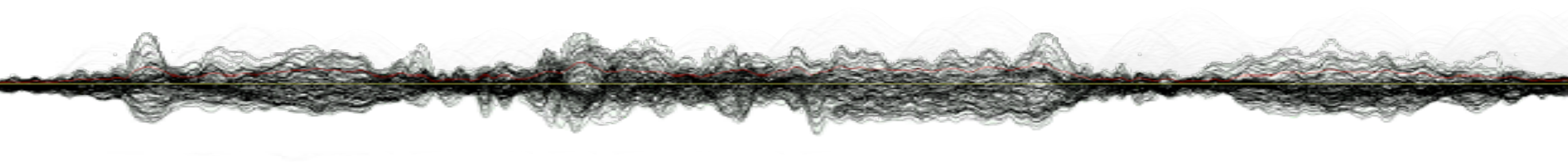


ME

JEAN-RÉMI KING

Setup

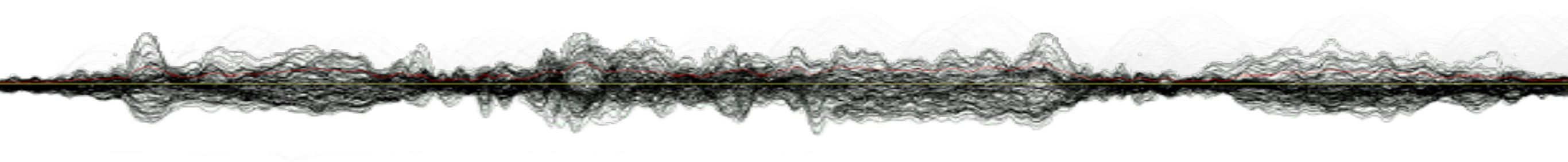
we take continuous speech, and annotate it for phoneme boundaries and phonetic information



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

Setup

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



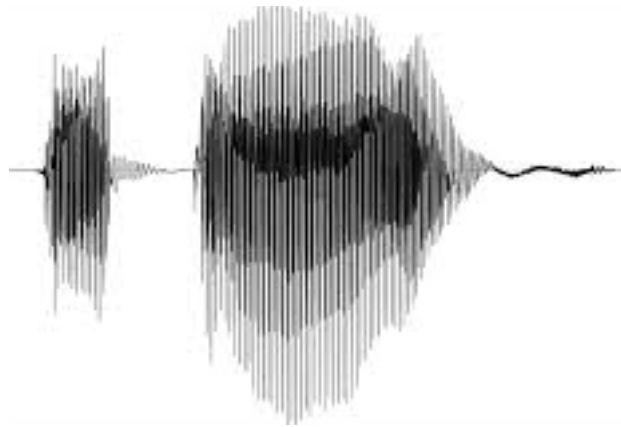
- 24 participants
- 1 hour recording
- ~40,000 phonemes per participant

