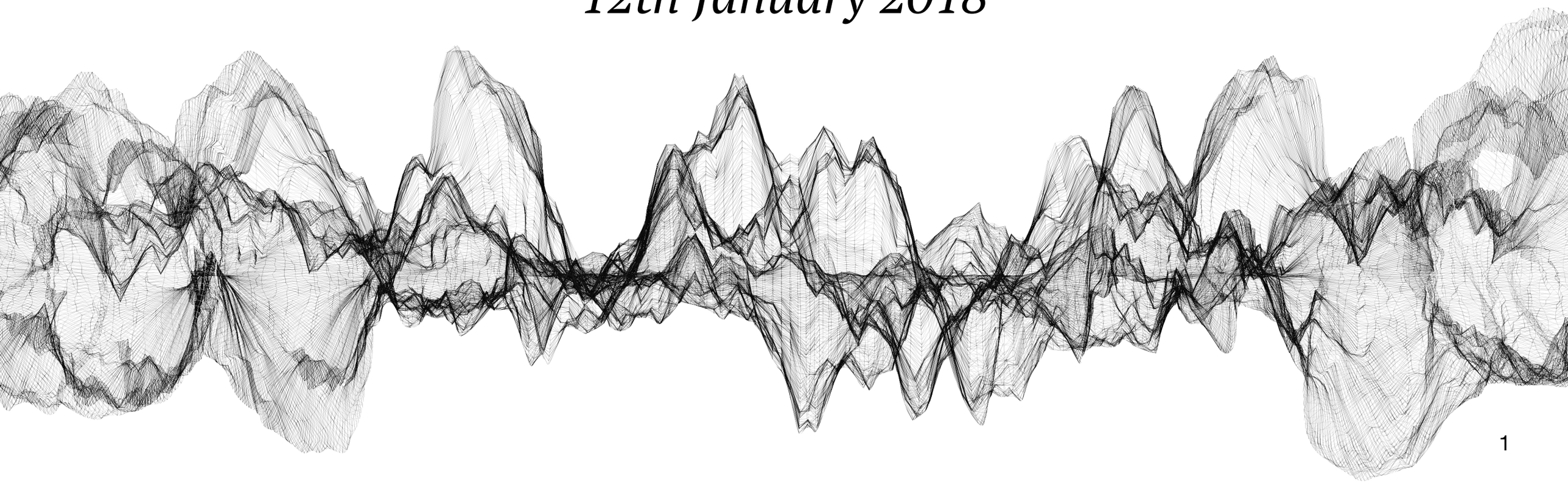
NEW YORK UNIVERSITY

Postdictive processing in spoken word recognition

Laura Gwilliams

12th January 2018



Collaborators

Tal Linzen

David Poeppel

Alec Marantz

Gwilliams, Poeppel, Marantz & Linzen (2018)
Gwilliams, Linzen, Poeppel & Marantz (Submitted)

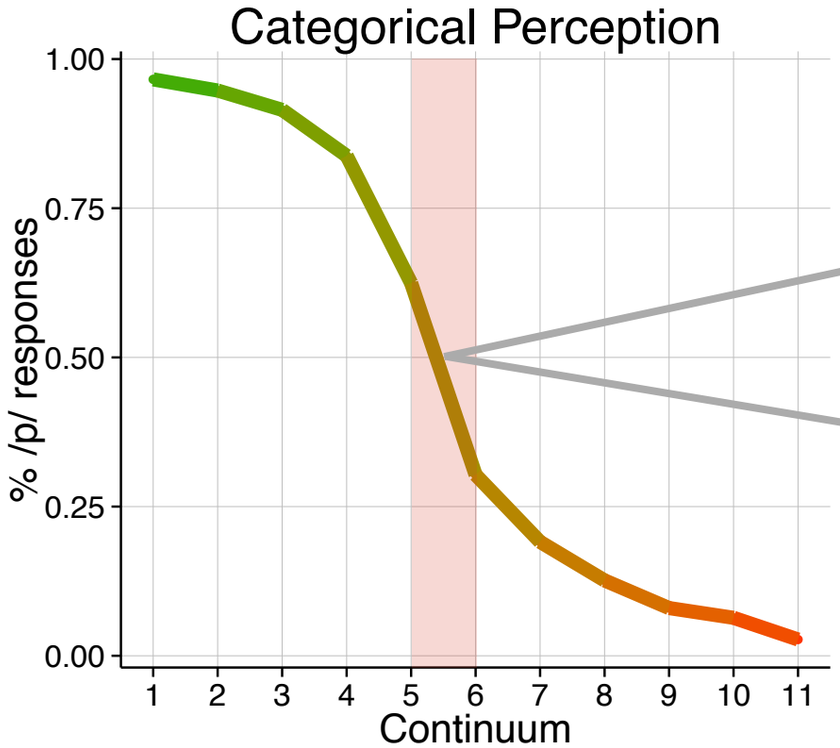
Future Influences on Perception

- 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 **ee** t

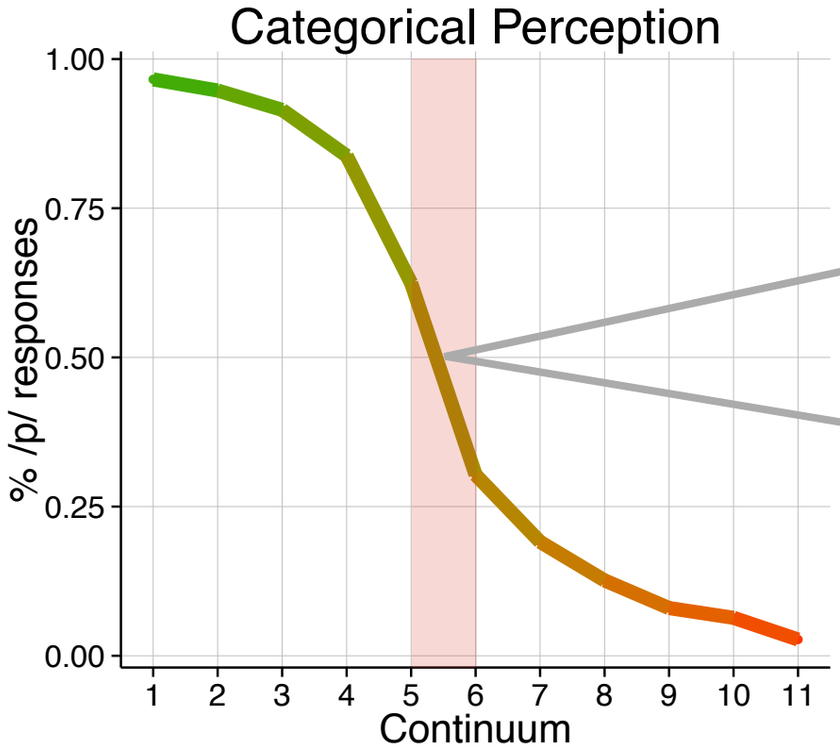
b **ai** d

(this is a parakeet)

(this is a barricade)

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

Future Influences on Perception



p **p** **p** **p** **b**

p a r a k **ee** t

b a r a k **ai** d

(this is a parakeet)

(this is a barricade)

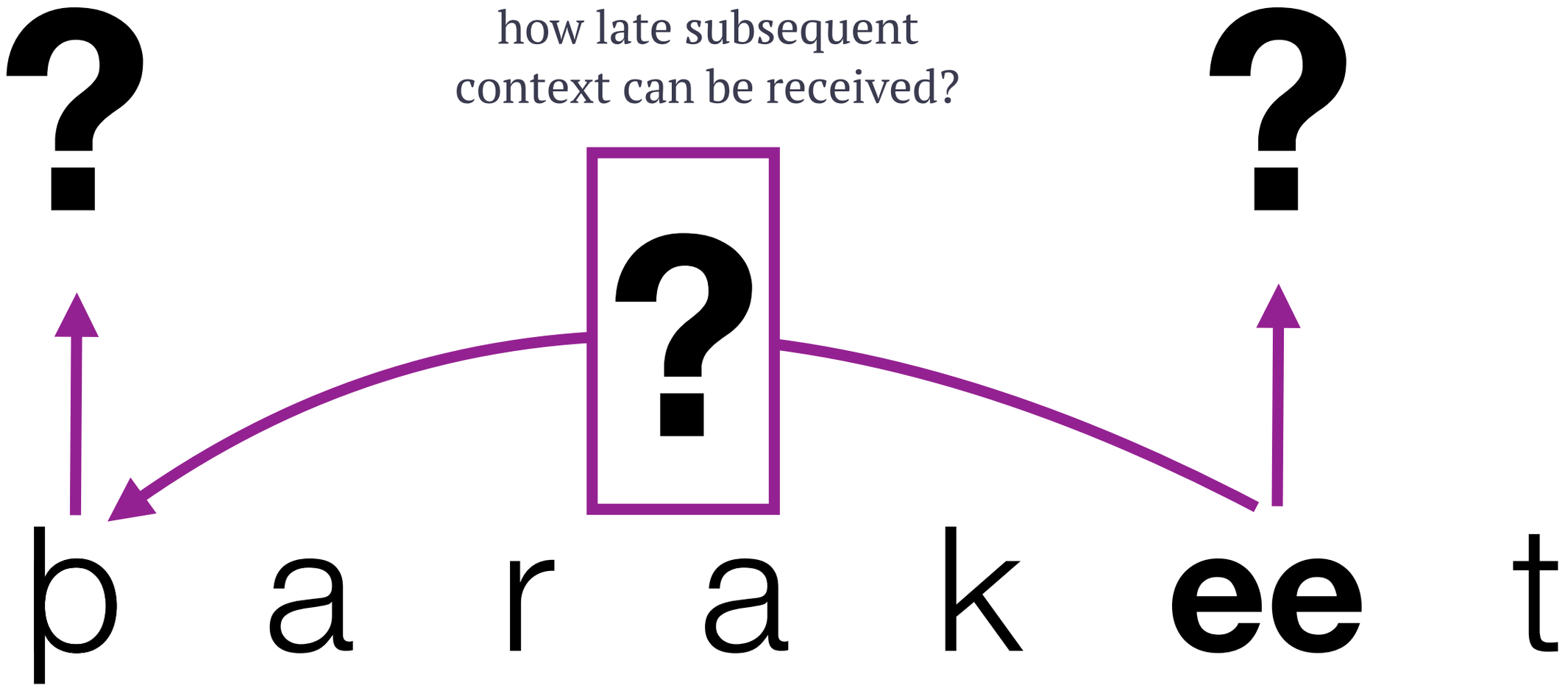
“**P**oint **o**f **D**isambiguation” (POD)

Today's Questions

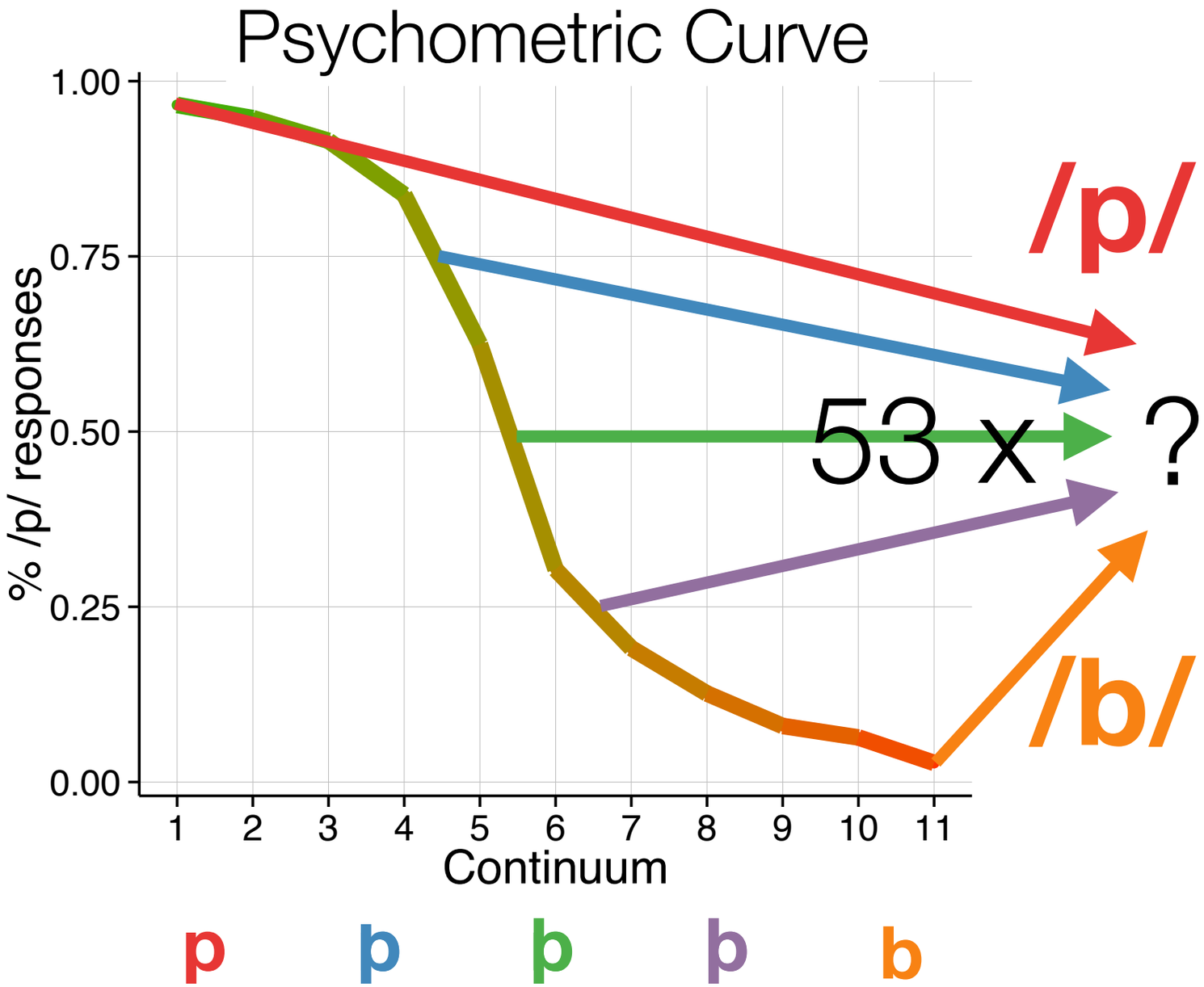
How does the auditory cortex **respond** to phonological ambiguity?

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

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



Design & Materials



/p/

53 x ?