Decoding from the MEG Signal

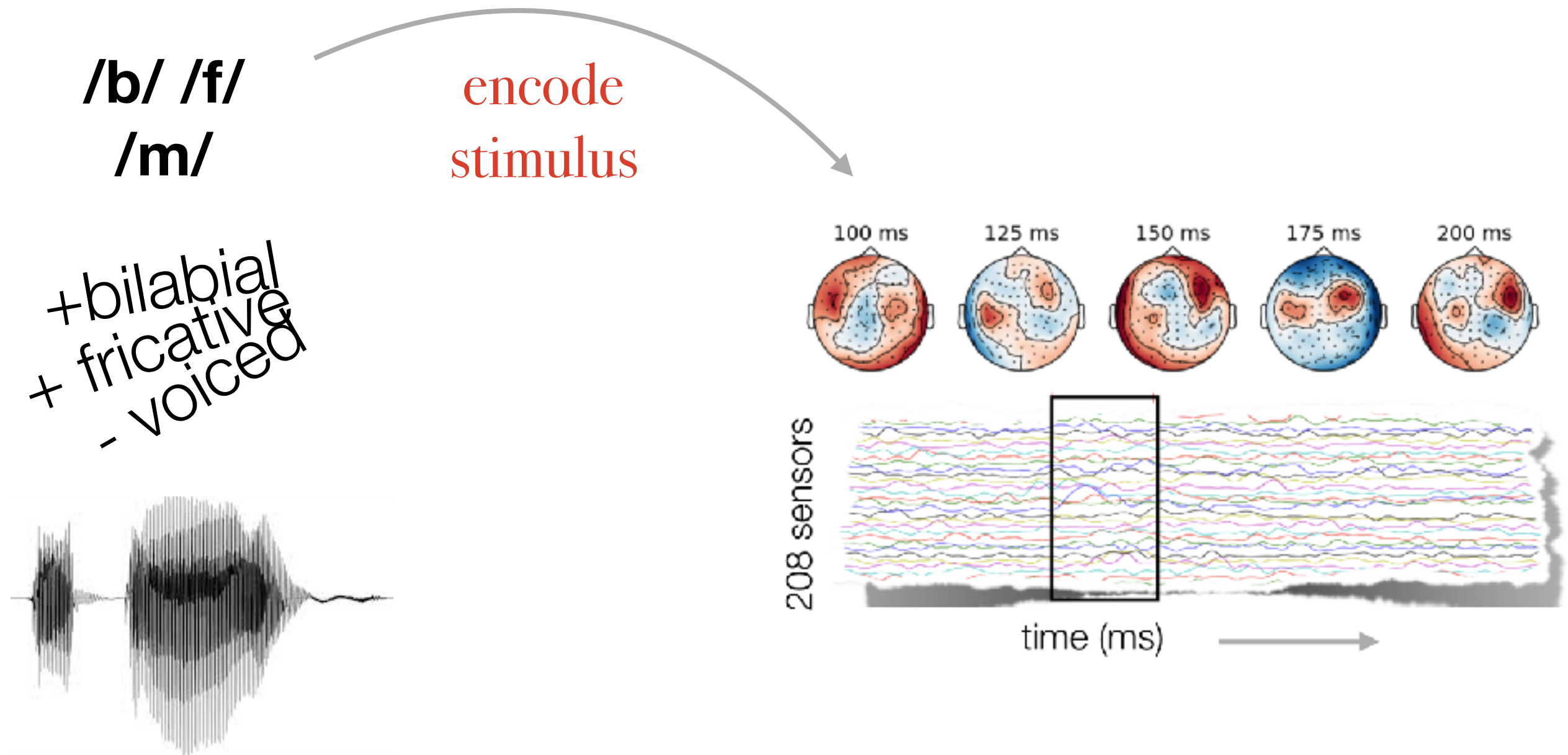
Decoding from the MEG Signal

**/b/ /f/
/m/**

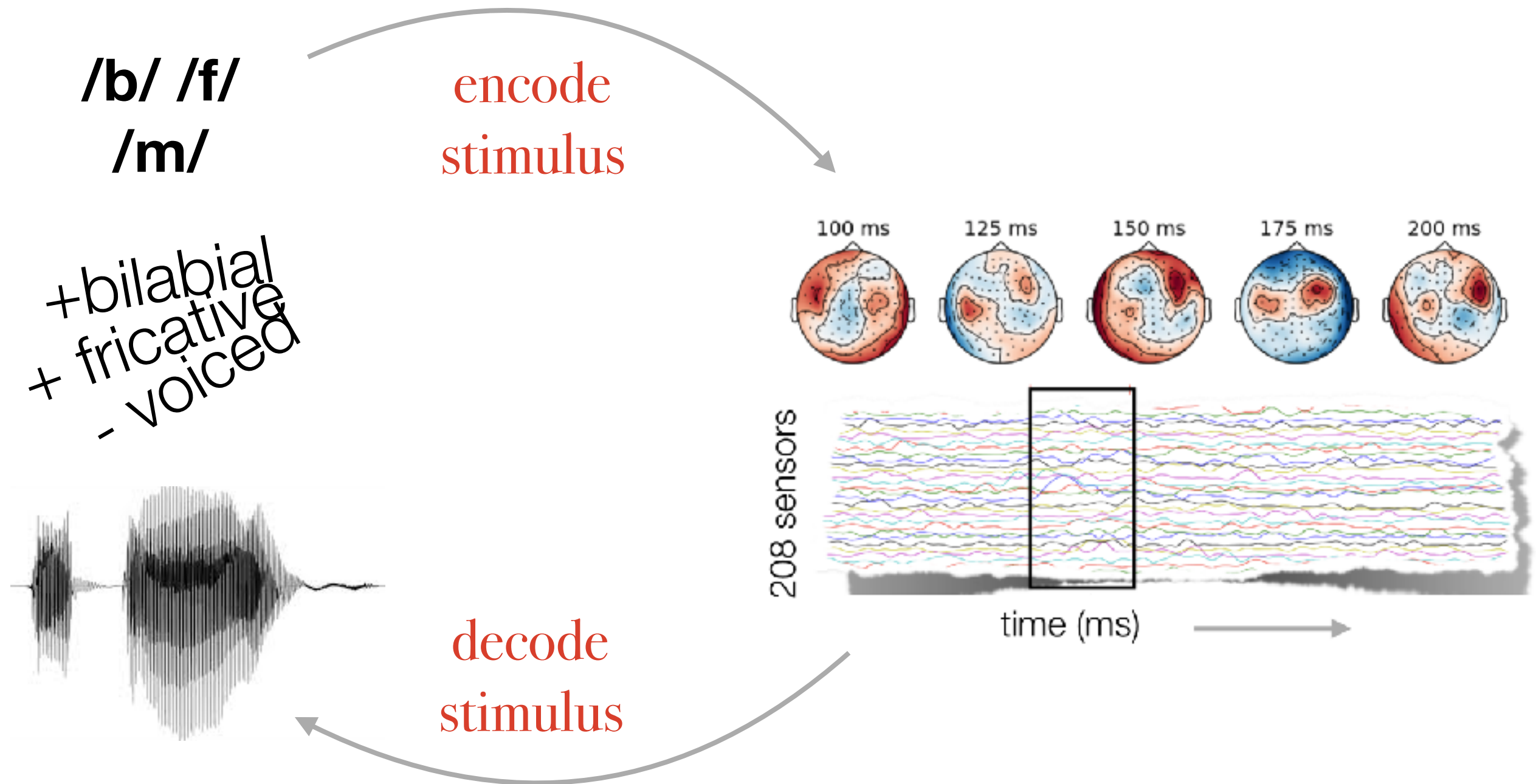
+bilabial
+fricative
-voiced



Decoding from the MEG Signal

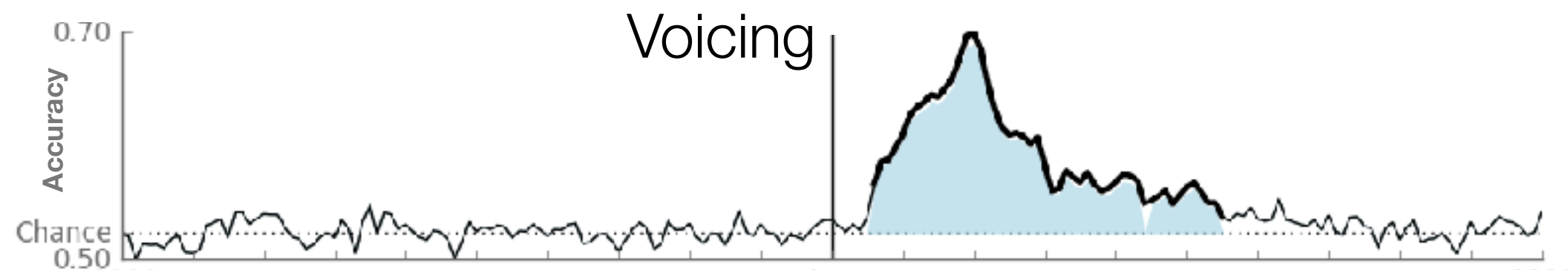


Decoding from the MEG Signal

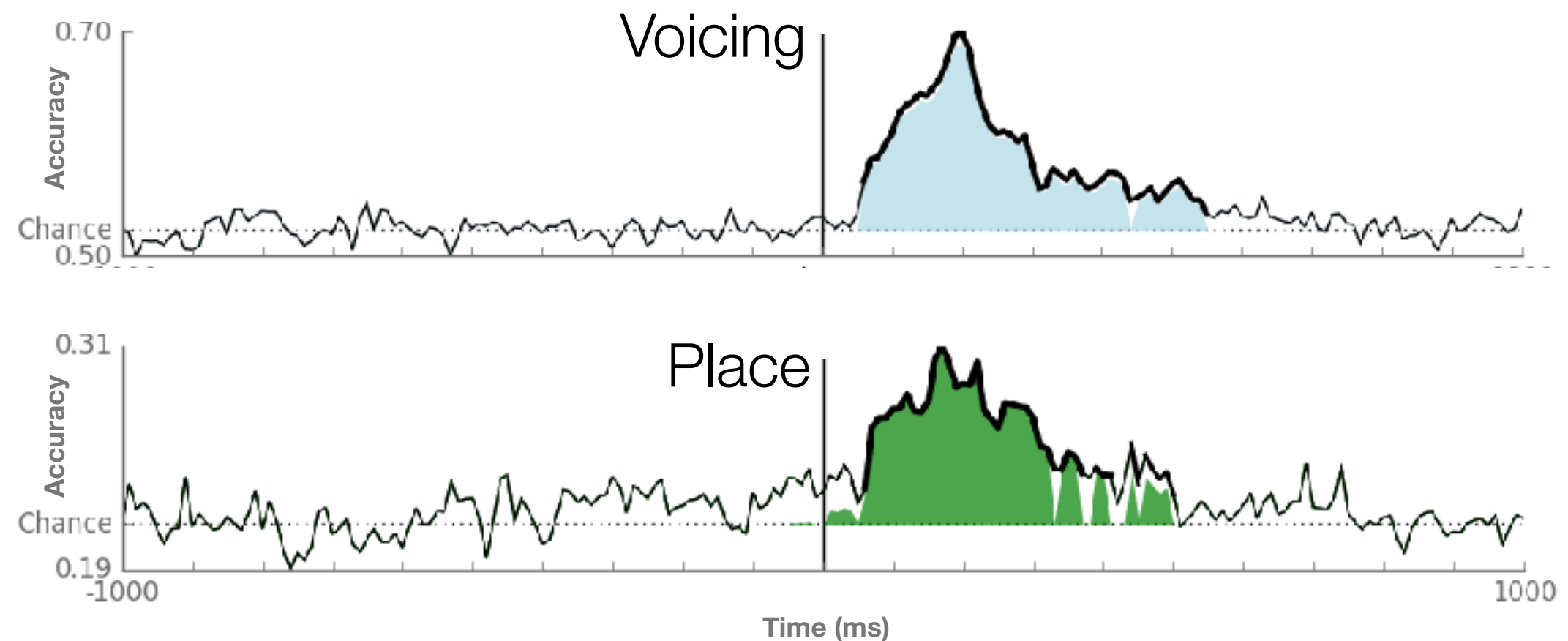


Decode Phonetic Features from the MEG Signal

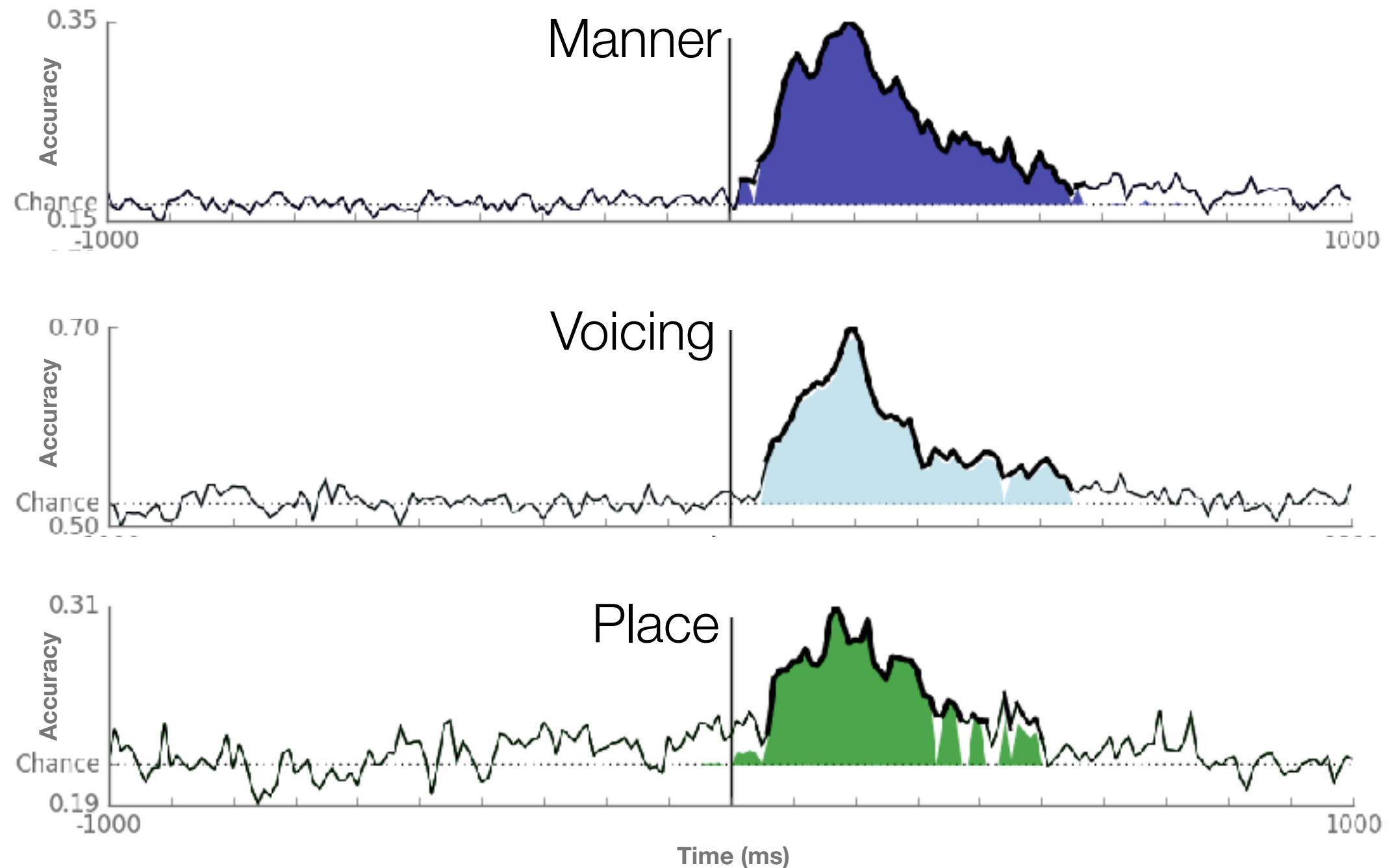
Decode Phonetic Features from the MEG Signal



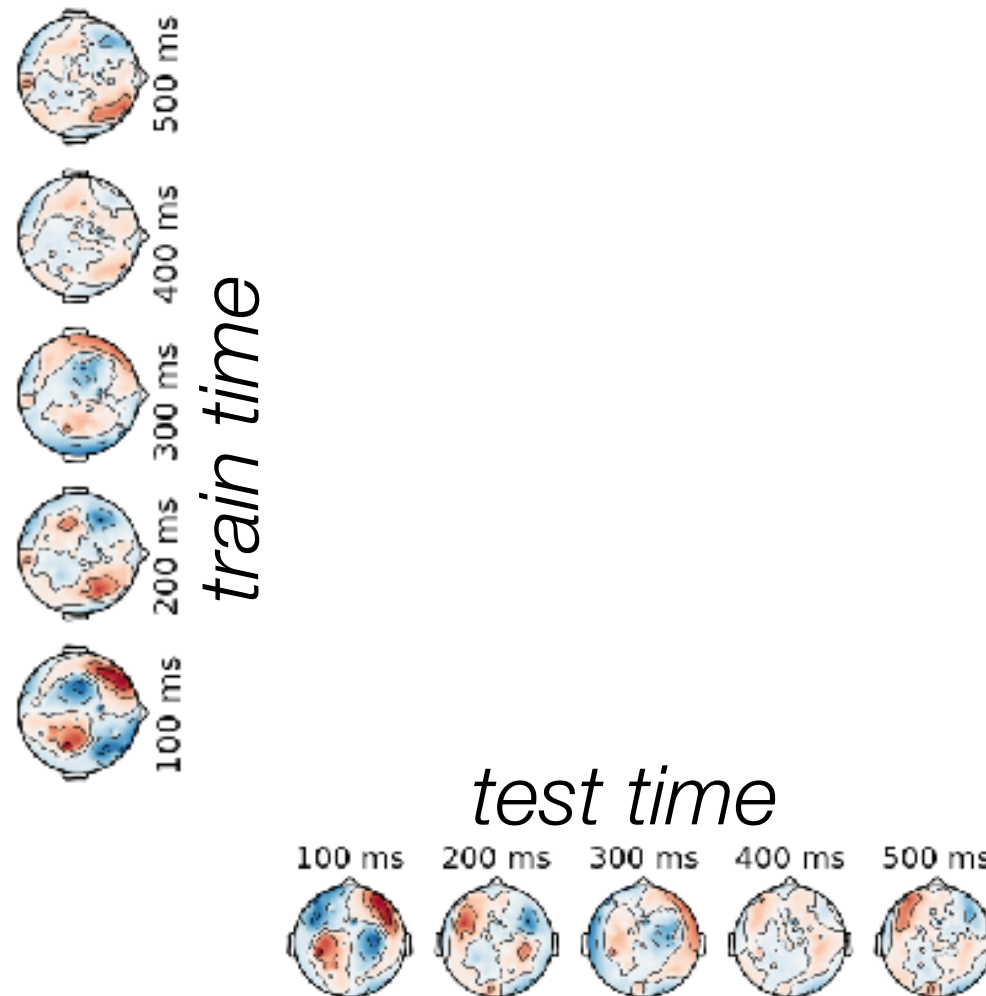
Decode Phonetic Features from the MEG Signal



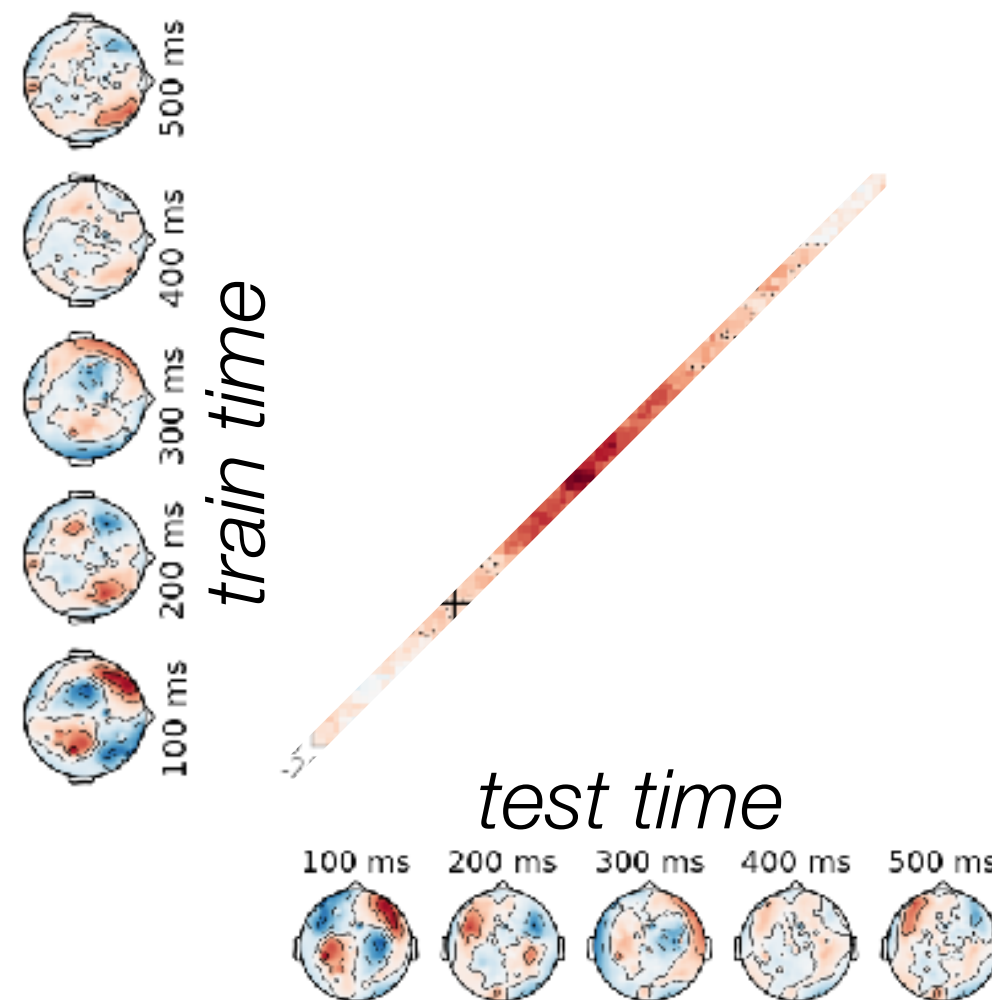
Decode Phonetic Features from the MEG Signal



Decode the *Dynamics* of Phonetic Feature Processing

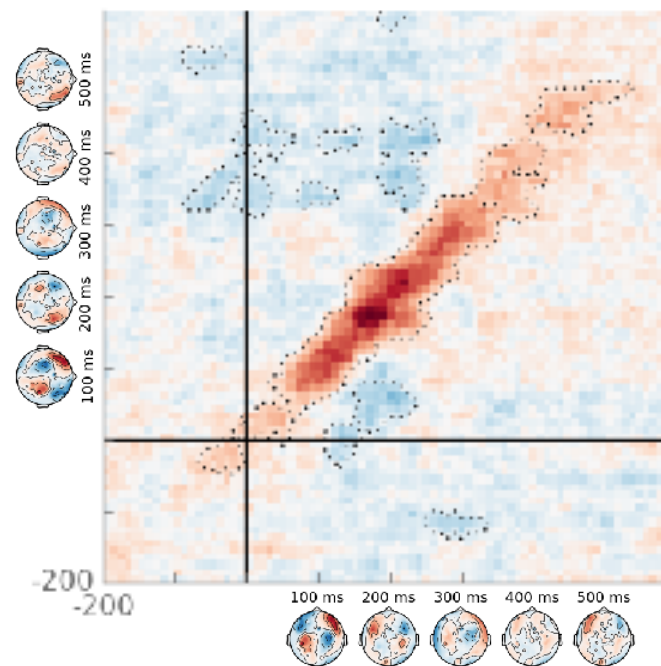


Decode the *Dynamics* of Phonetic Feature Processing



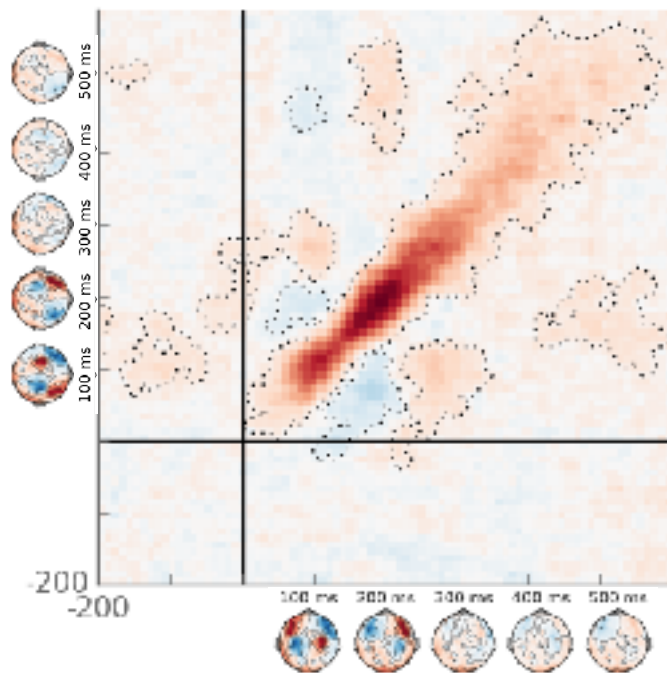
Decode the *Dynamics* of Phonetic Feature Processing