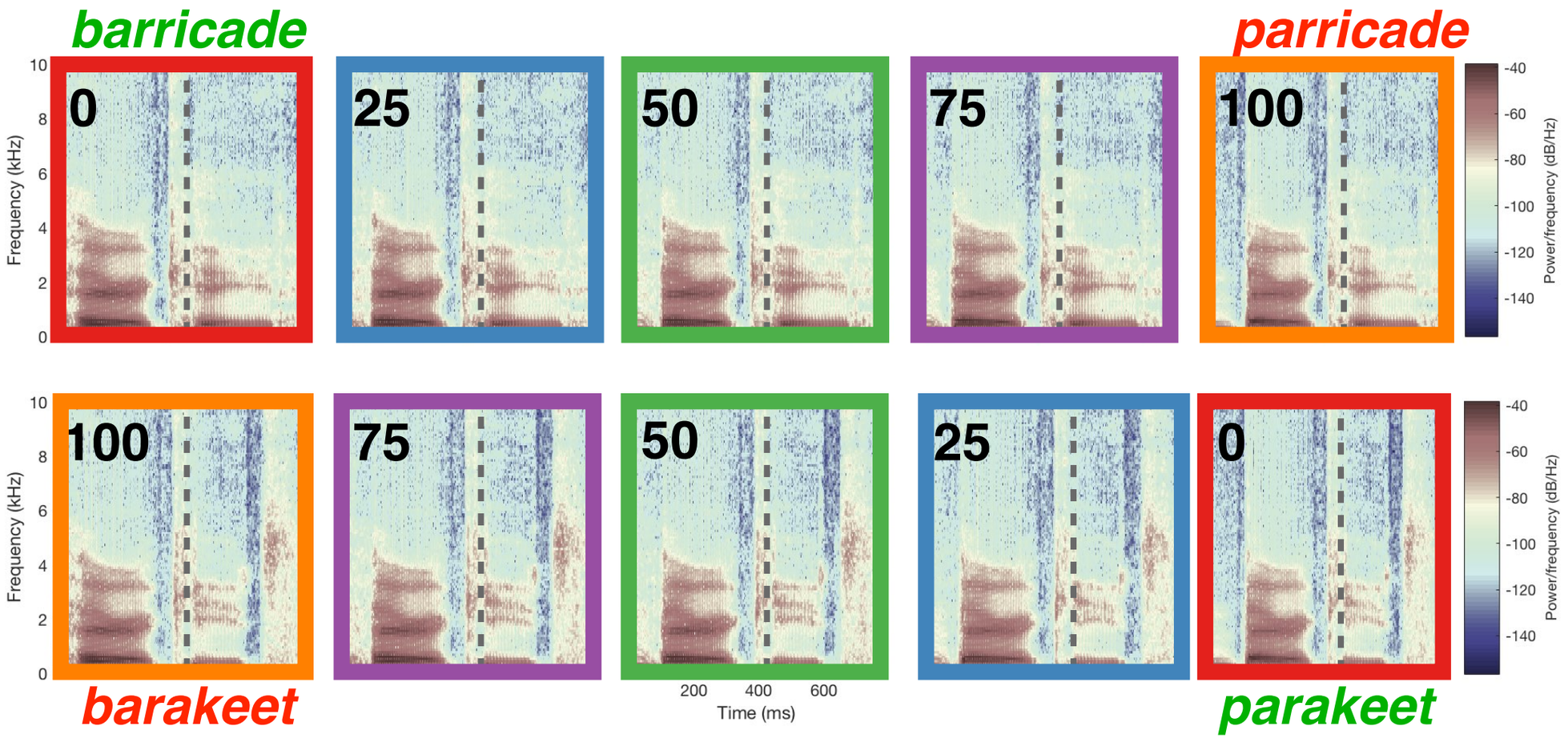
/b/

a r a k **ee** t

a r a k **ai** d

(see also McMurray et al., 2009)

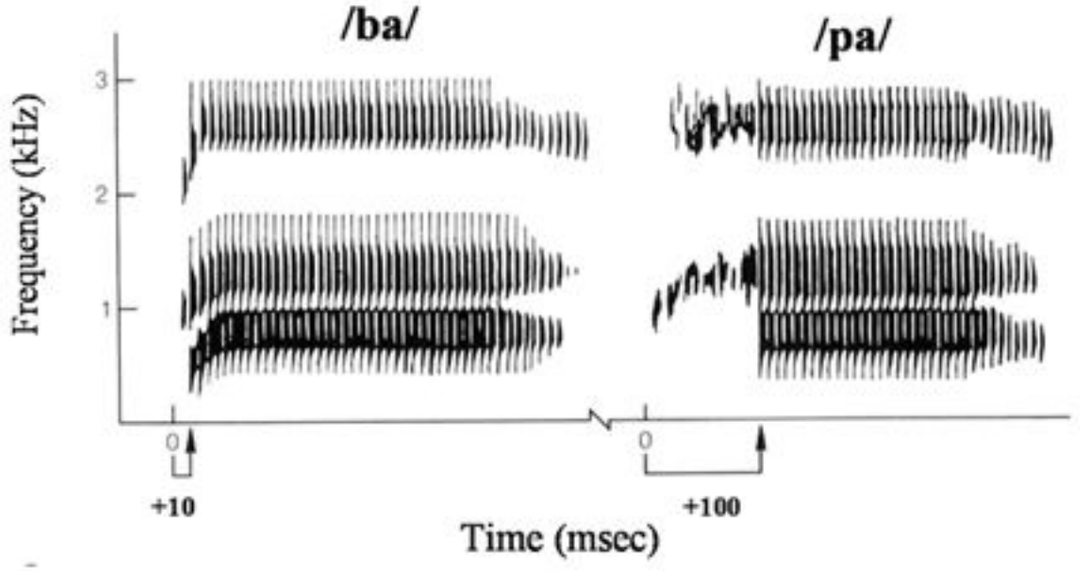
Design & Materials



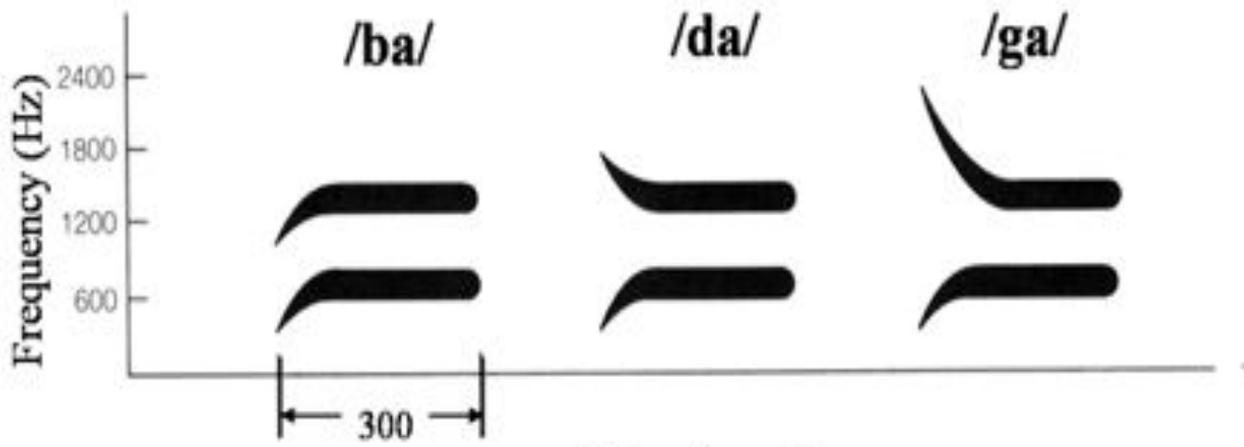
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms
- VOT (31 pairs) {p-b, t-d, k-g} and POA (22 pairs) {t-k, p-t}

Design & Materials

Contrast: Voice Onset Time (VOT)

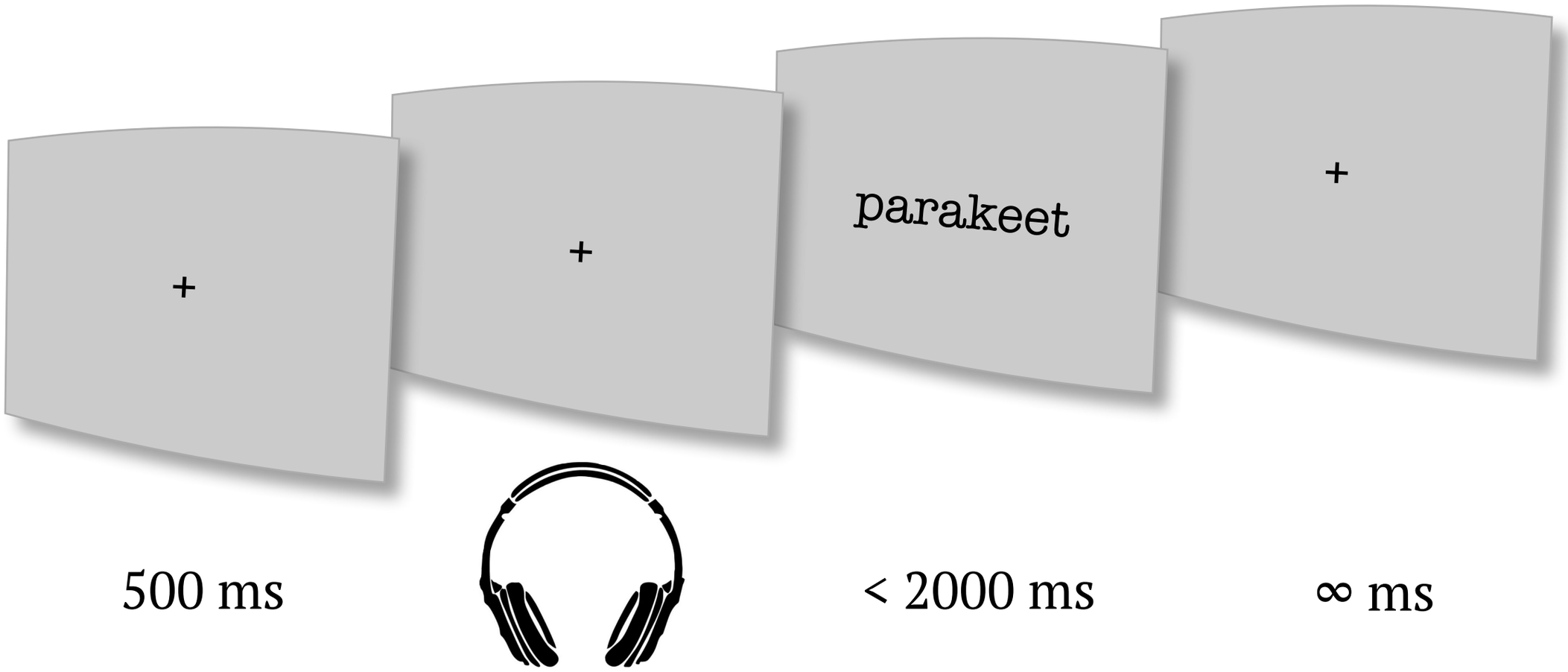


Contrast: Place of Articulation (POA)



- 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

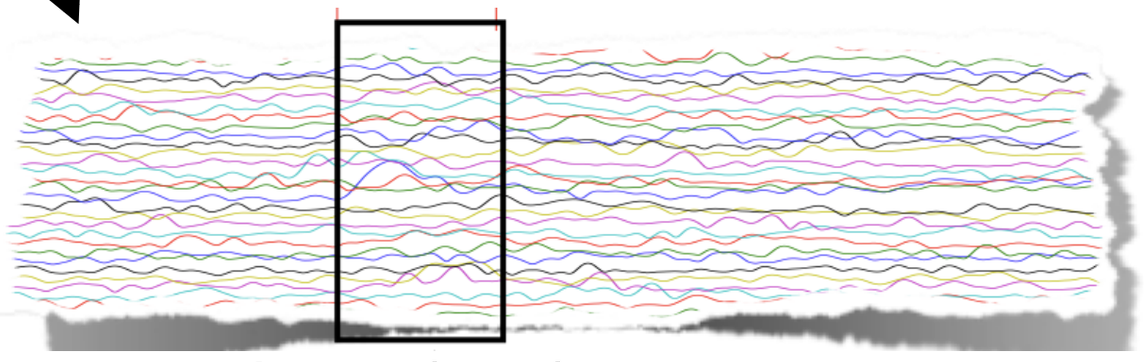


Procedure & Analysis

x 25

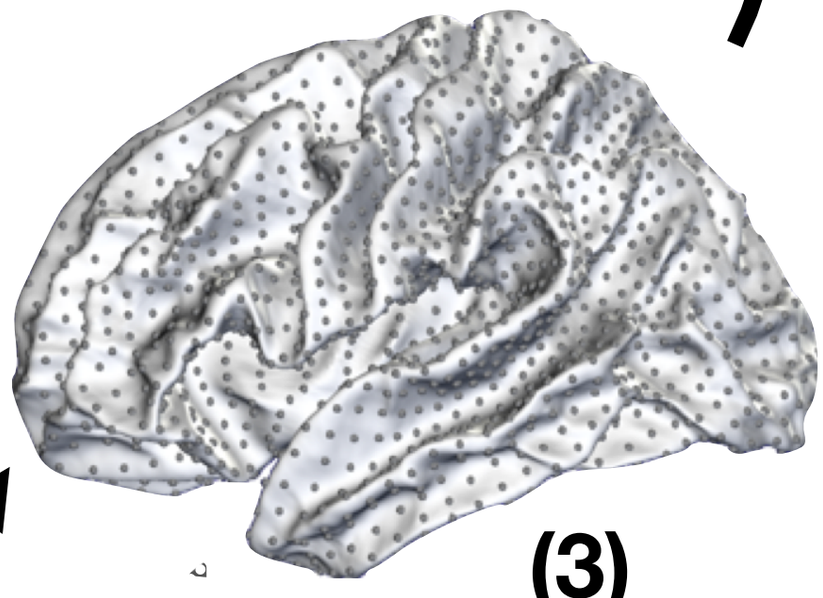
(1)

208 sensors



time (ms)

(2)



(3)

(4)

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

Today's Questions

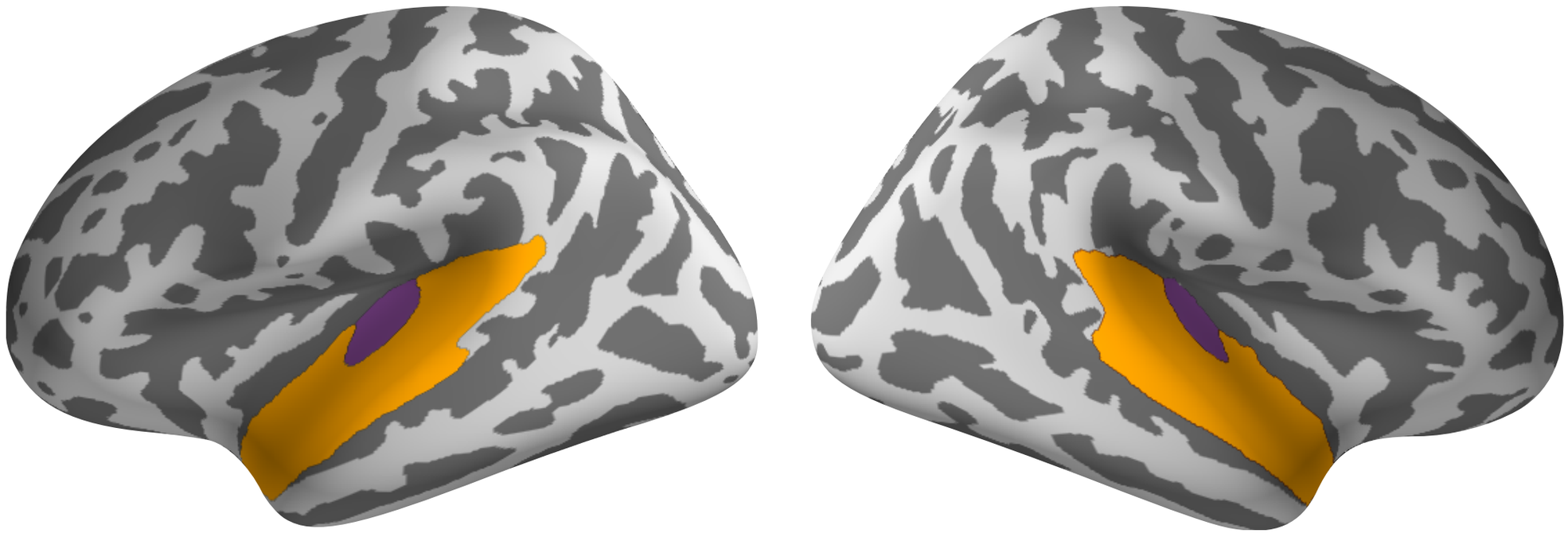
How does the auditory cortex **respond** to phonological ambiguity?