Place

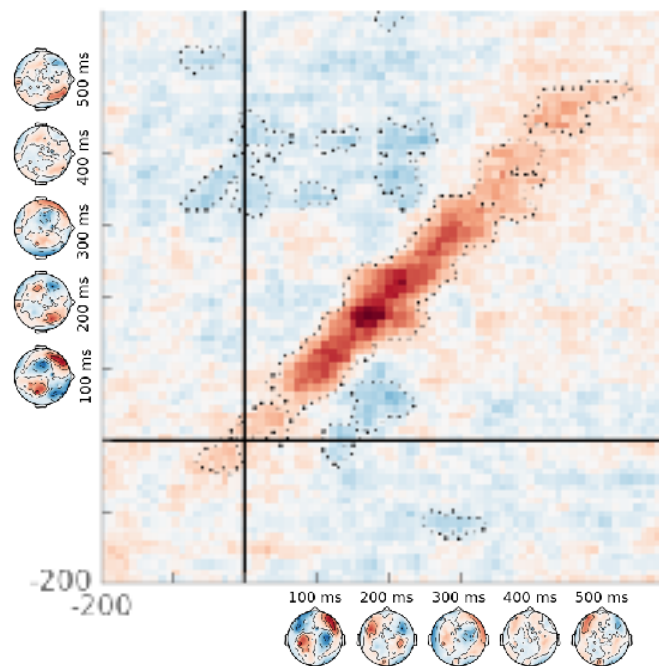


Decode the *Dynamics* of Phonetic Feature Processing

Manner

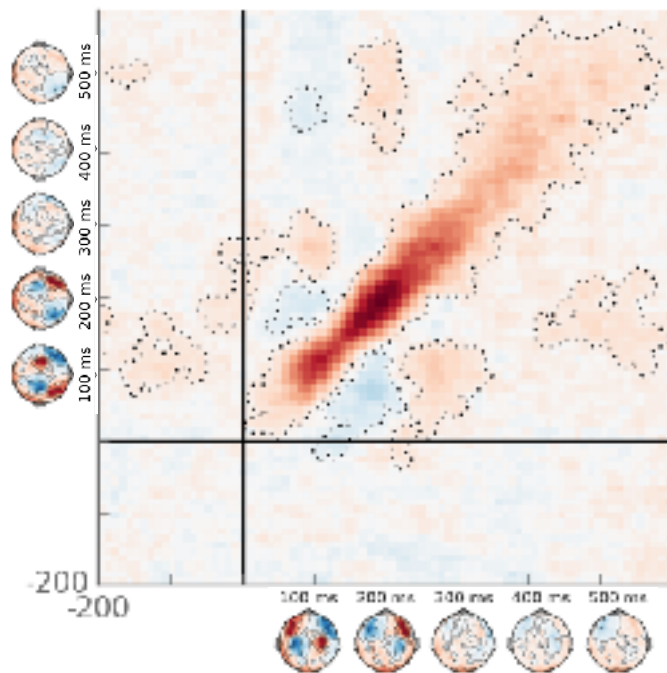


Place

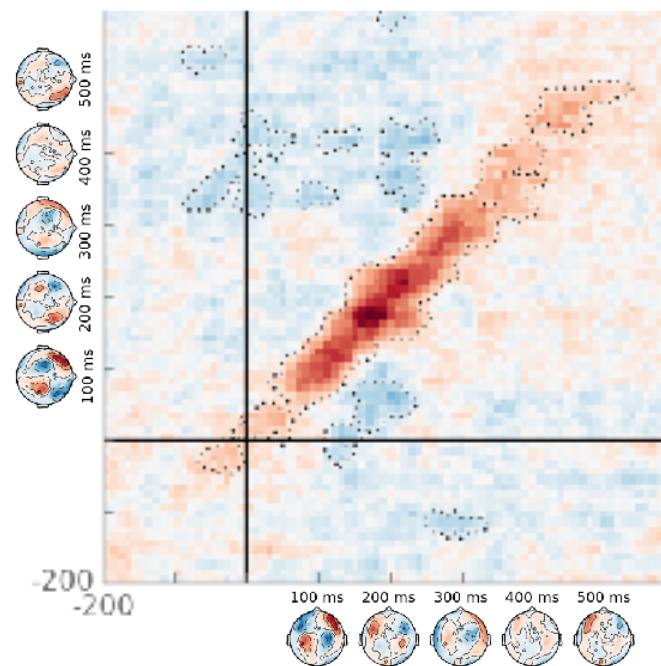


Decode the *Dynamics* of Phonetic Feature Processing

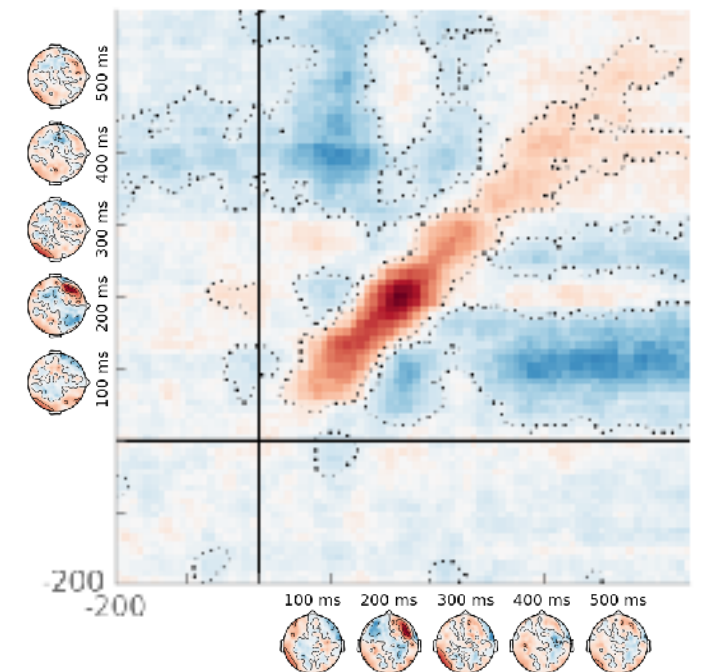
Manner



Place



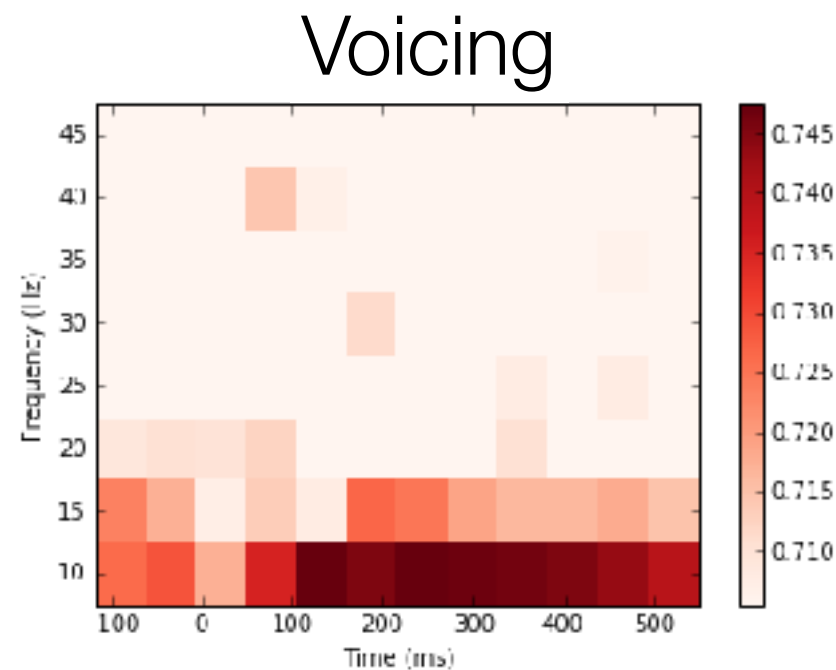
Voicing



Decode the *Frequency* of Phonetic Feature Processing

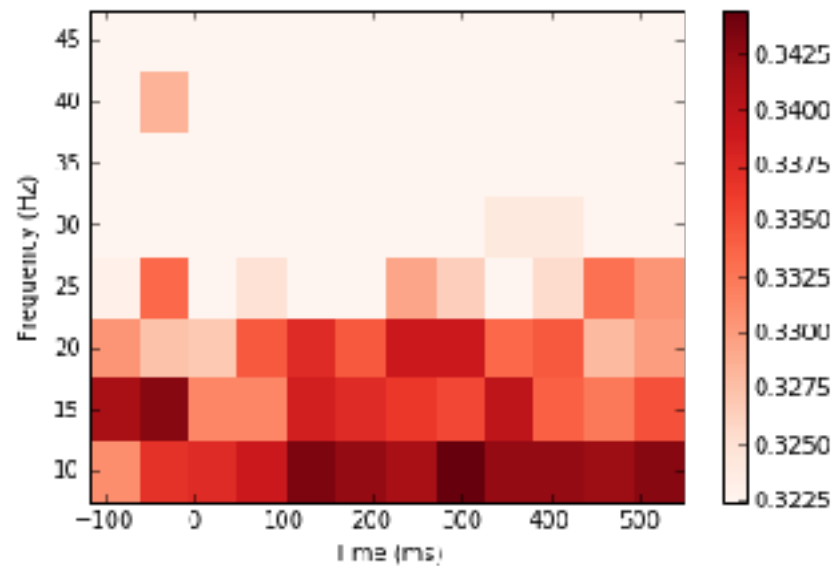


Decode the *Frequency* of Phonetic Feature Processing

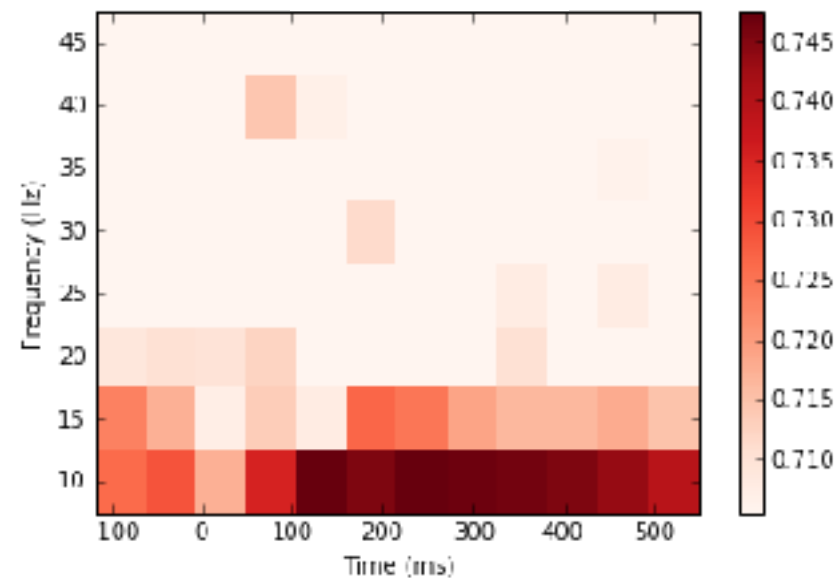


Decode the *Frequency* of Phonetic Feature Processing

Manner

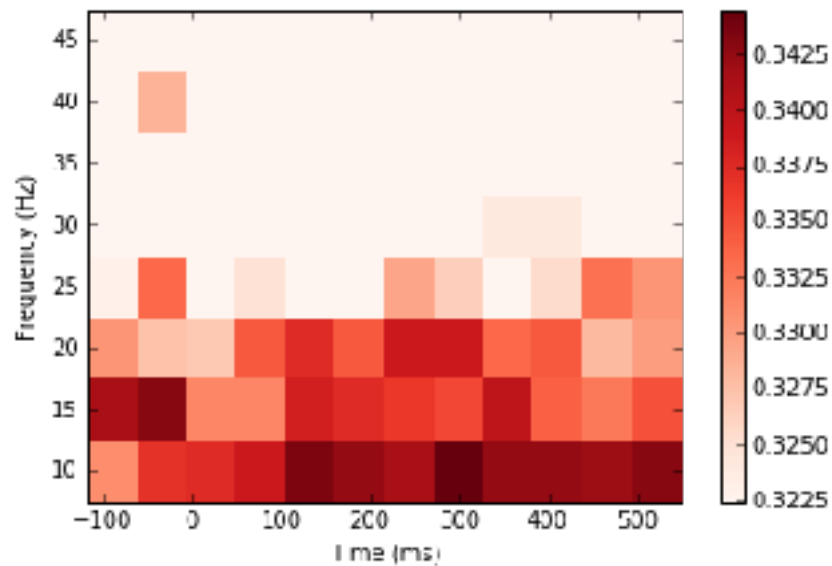


Voicing

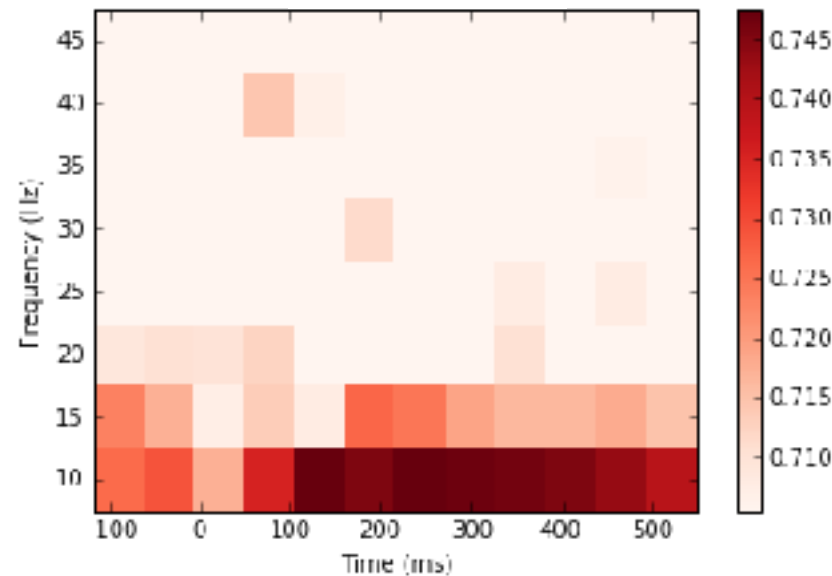


Decode the *Frequency* of Phonetic Feature Processing

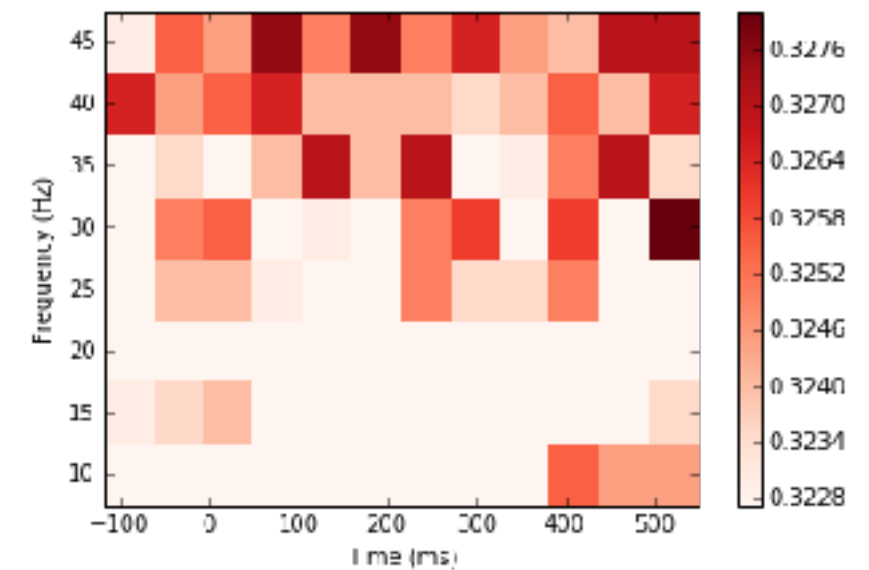
Manner



Voicing



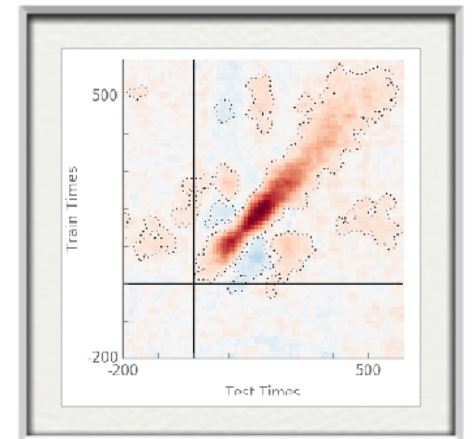
Place



Promising Work in Progress

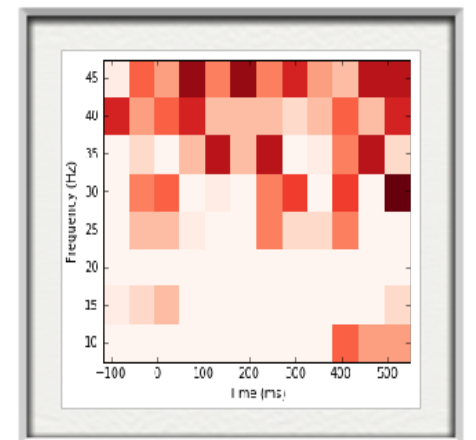
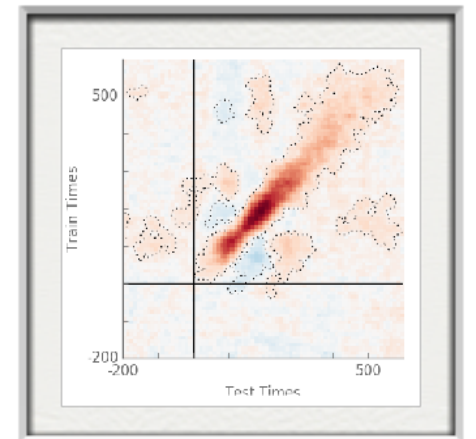
Promising Work in Progress

- Phonetic features appear to elicit different temporal dynamics



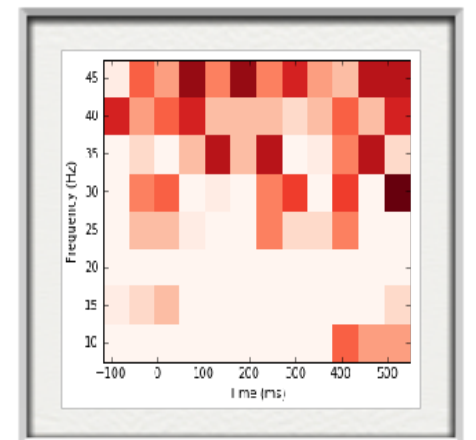
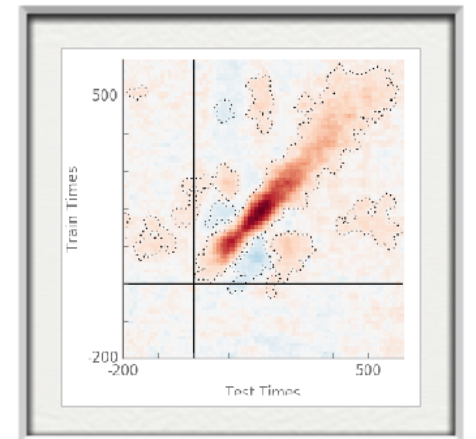
Promising Work in Progress

- Phonetic features appear to elicit different temporal dynamics
- And different spectral profiles



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

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?

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](https://twitter.com/GwilliamsL)



✉ laura.gwilliams@nyu.edu
🐦 [@GwilliamsL](https://twitter.com/GwilliamsL)

Thank you!