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

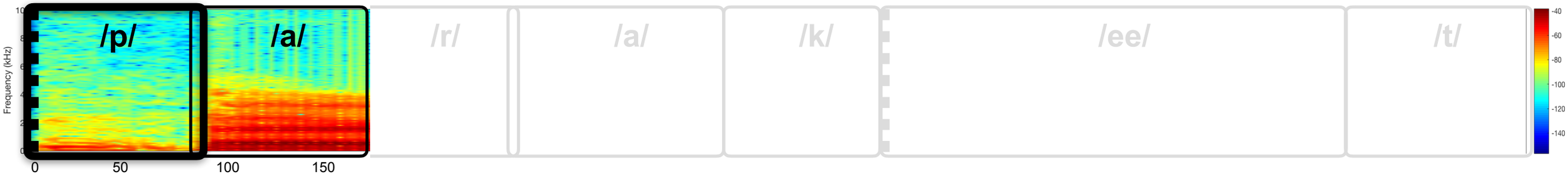
Simos et al. 1998, Ackermann et al. 1999, Obleser et al. 2003, Papanicolaou et al. 2003, Obleser et al. 2004 Mesgarani et al. 2014, Di Liberto et al. 2015

p p b b b

Ambiguity at Onset

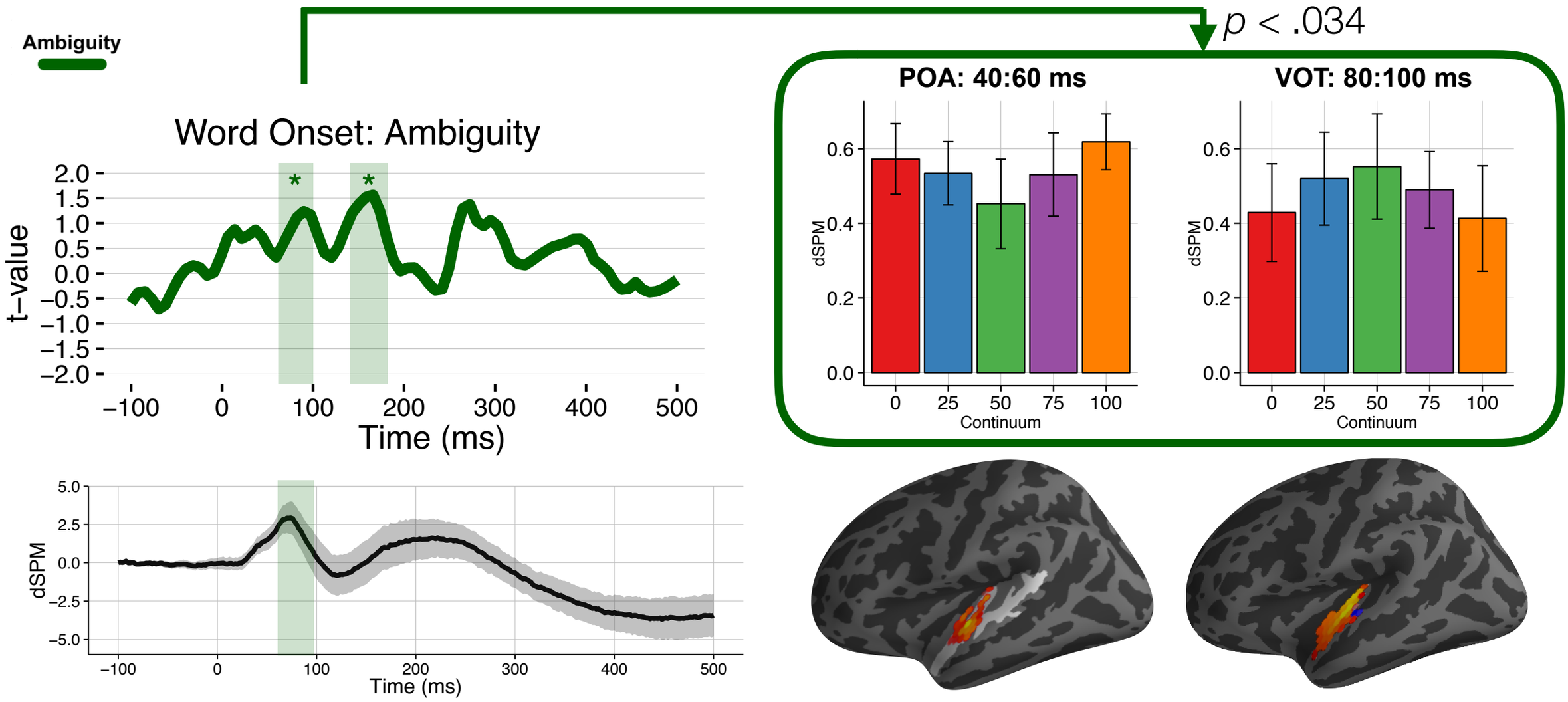


- Time-window: 0-200 ms after word onset
- Region: **Heschl's gyrus** & **superior temporal gyrus** bilaterally

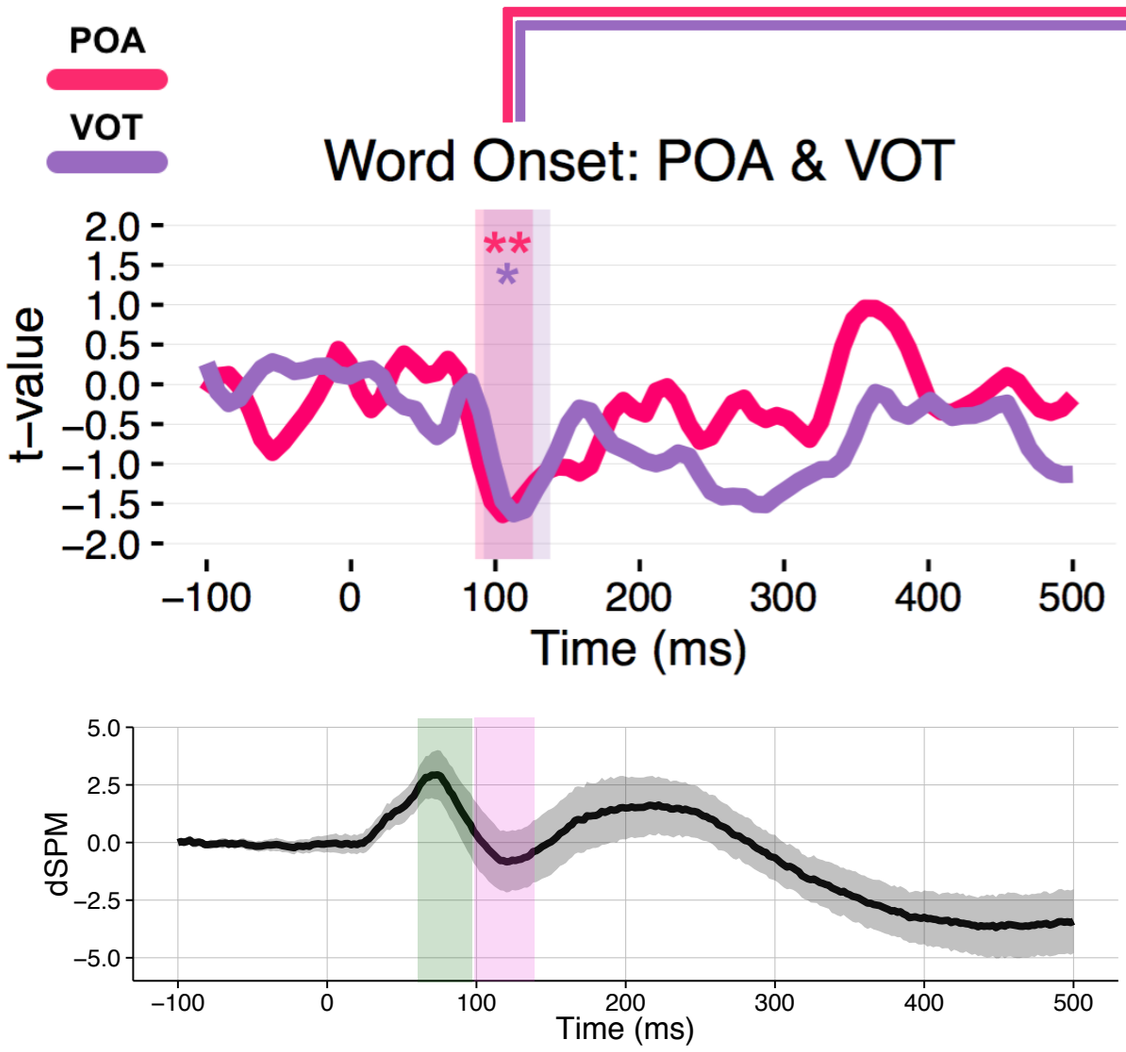


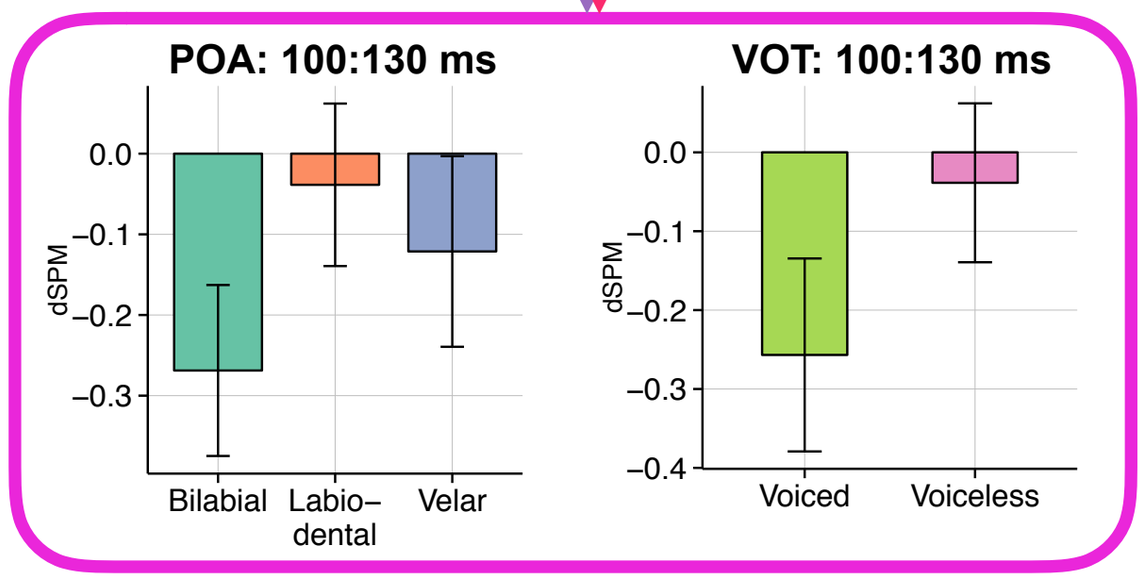
p p b b b

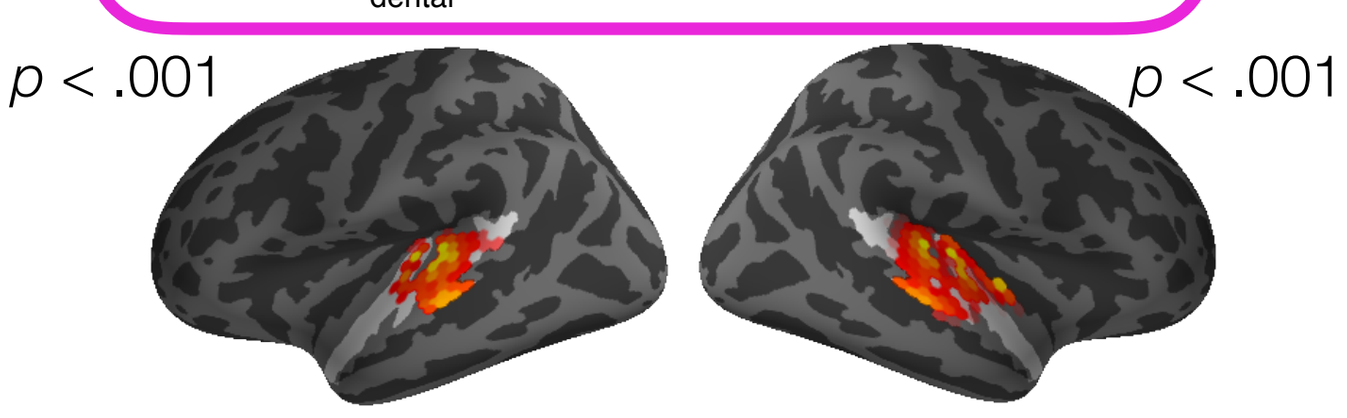
Ambiguity at Onset

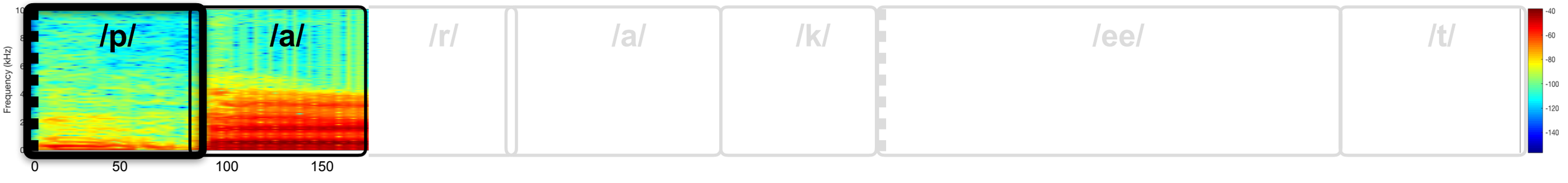


Ambiguity at Onset

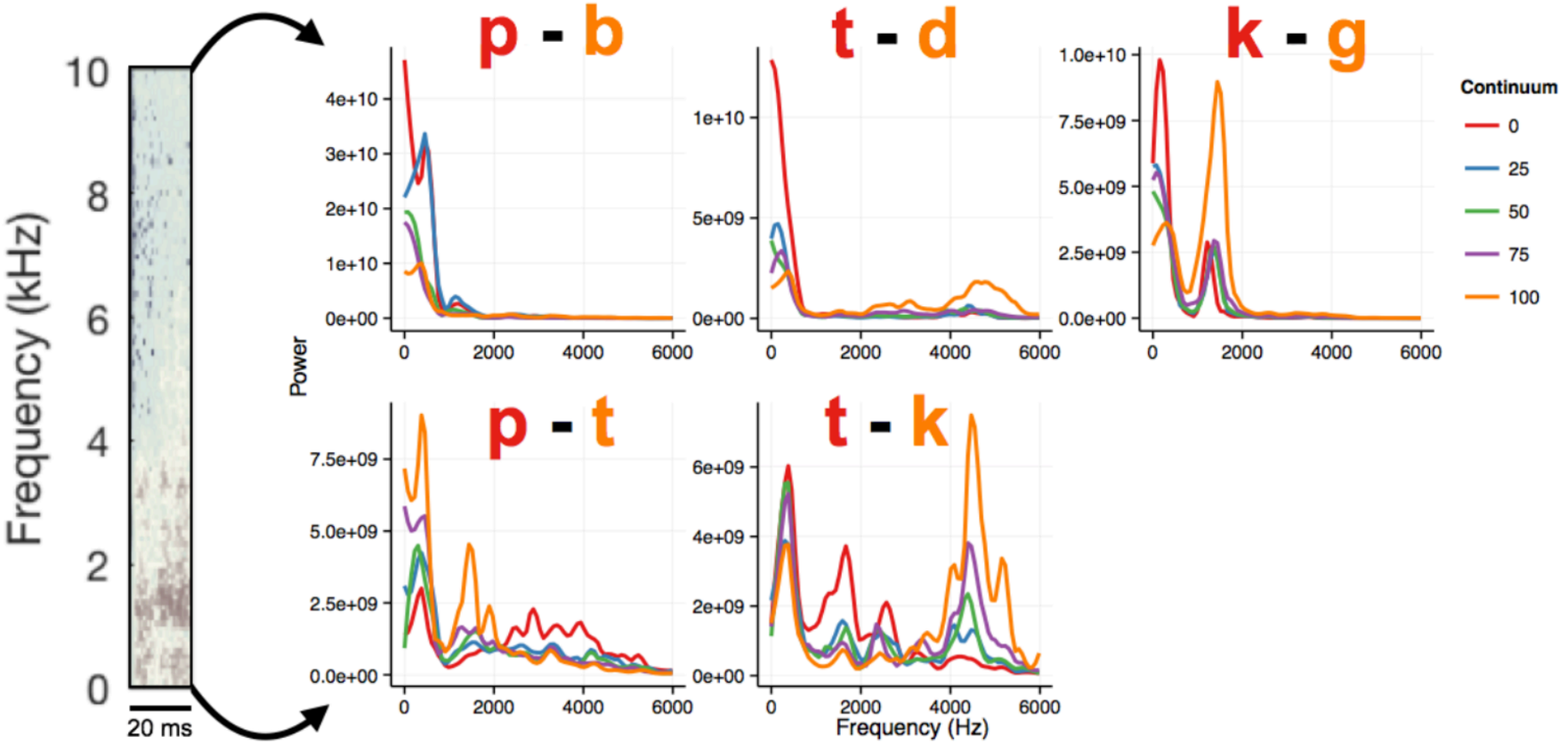


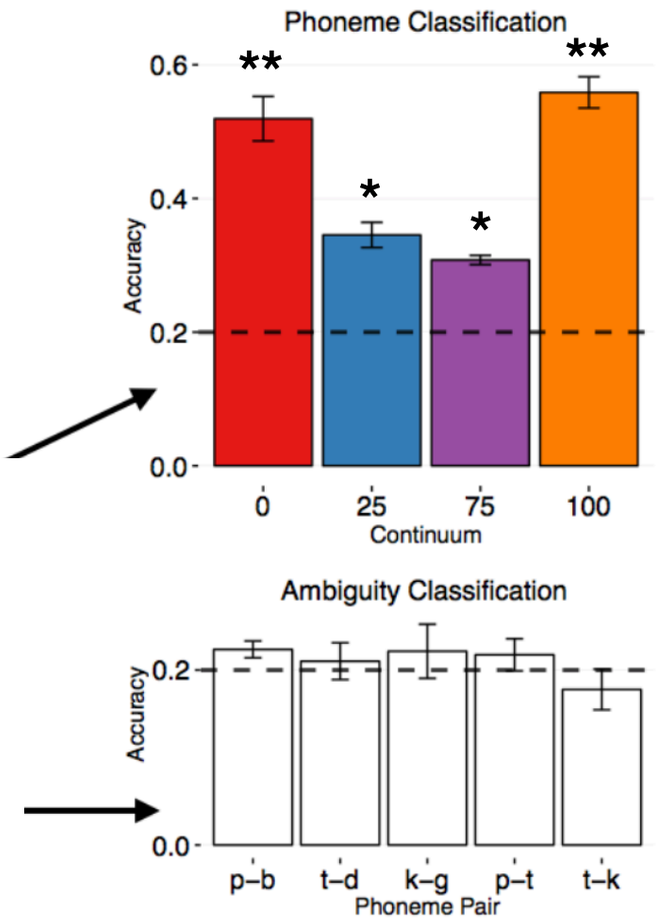




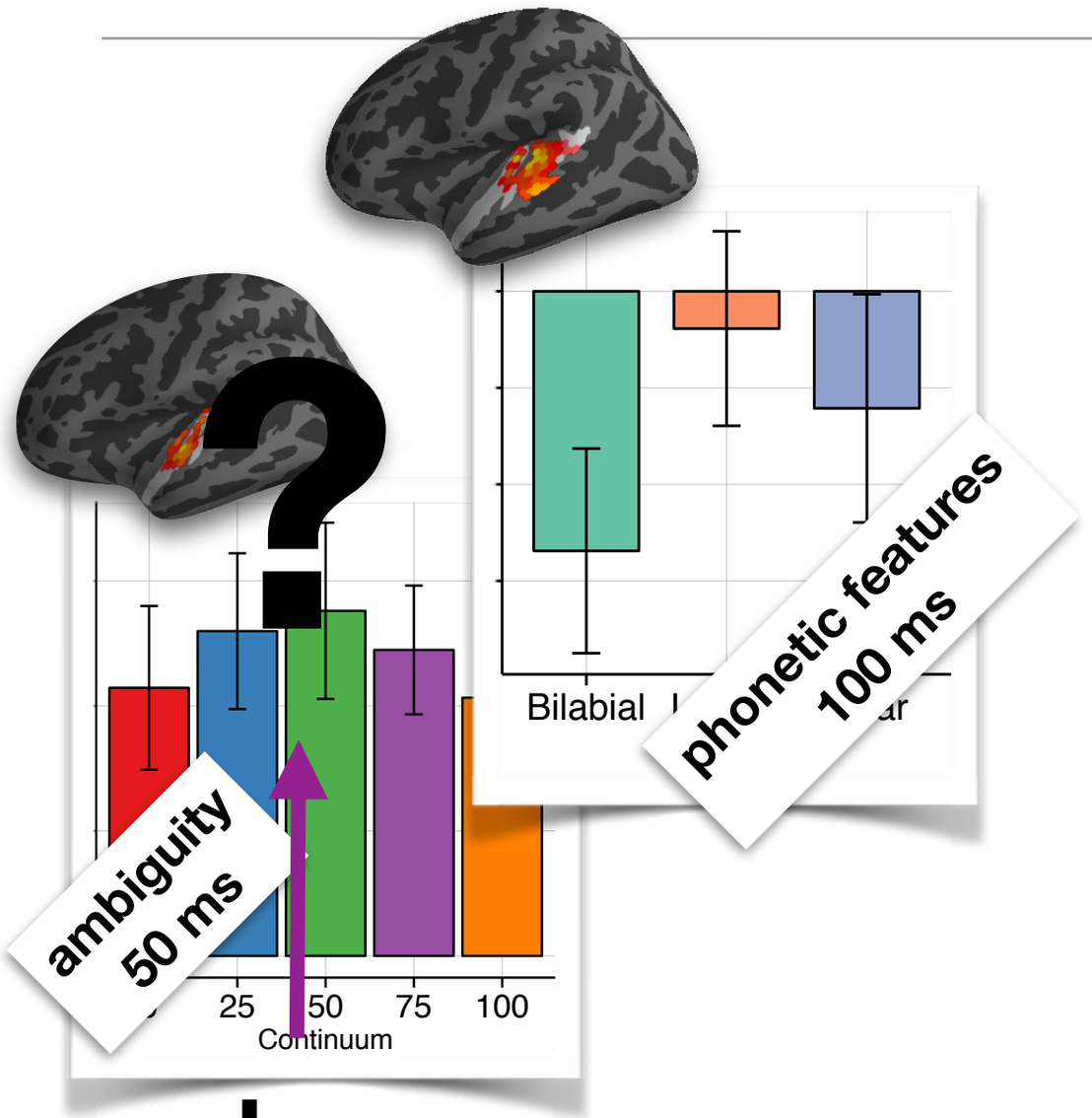


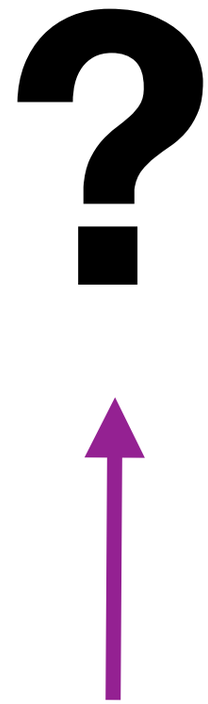
Ambiguity at Onset





Interim Conclusion



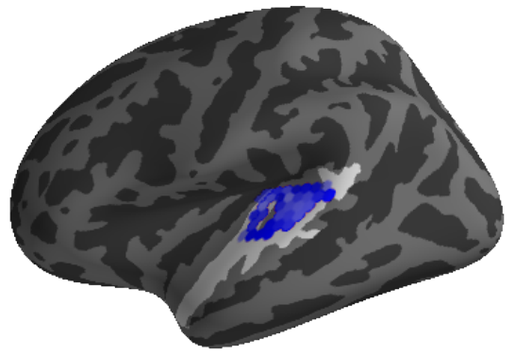