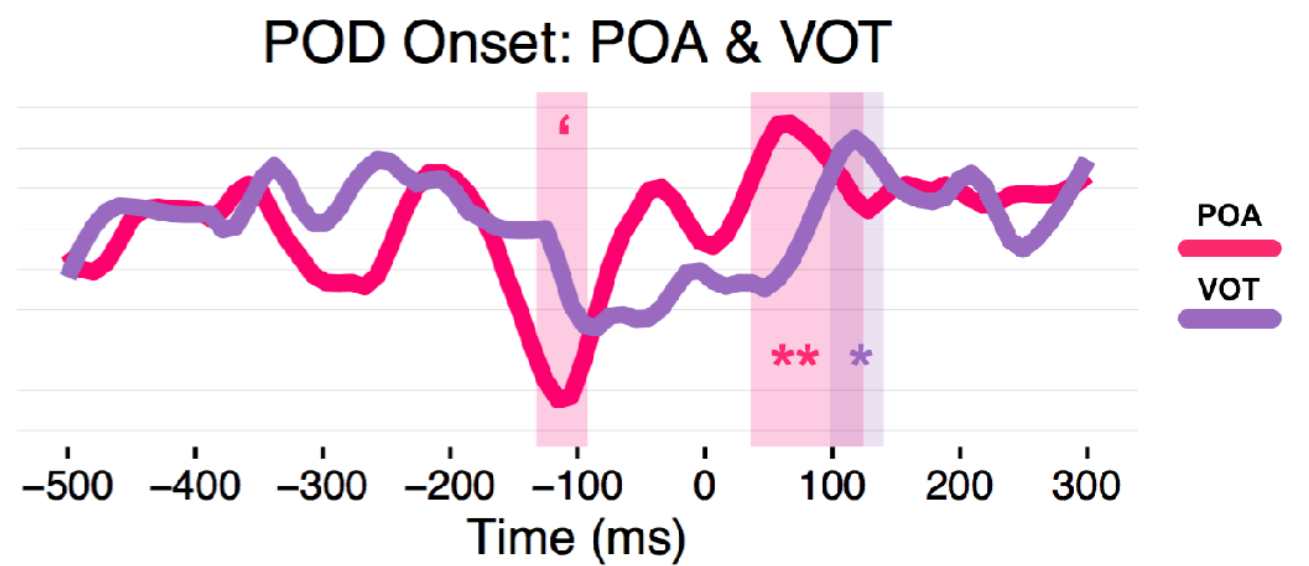
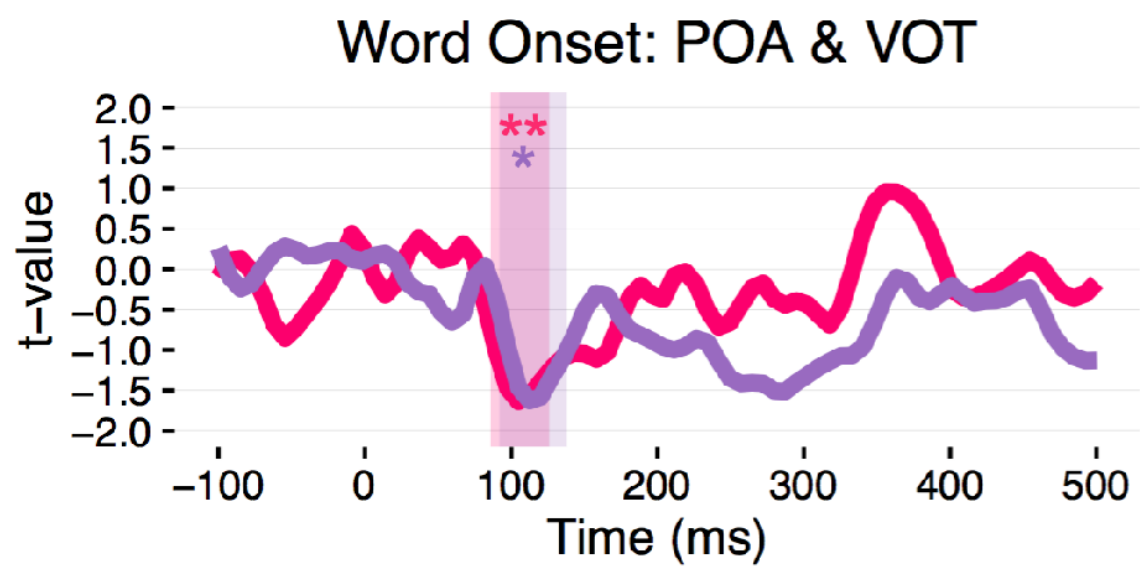
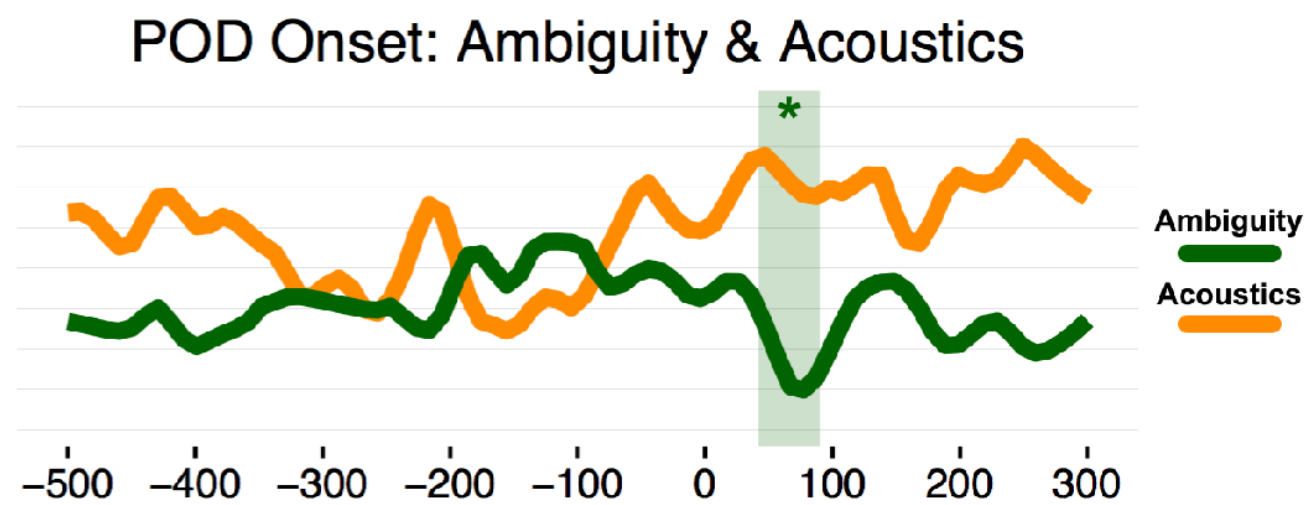
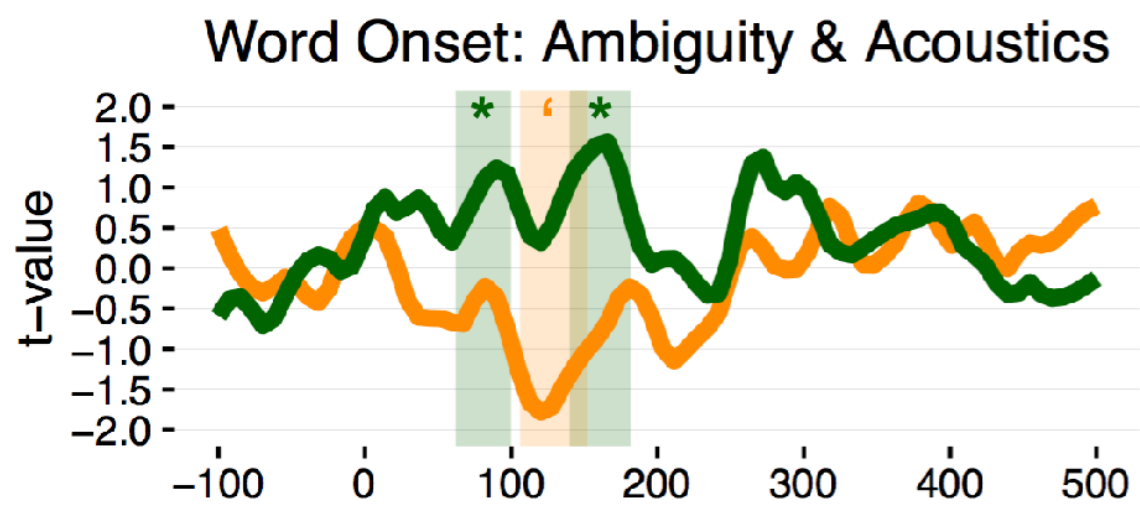