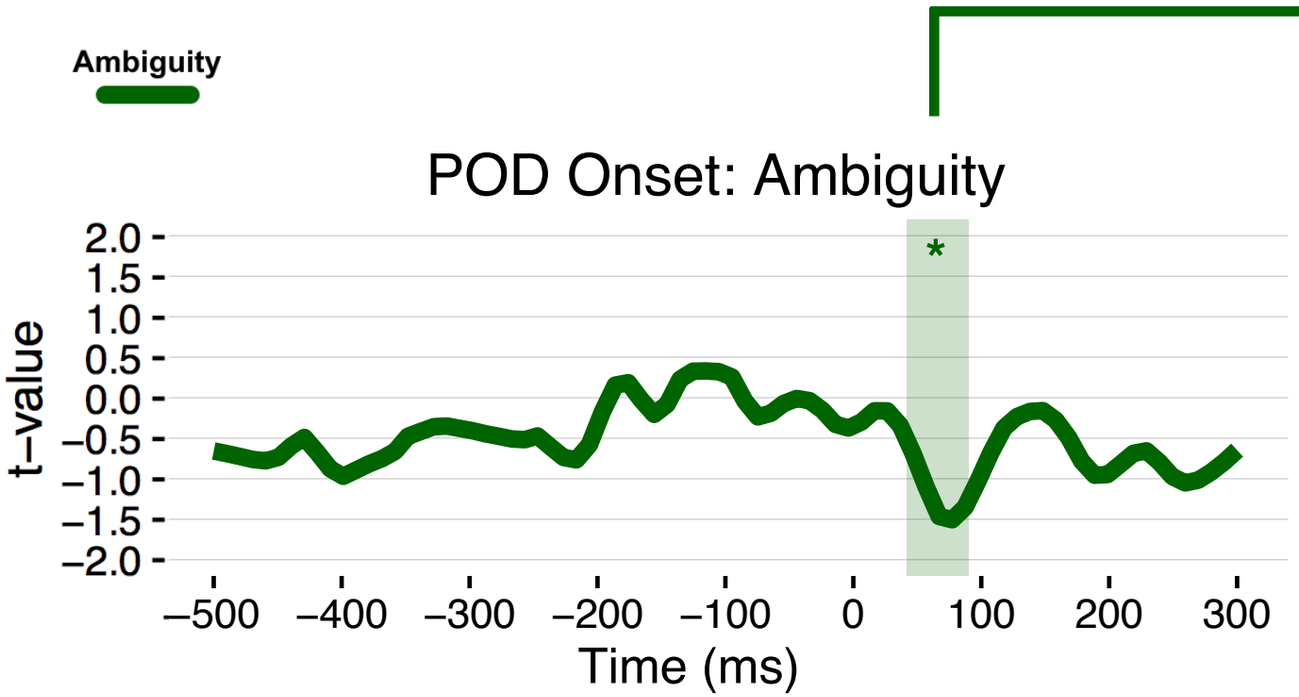
p a r a k **ee** t

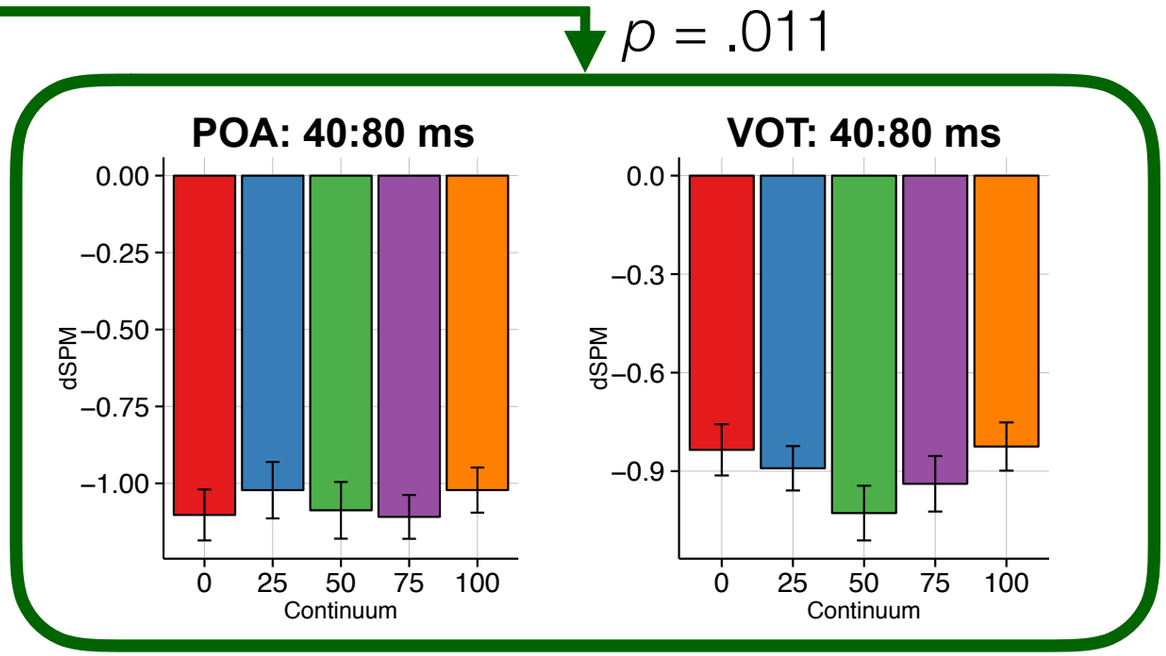
Today's Questions

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

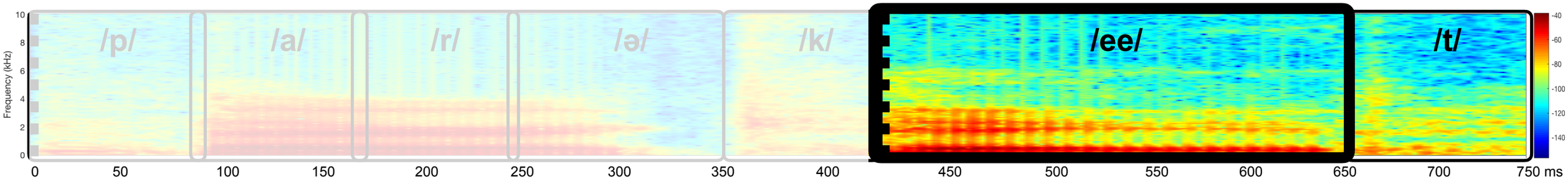
Ambiguity at POD



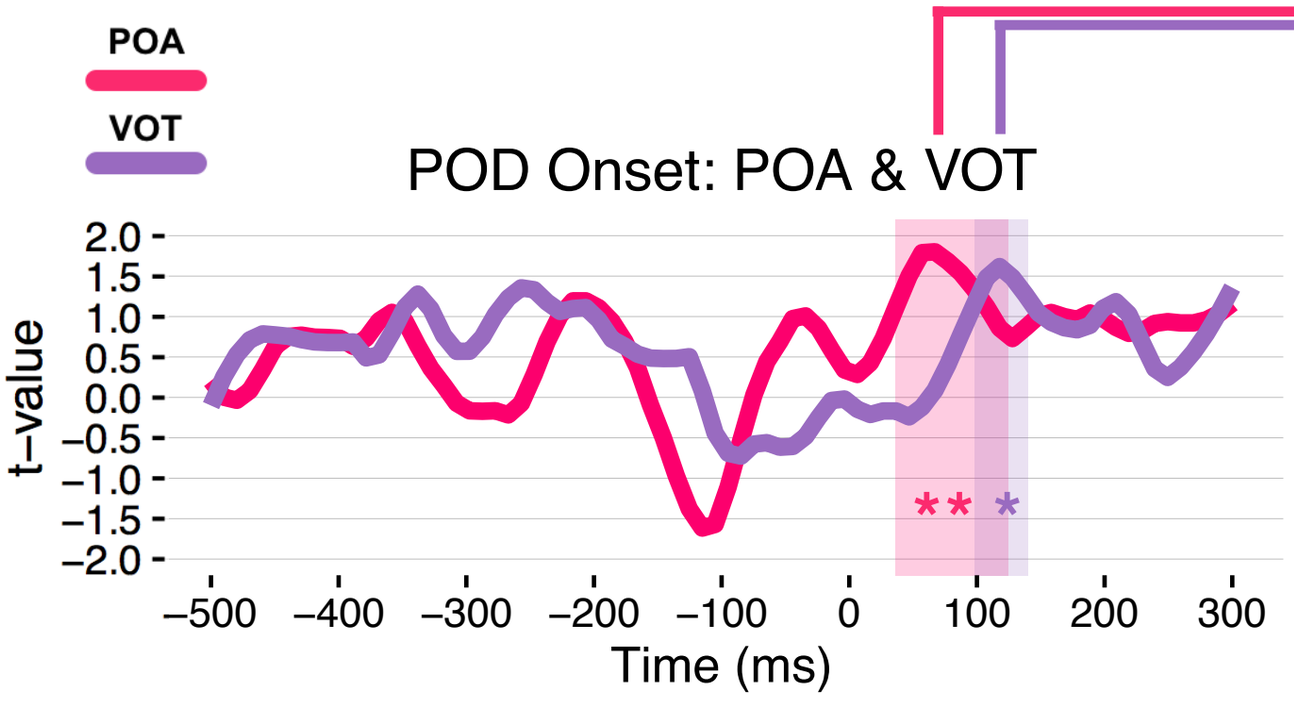


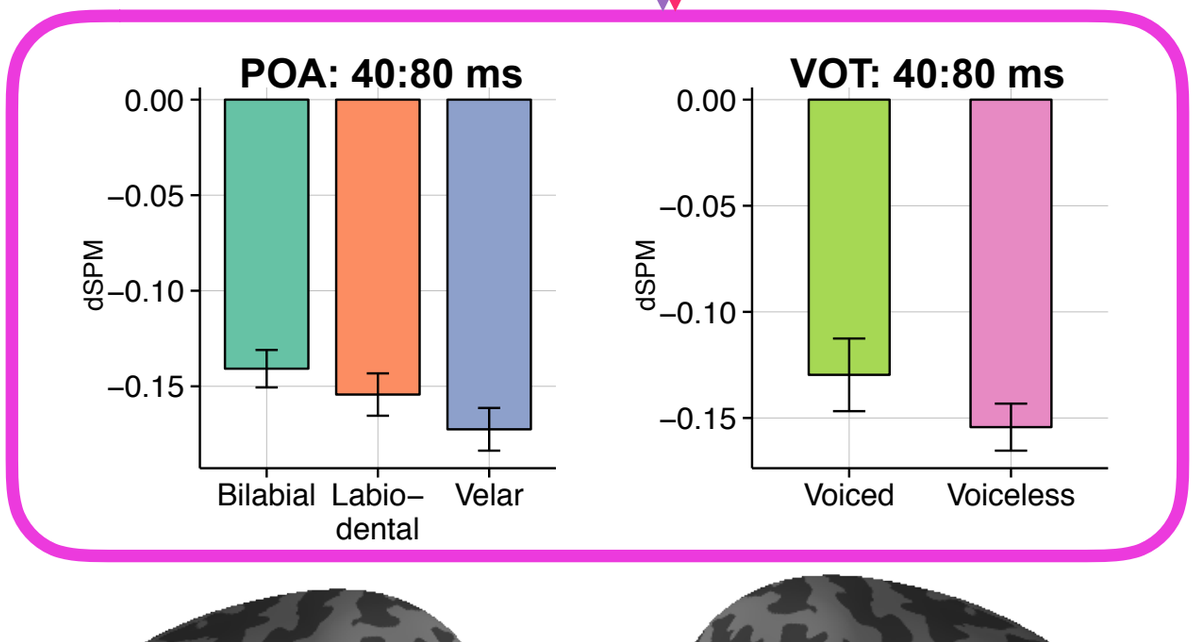


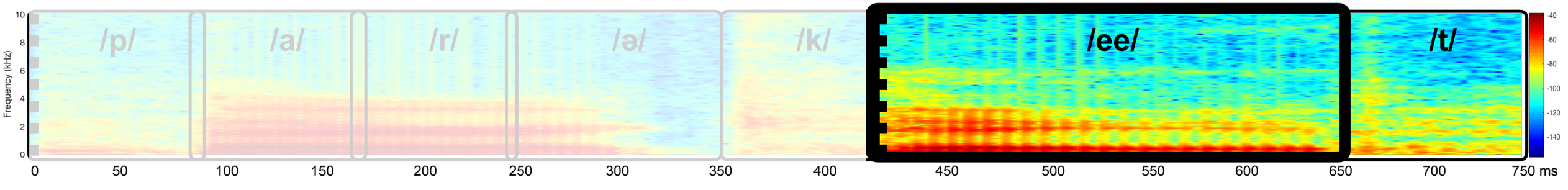
- Time-window: 0-200 ms after POD onset
- Region: **Heschl's gyrus** & **superior temporal gyrus** bilaterally



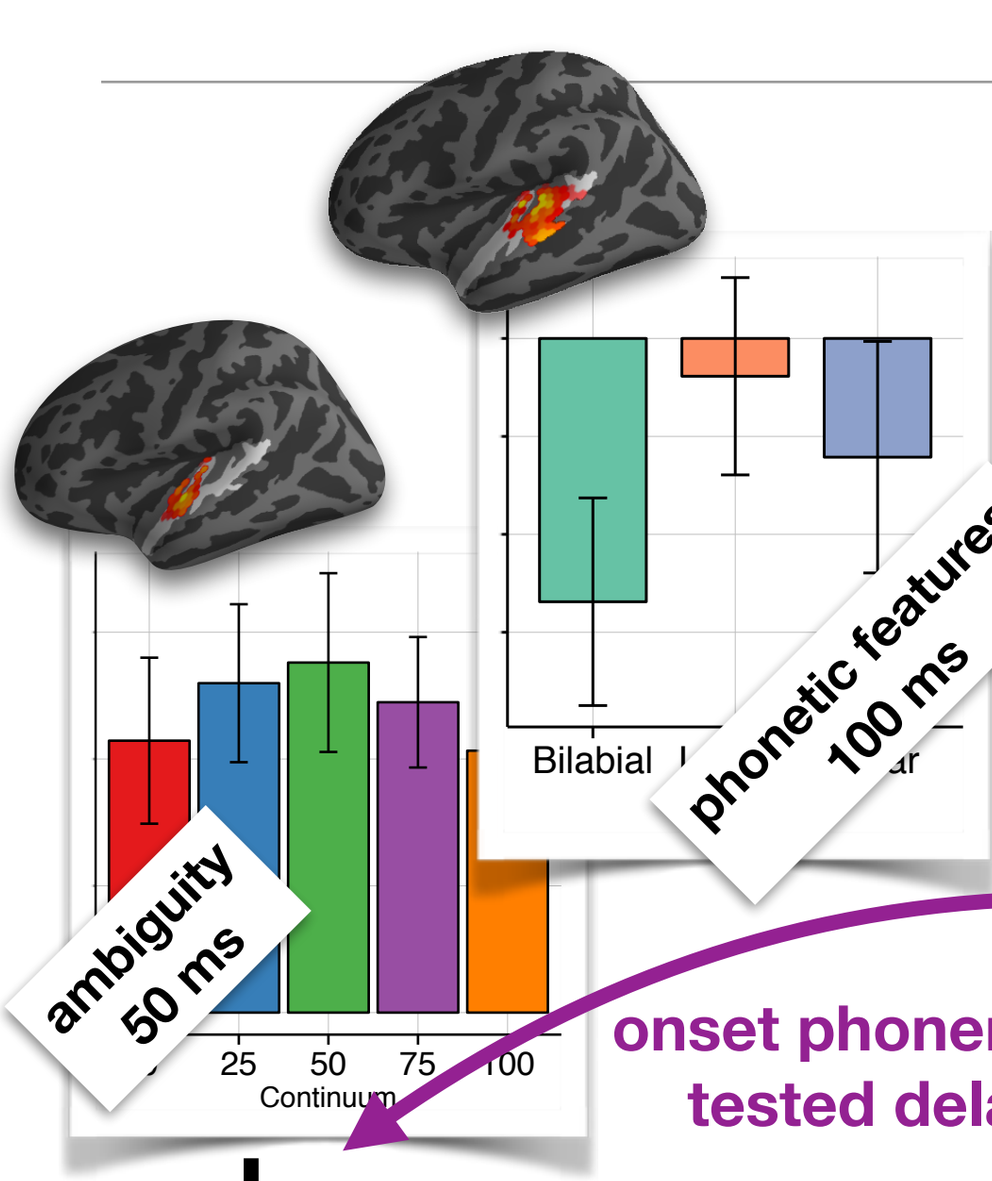
Ambiguity at POD

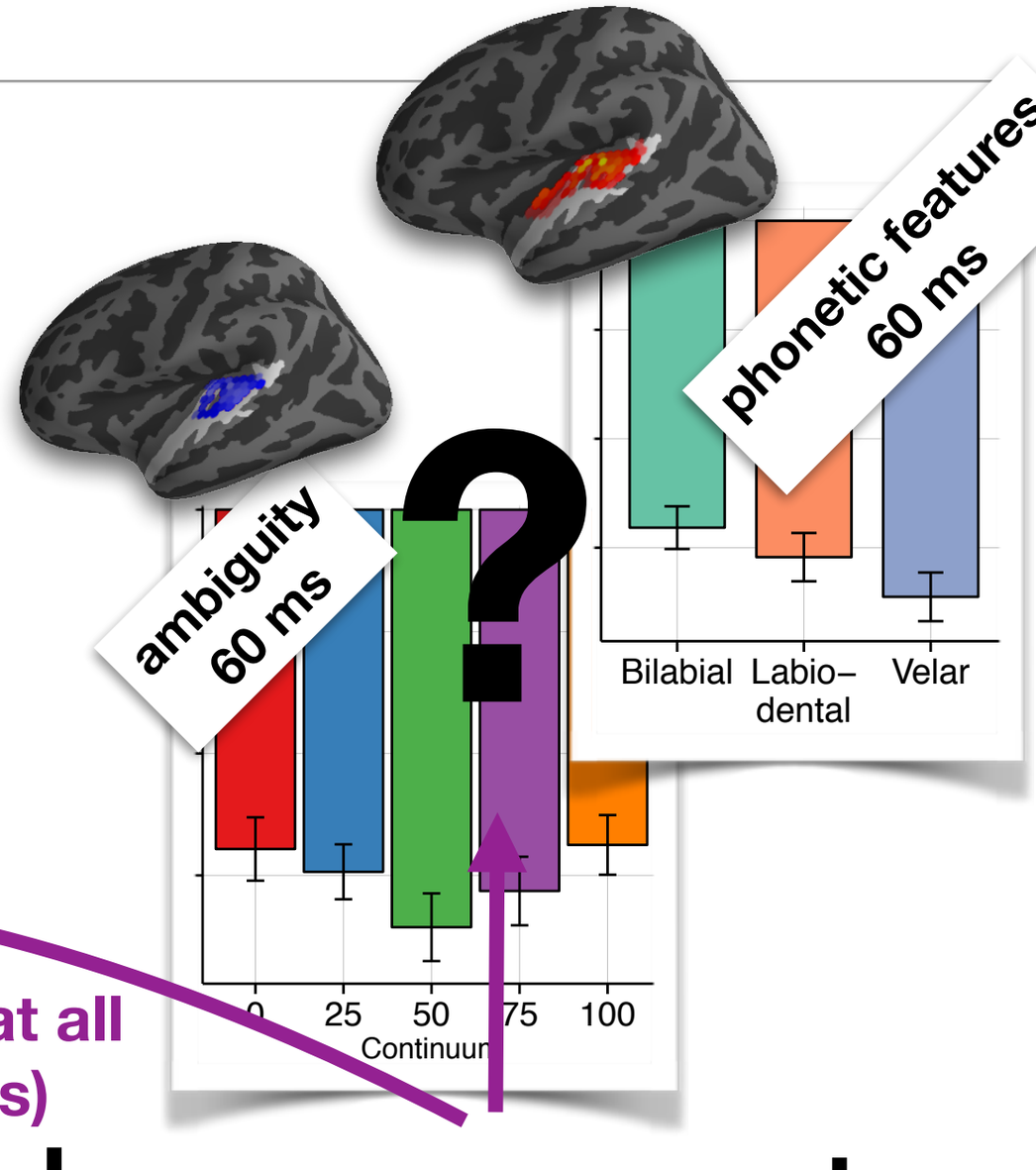






Interim Conclusion

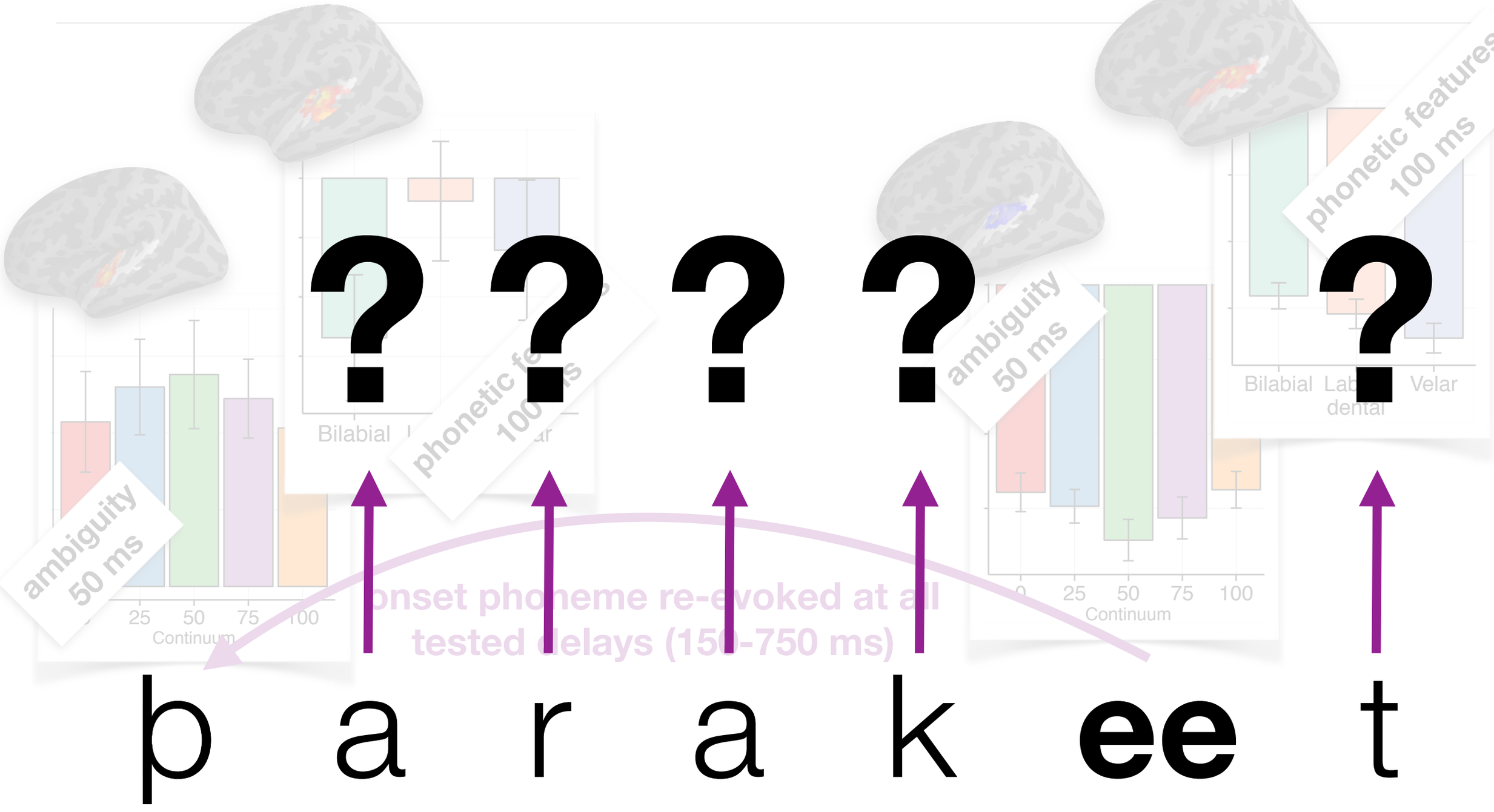




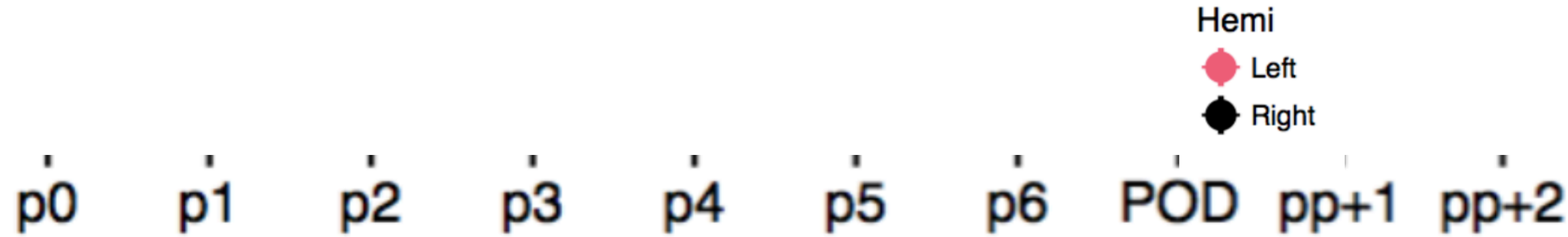
onset phoneme re-evoked at all tested delays (150-750 ms)