NEW YORK UNIVERSITY

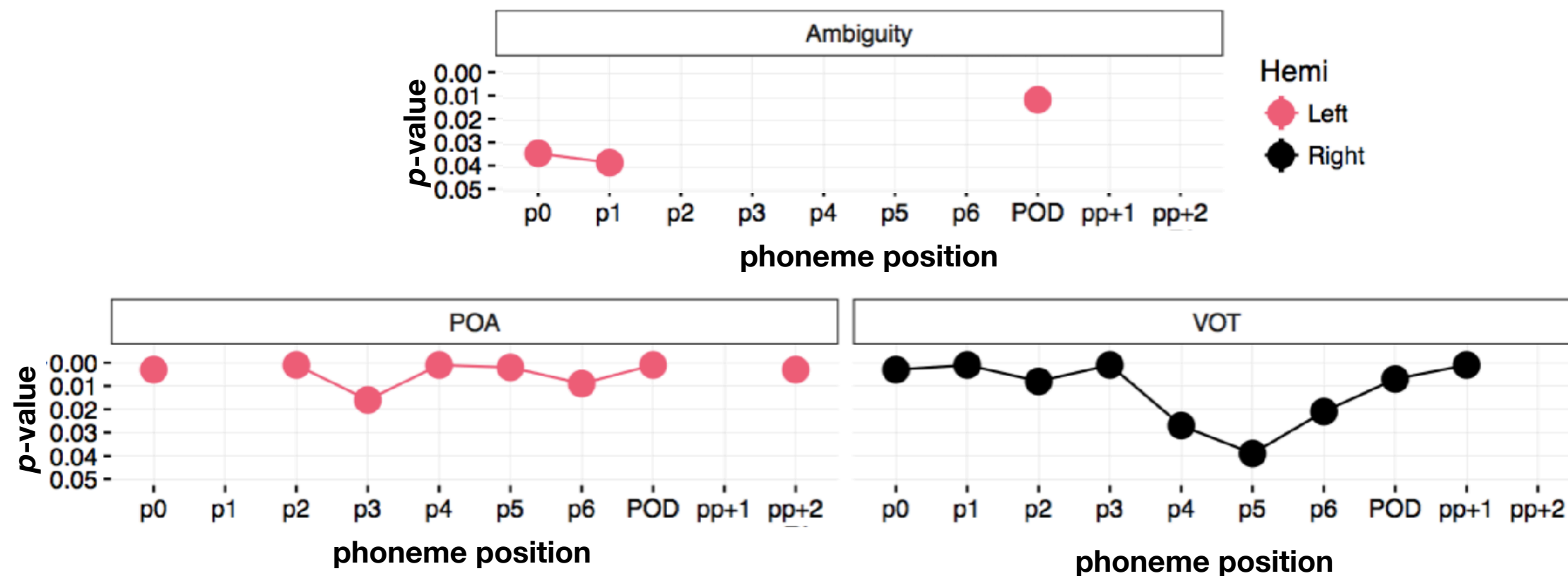
Example Stimuli Pairs

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

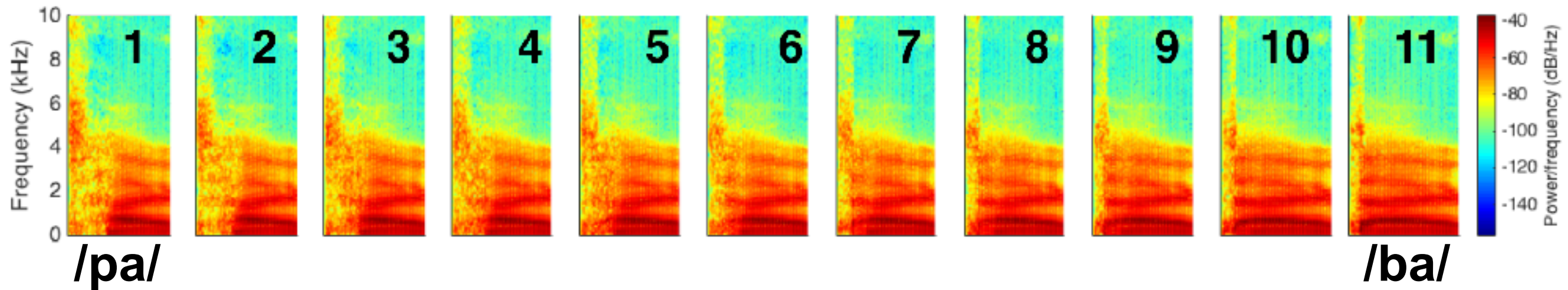


Reactivation in Intermediate Positions

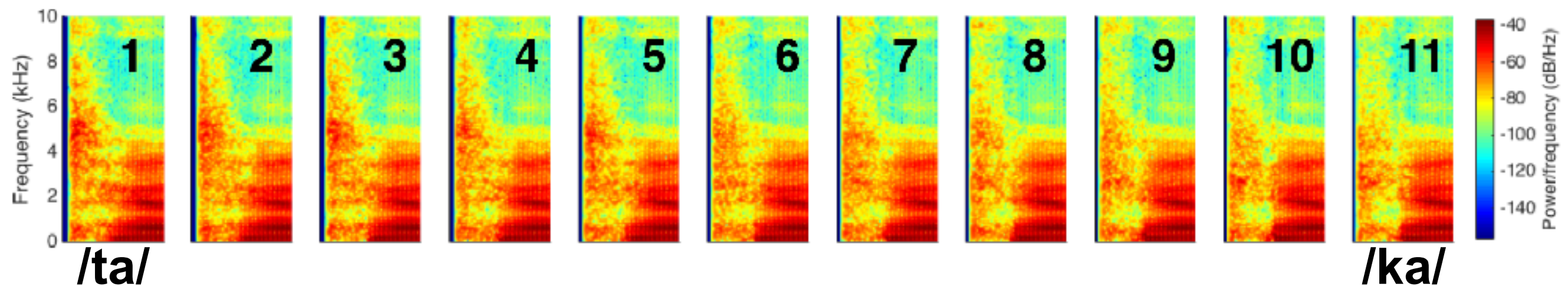


Experiment 1: Design & Materials

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

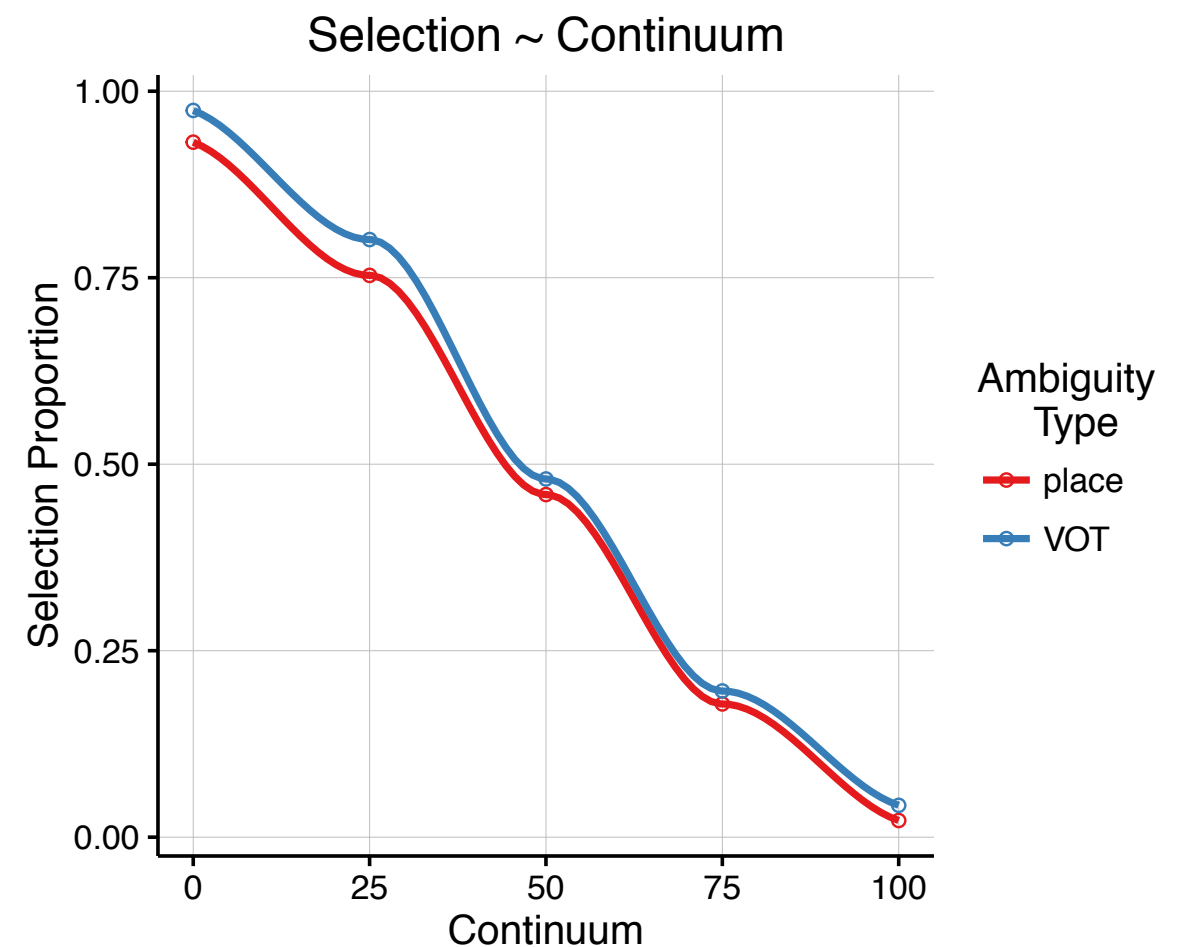
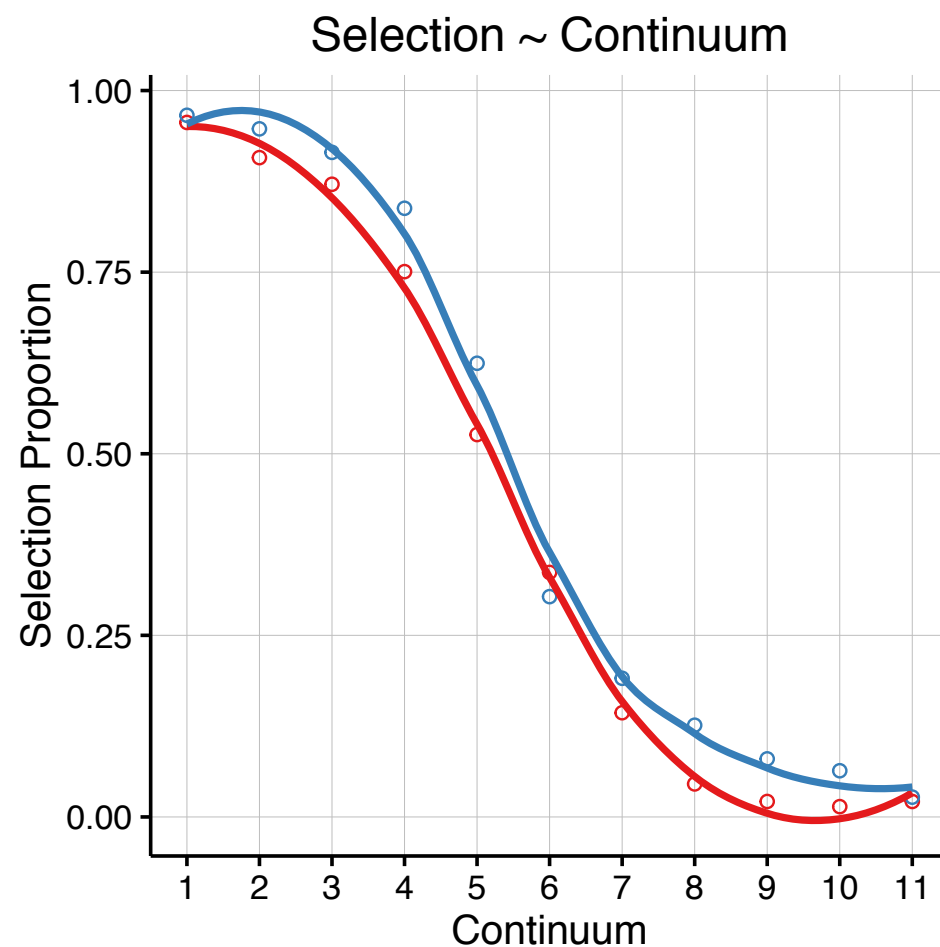