p a r a k **ee** t

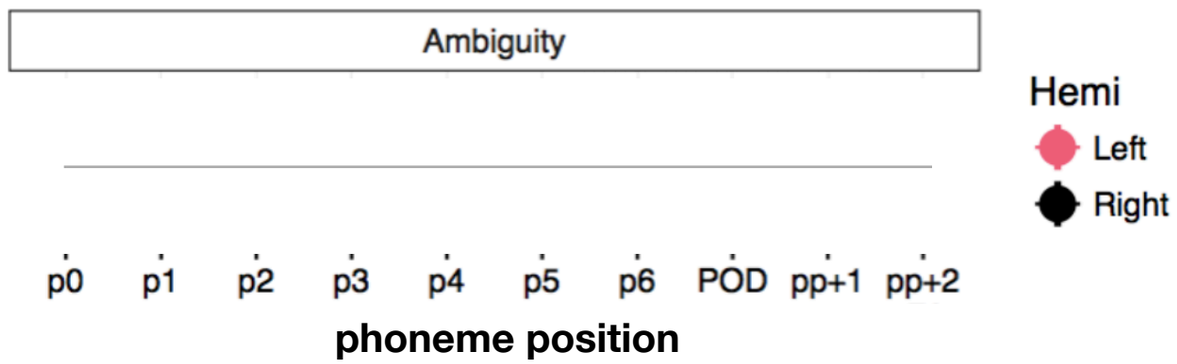
Interim Conclusion



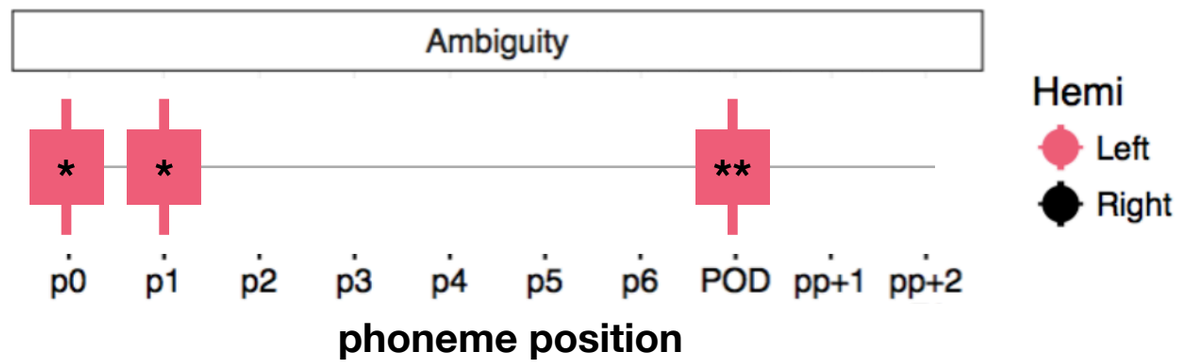
Reactivation in Intermediate Positions



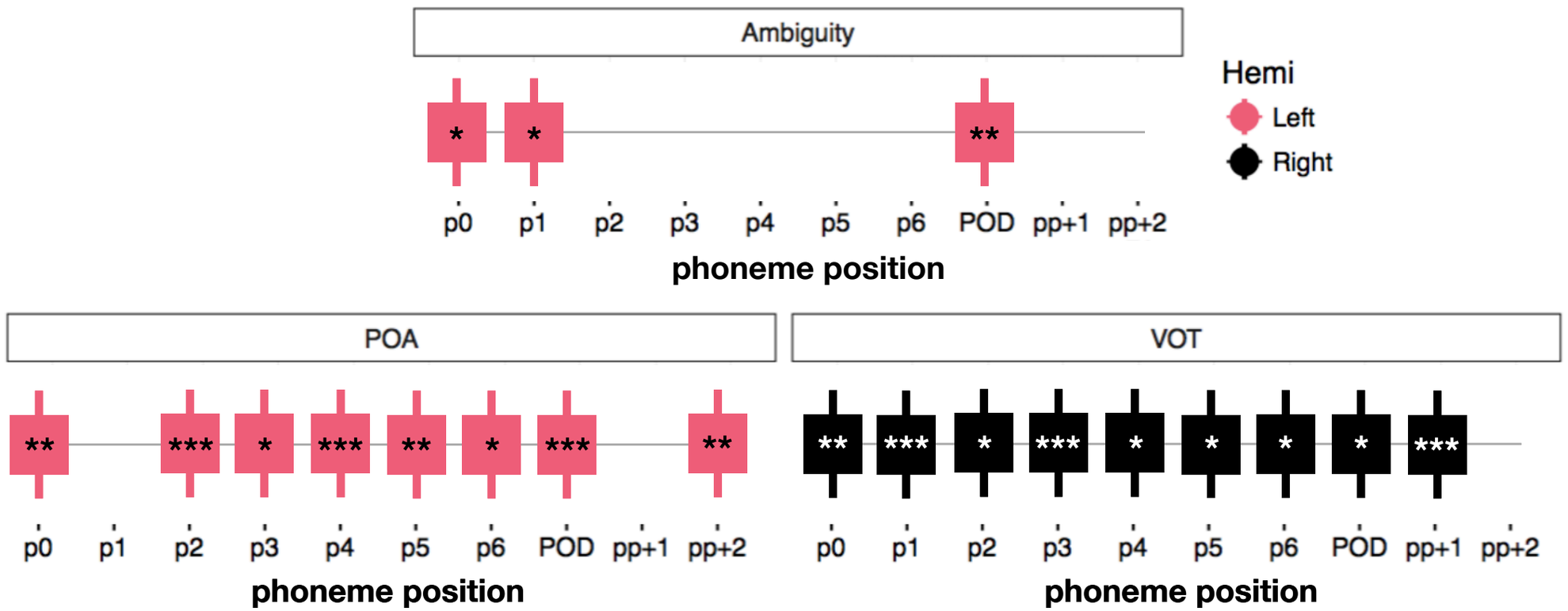
Reactivation in Intermediate Positions



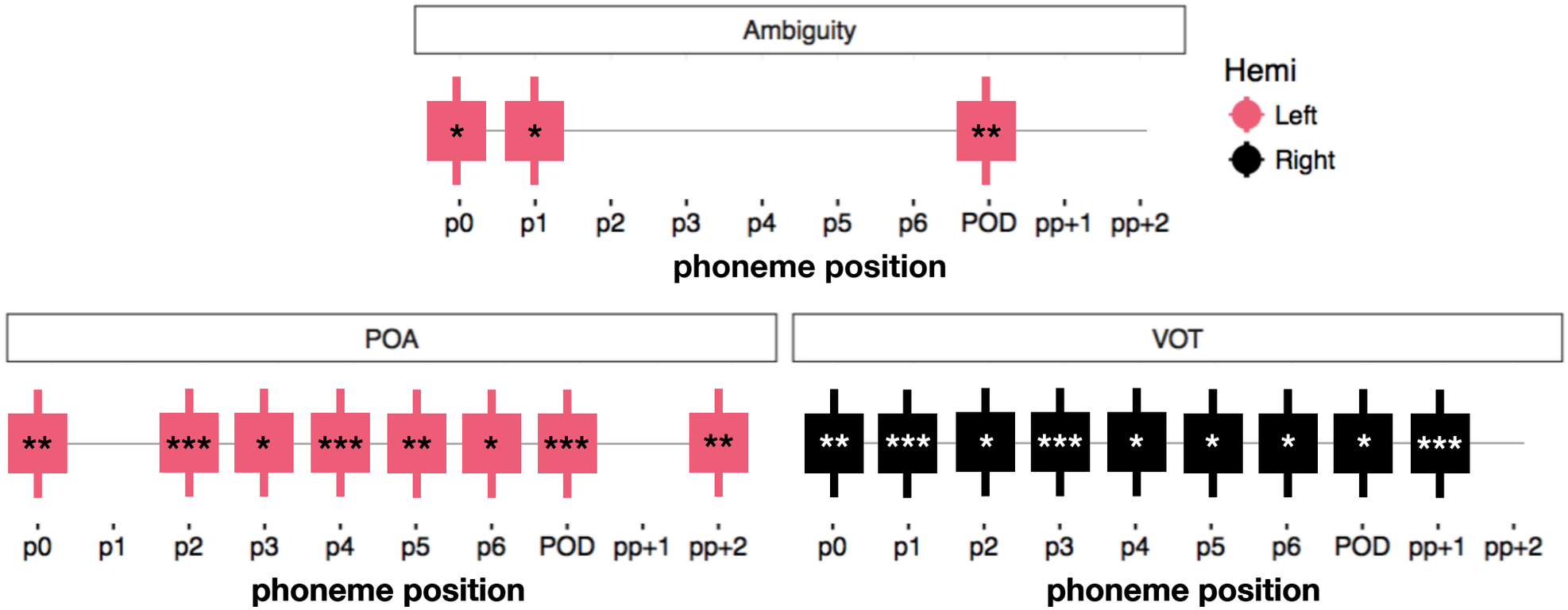
Reactivation in Intermediate Positions



Reactivation in Intermediate Positions

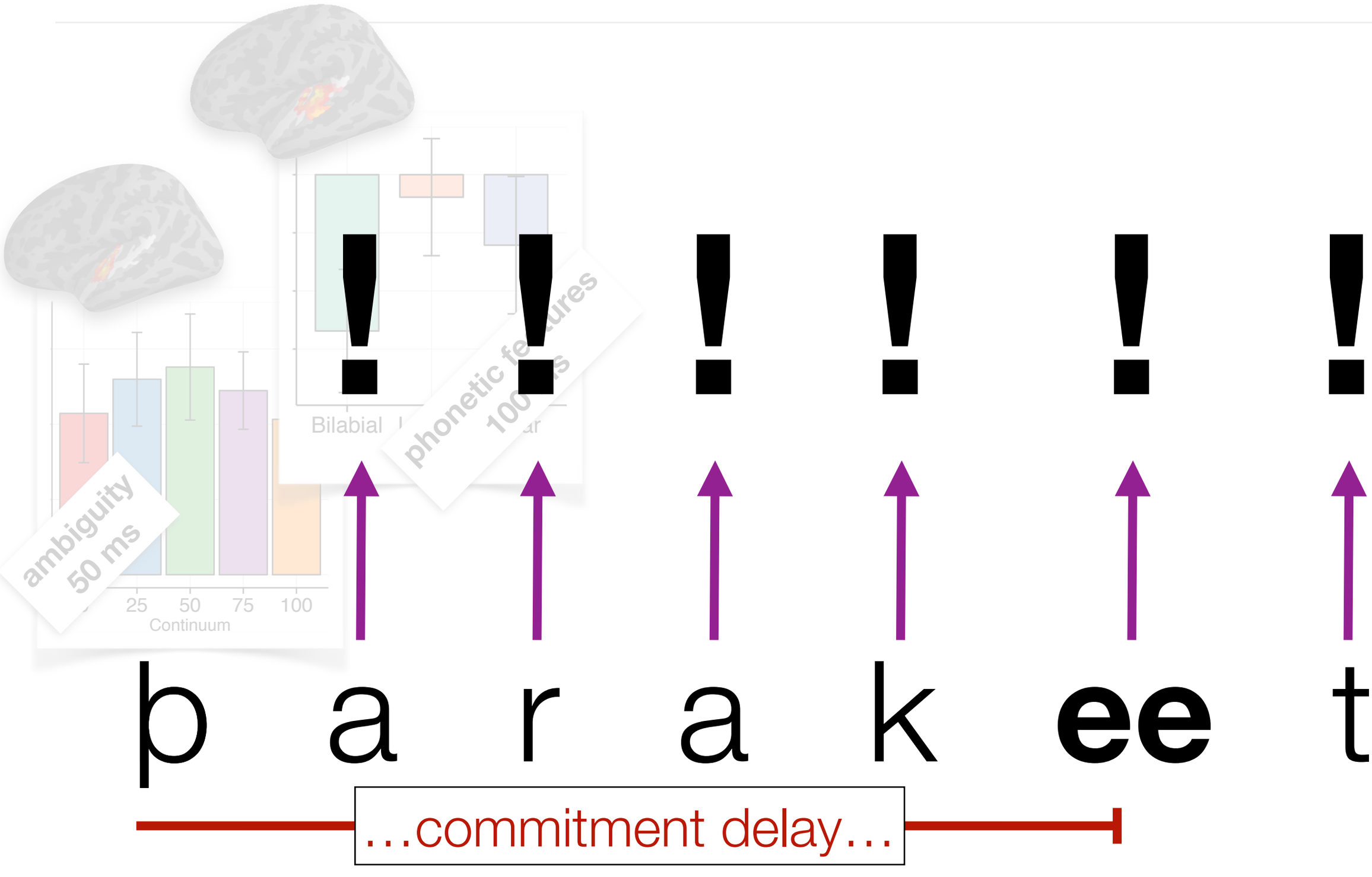


Reactivation in Intermediate Positions



- 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

Two states of the world

No commitment

Commitment

ballet

prove

pin

bath

pacify

bond

palate

beef

book

p

pants

balance

bind

paddle

boast

poke

panda

ballet

prove

pin

bath

pacify

bond

palate

beef

book

b p

b p

pants

balance

bind

paddle

boast

poke

panda

ballet

prove

pin

bath

pacify

bond

palate

beef

book

b

pants

balance

bind

paddle

boast

poke

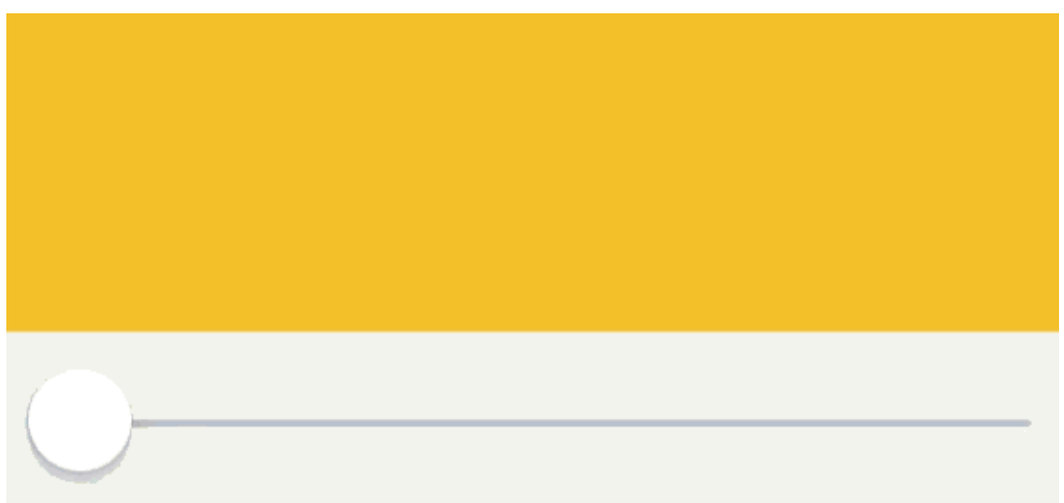
panda

Two states of the world

$$P(\varphi_a|A)$$

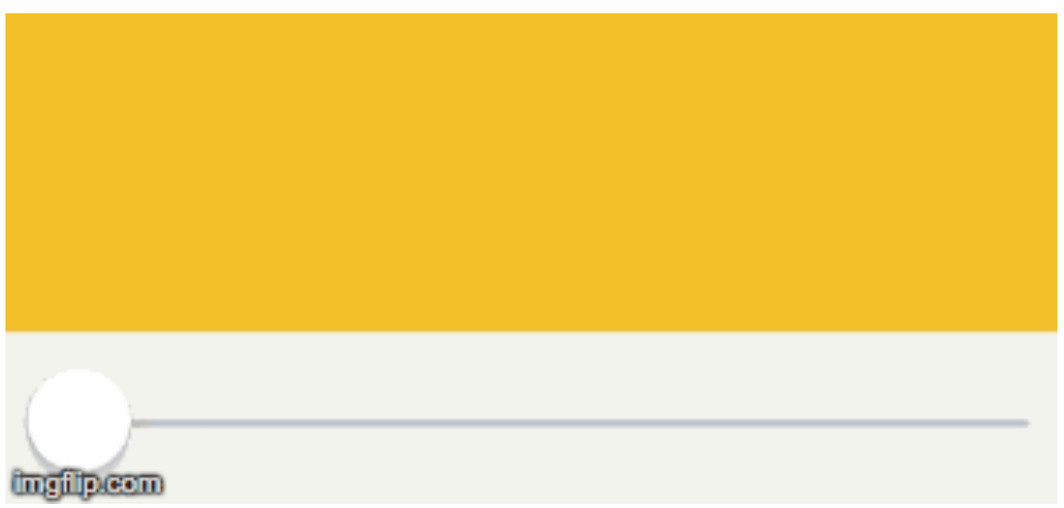
φa = phoneme$_a$ A = acoustic input

No commitment



- 1+ cohort of words
- continuous acoustic term

Commitment



- 1 cohort of words
- binary acoustic term

Critical Variables

- **Surprisal**:

Probability of an outcome

$$-log_2 \frac{f(\varphi_1, \ldots, \varphi_t)}{f(\varphi_1, \ldots, \varphi_{t-1})}$$

- **Entropy:**

Uncertainty over future input

$$-\sum_{w \in C} P(w|C) log_2 P(w|C)$$

Critical Variables

- **Surprisal**:

No commitment
Commitment

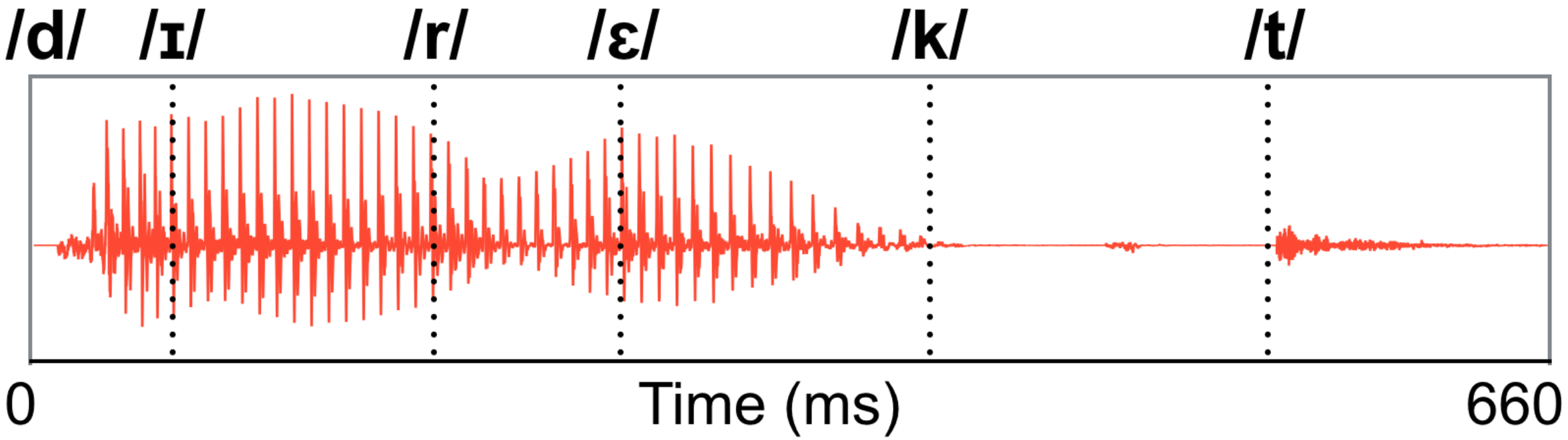
$$-log_2\left(P(\varphi_a|A)\frac{f(\varphi_a,\varphi_2,\ldots,\varphi_t)}{f(\varphi_a,\varphi_2,\ldots,\varphi_{t-1})}Q_a^t+\right.$$
$$\left.P(\varphi_b|A)\frac{f(\varphi_b,\varphi_2,\ldots,\varphi_t)}{f(\varphi_b,\varphi_2,\ldots,\varphi_{t-1})}Q_b^t\right)$$

- **Entropy:**

No commitment
Commitment

$$P(w|C,A) = P(w|C_a)P(\varphi_a|A)+$$
$$P(w|C_b)P(\varphi_b|A)$$

Procedure & Analysis



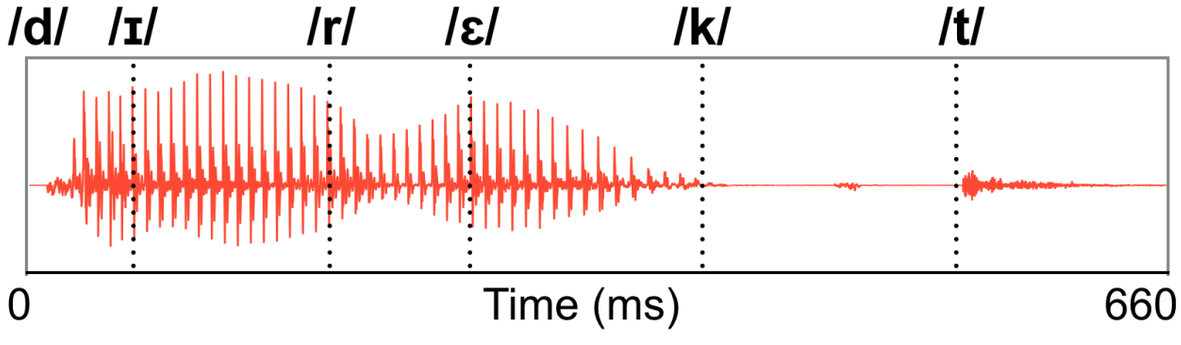
Model Setup

- **Critical variables**:

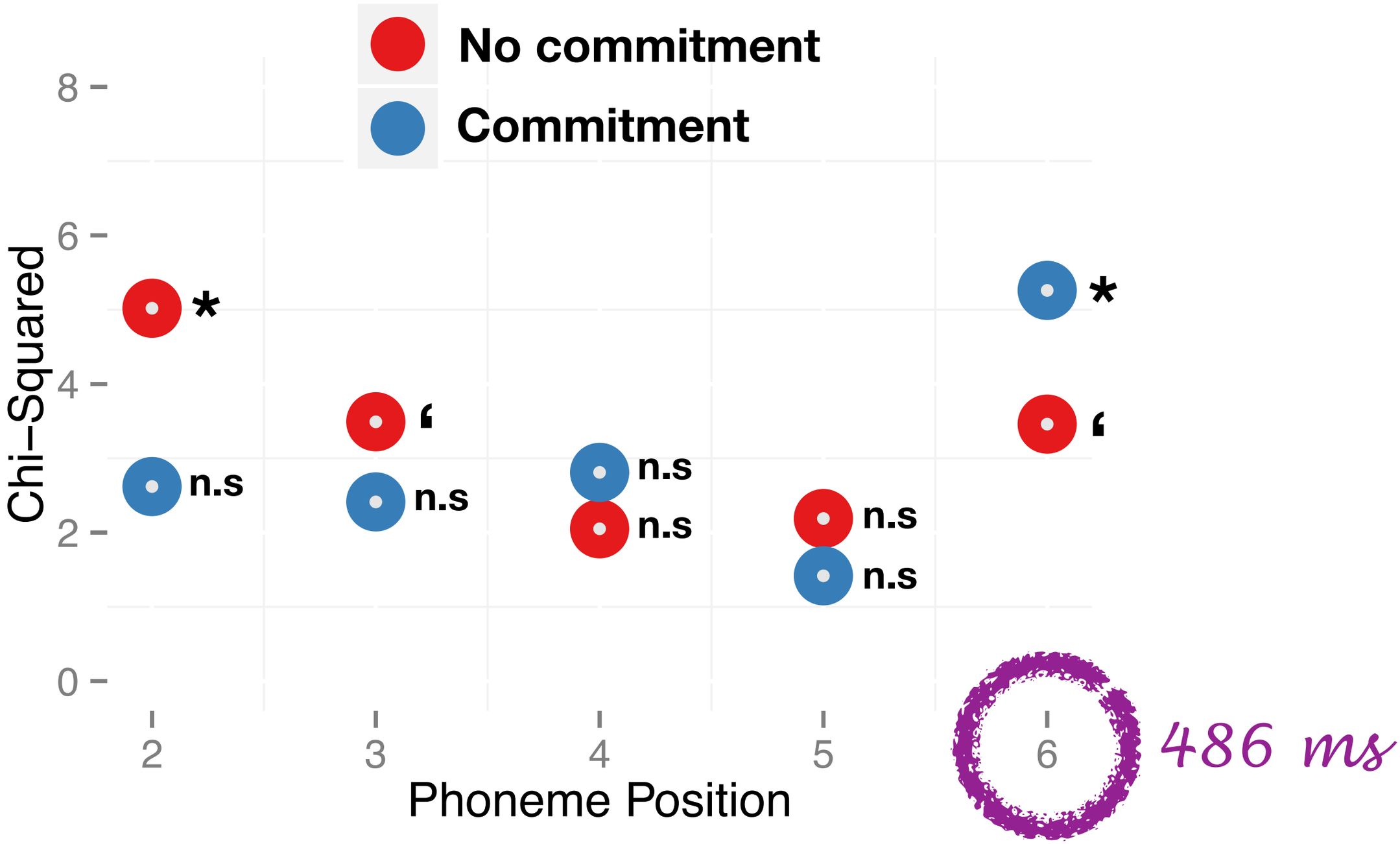
no commitment entropy
no commitment surprisal
commitment entropy
commitment surprisal

- **Control variables**:

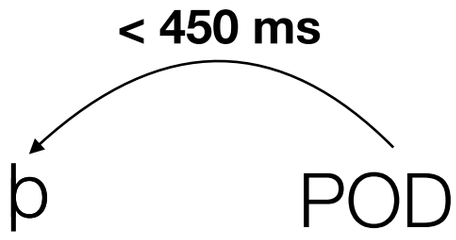
phoneme latency (ms)
phoneme latency (number of phonemes)
trial number
block number
stimulus amplitude
phoneme pair
ambiguity



Results

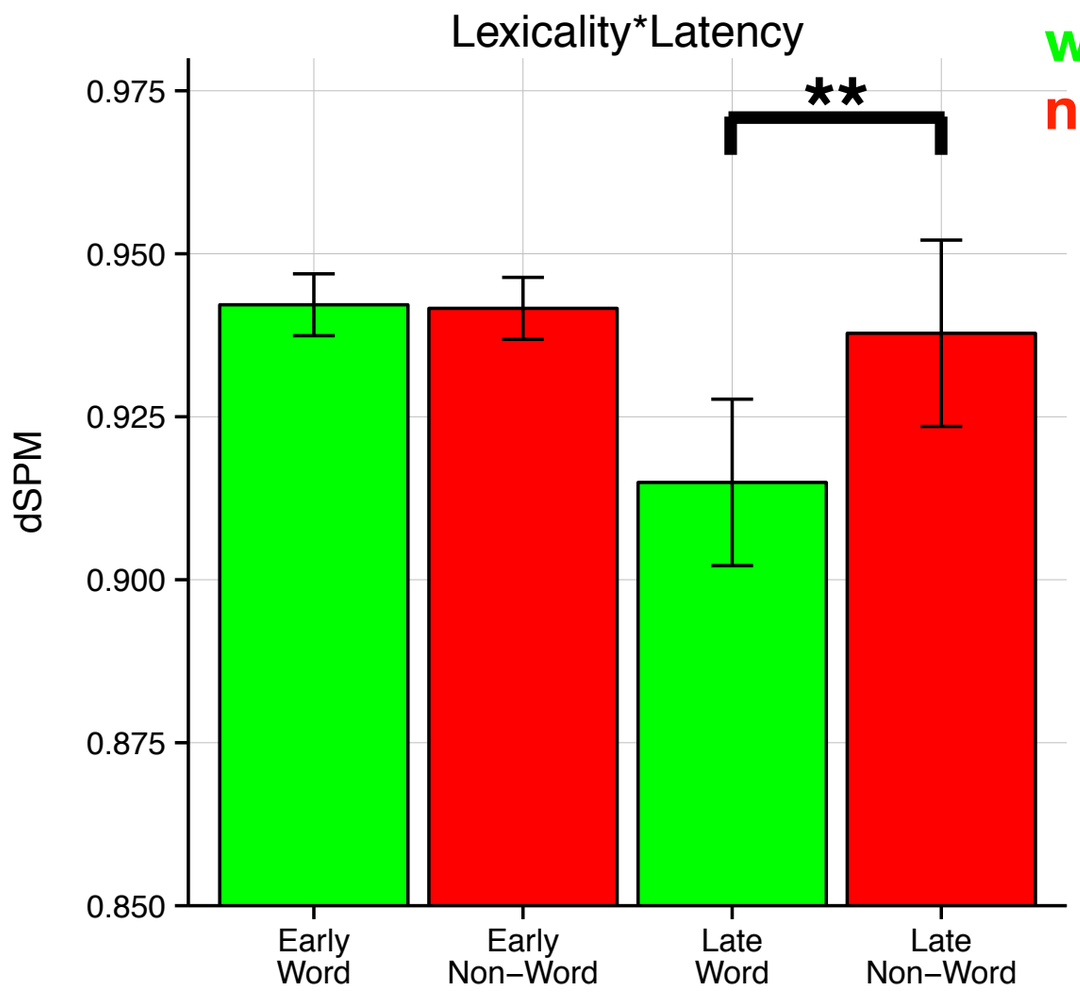

No commitment
Commitment
Chi–Squared
8
6
4
2
0
*
n.s
‘
n.s
n.s
n.s
n.s
*
‘
2
3
4
5
6
Phoneme Position
486 ms

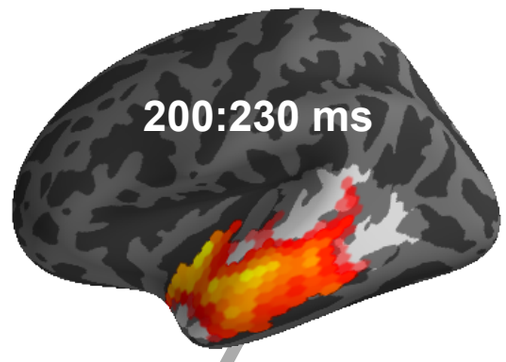
Commitment Before POD

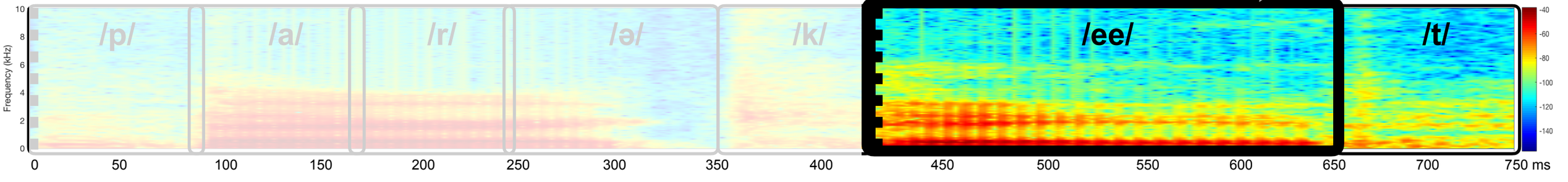


Early: POD <u>earlier</u> than 450 ms after word onset

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

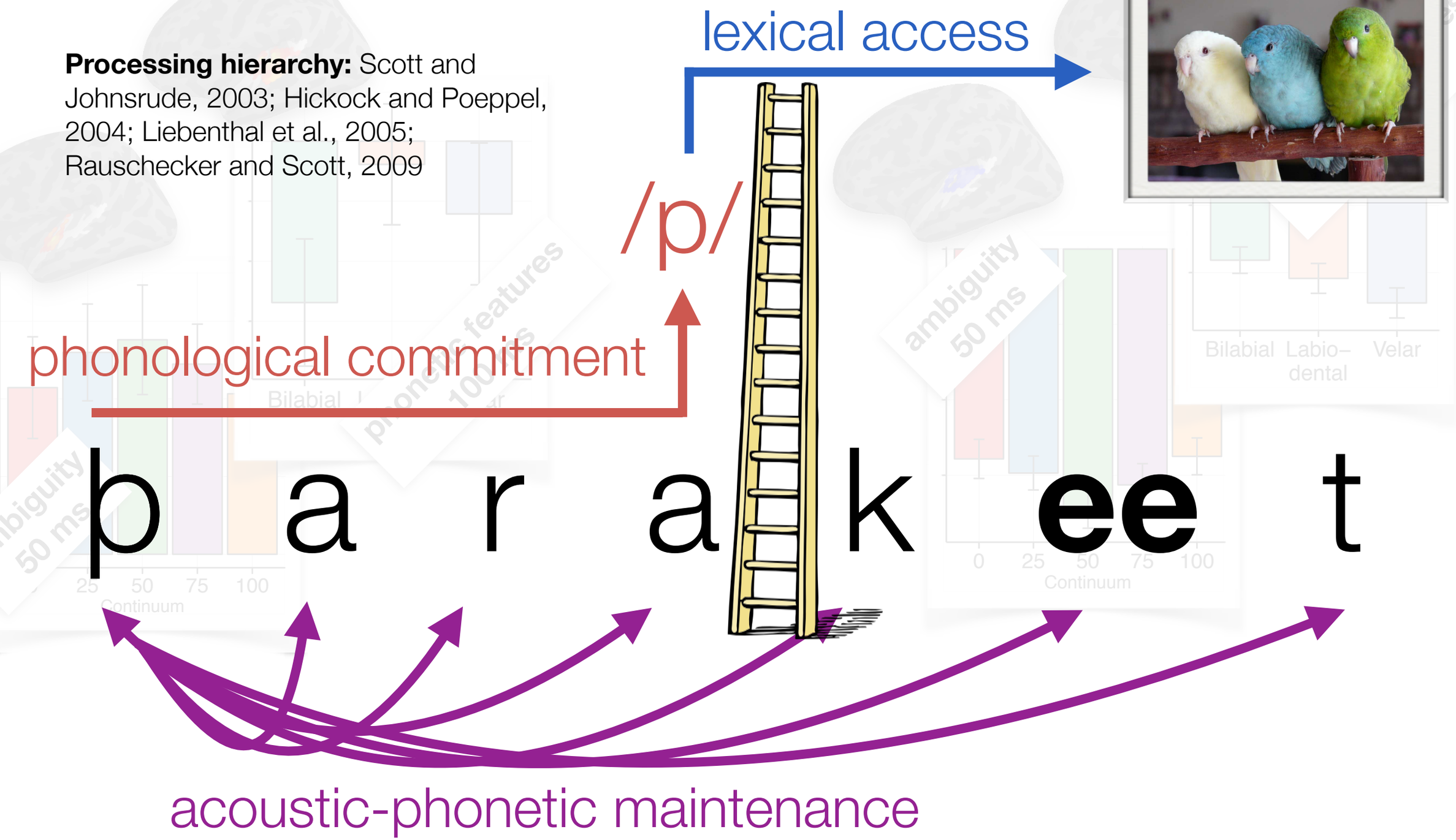






Interpretation

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



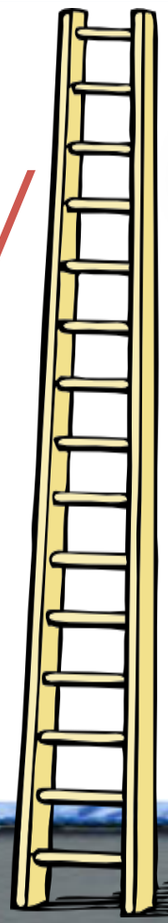
Interpretation

Processing is not purely feed forward, or feed up: TRACE model: McClelland and Elman, 1986; McMurray et al. 2009. Pauseback and Scott, 2009

lexical access

/p/

phonological commitment



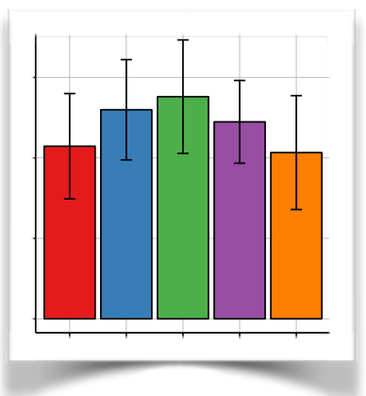
Analogies to other sources of ambiguity, such as homophone resolution: Twilley and Dixon, 2000; Rodd et al., 2010, Rodd, 2017

p a r a k **ee** t

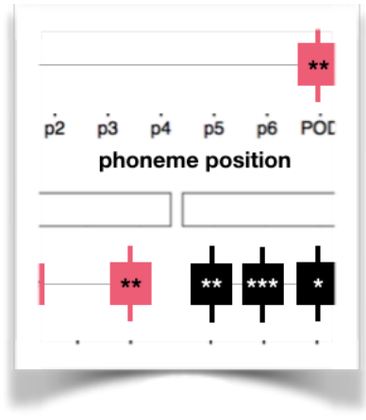
acoustic-phonetic maintenance

3 Take Home Messages

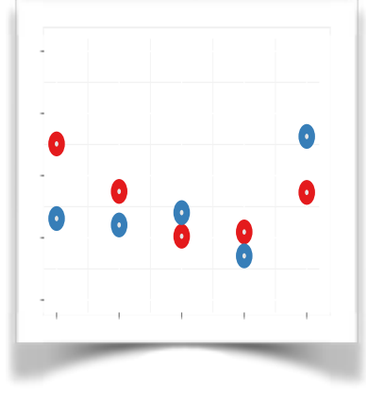
- Sensitivity to phoneme ambiguity ~50 ms after onset in primary auditory cortex



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



- The brain is not scared of commitment; it occurs ~500 ms after phoneme onset in superior temporal gyrus



With big thanks to:

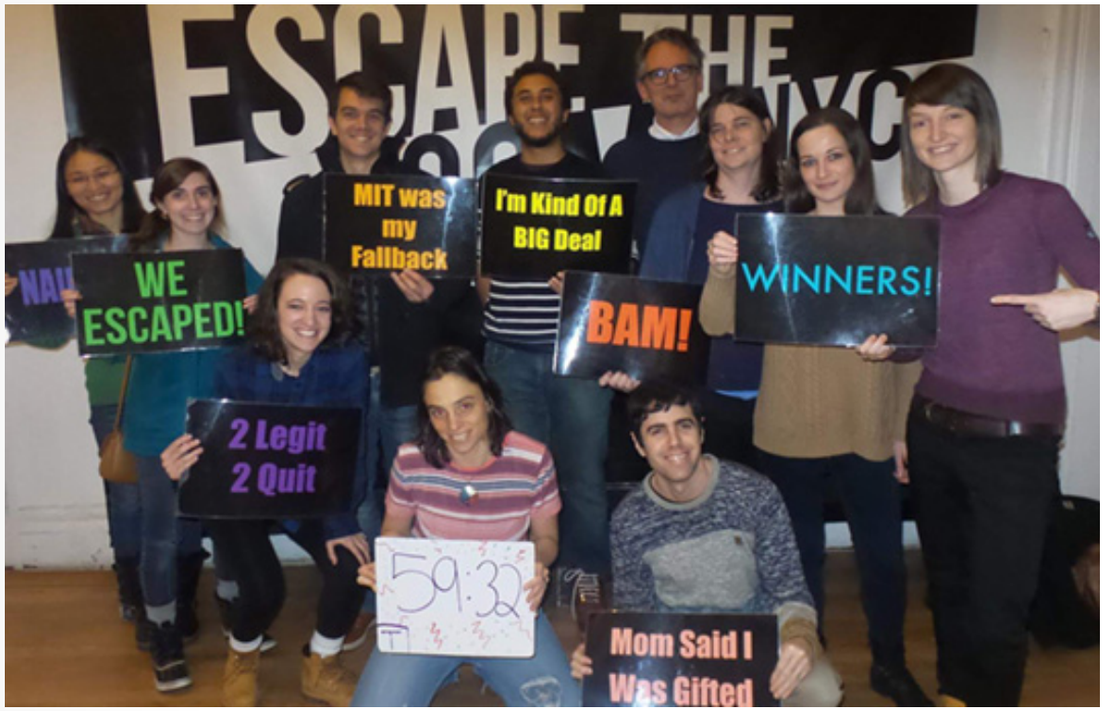
- My supervisors, **Alec Marantz** and **David Poeppel**, as well as everyone in the **Neuroscience of Language Lab** and **Poeppel Lab**!

جامعة نيويورك أبوظبي
NYU | ABU DHABI

NYU

Funding: G1001 Abu Dhabi Institute

laura.gwilliams@nyu.edu
@GwilliamsL



laura.gwilliams@nyu.edu
@GwilliamsL

Thank you!

NEW YORK UNIVERSITY