- Place of articulation (PoA) - {t-k, p-t}

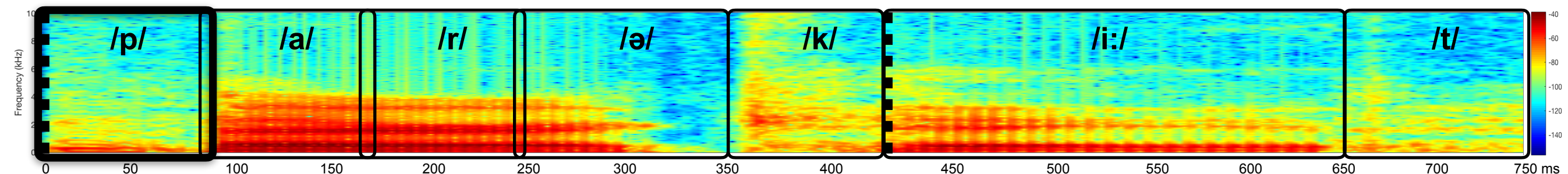
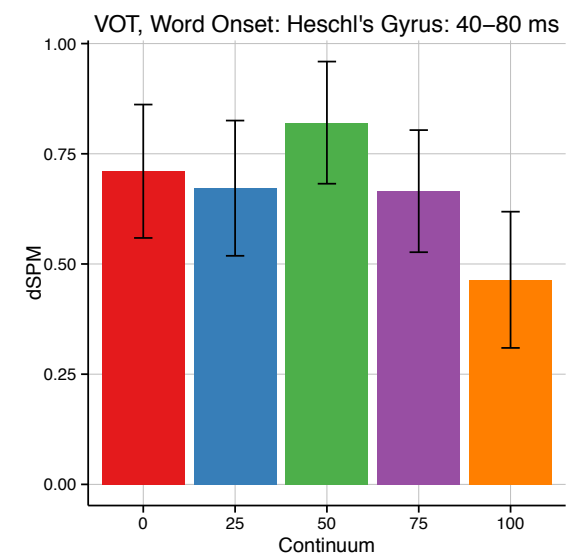
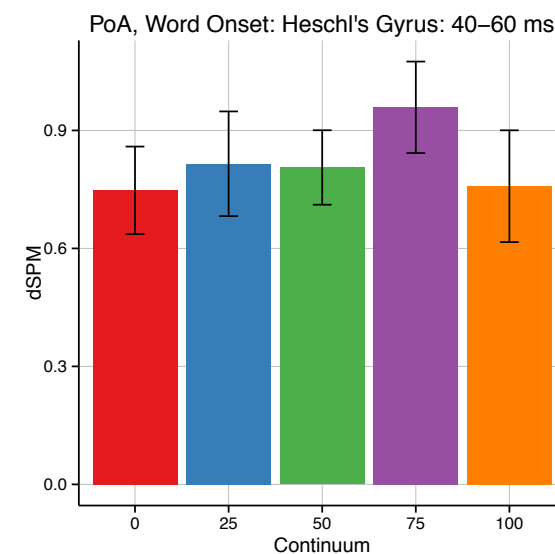
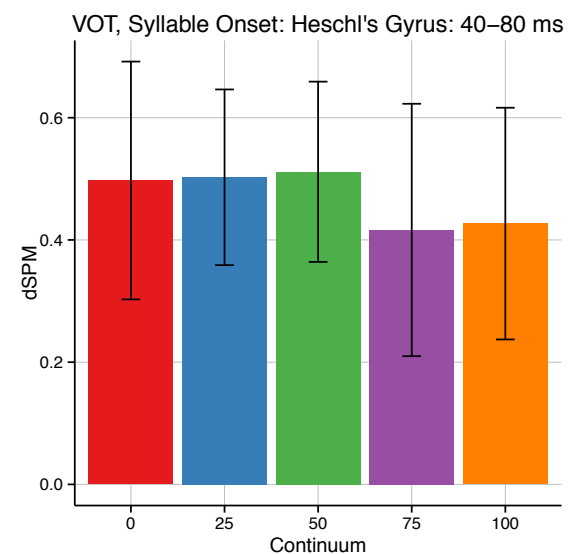
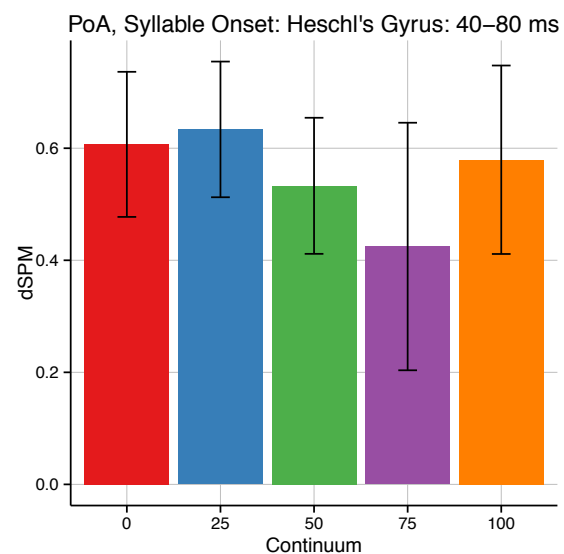
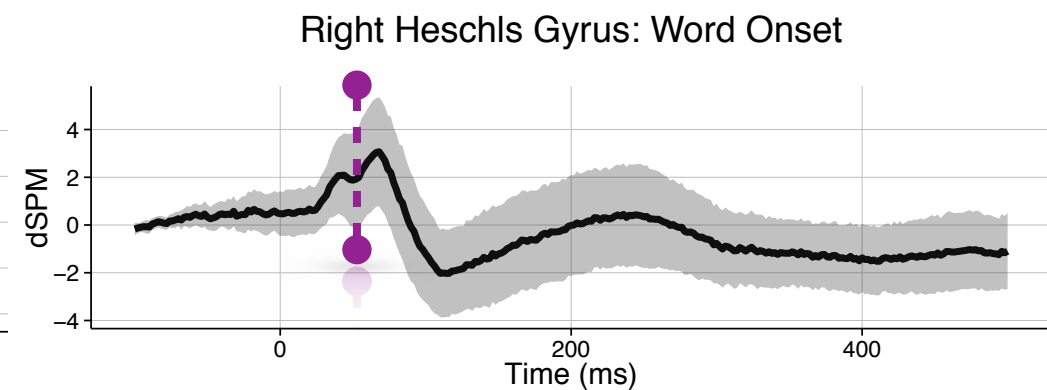
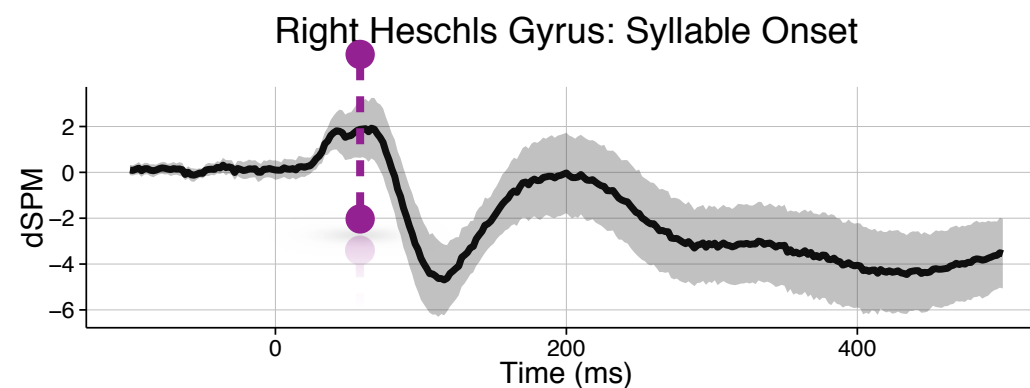


Experiment 1: Design & Materials

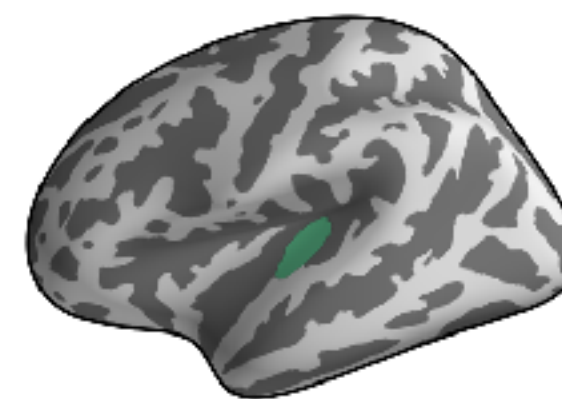
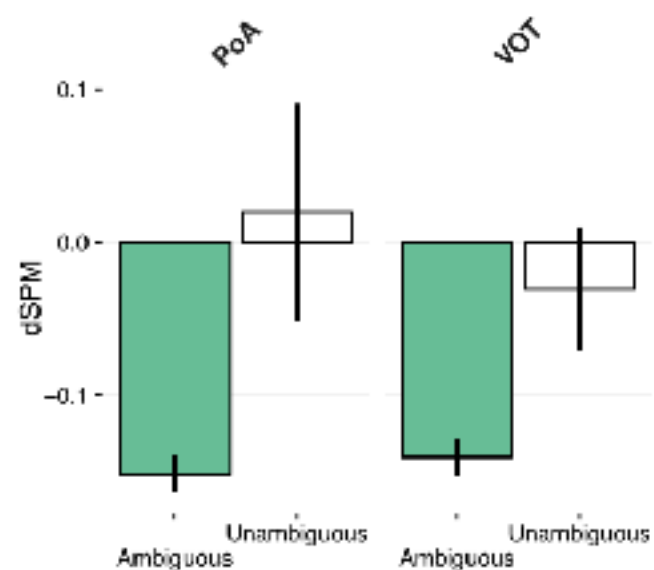
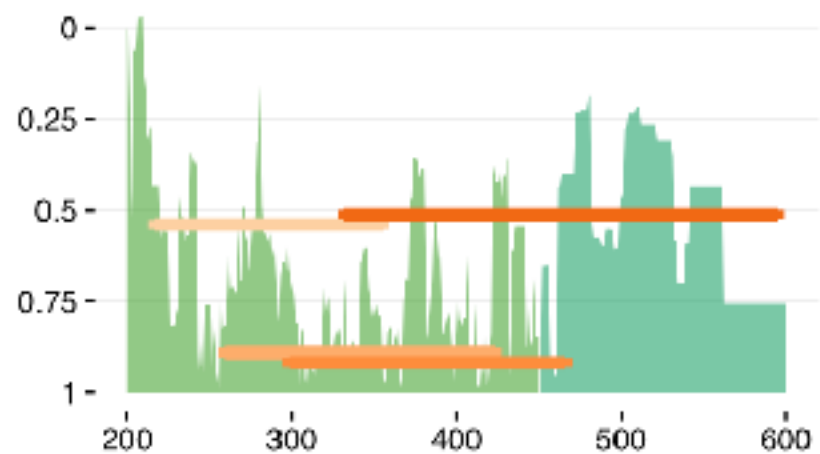
- Re-sampled the continuum to match perceptual categorisation



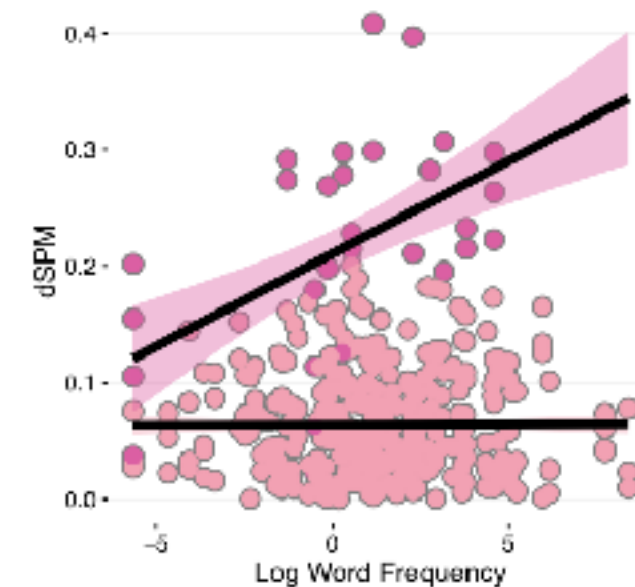
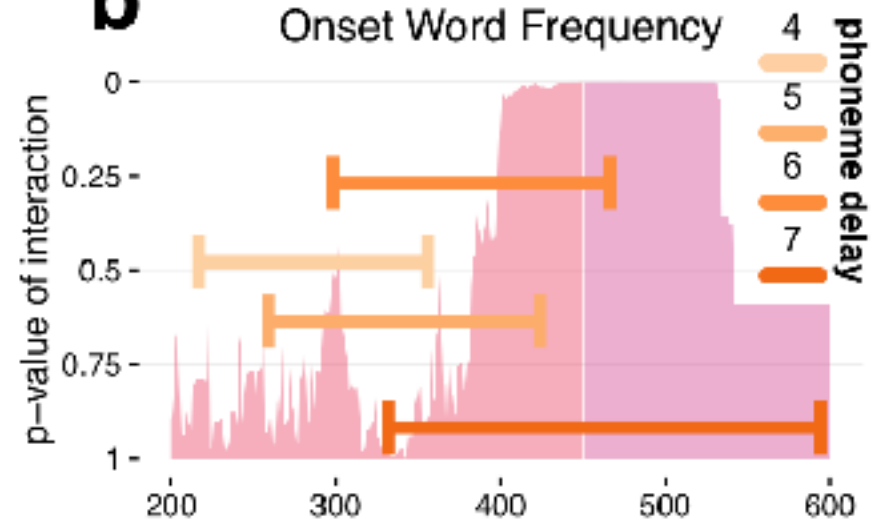
No Ambiguity Effect in Right Hemisphere



a Phoneme Ambiguity



b Onset Word Frequency



c Lexical Resolution

