



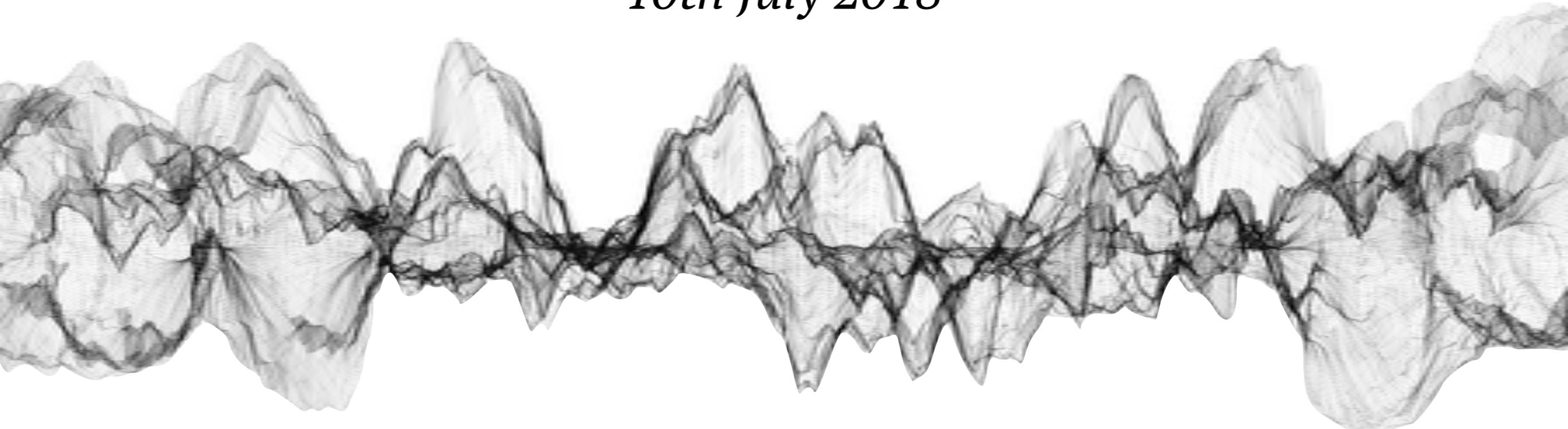
NEW YORK UNIVERSITY

# Towards a mechanistic account of speech comprehension in the human brain

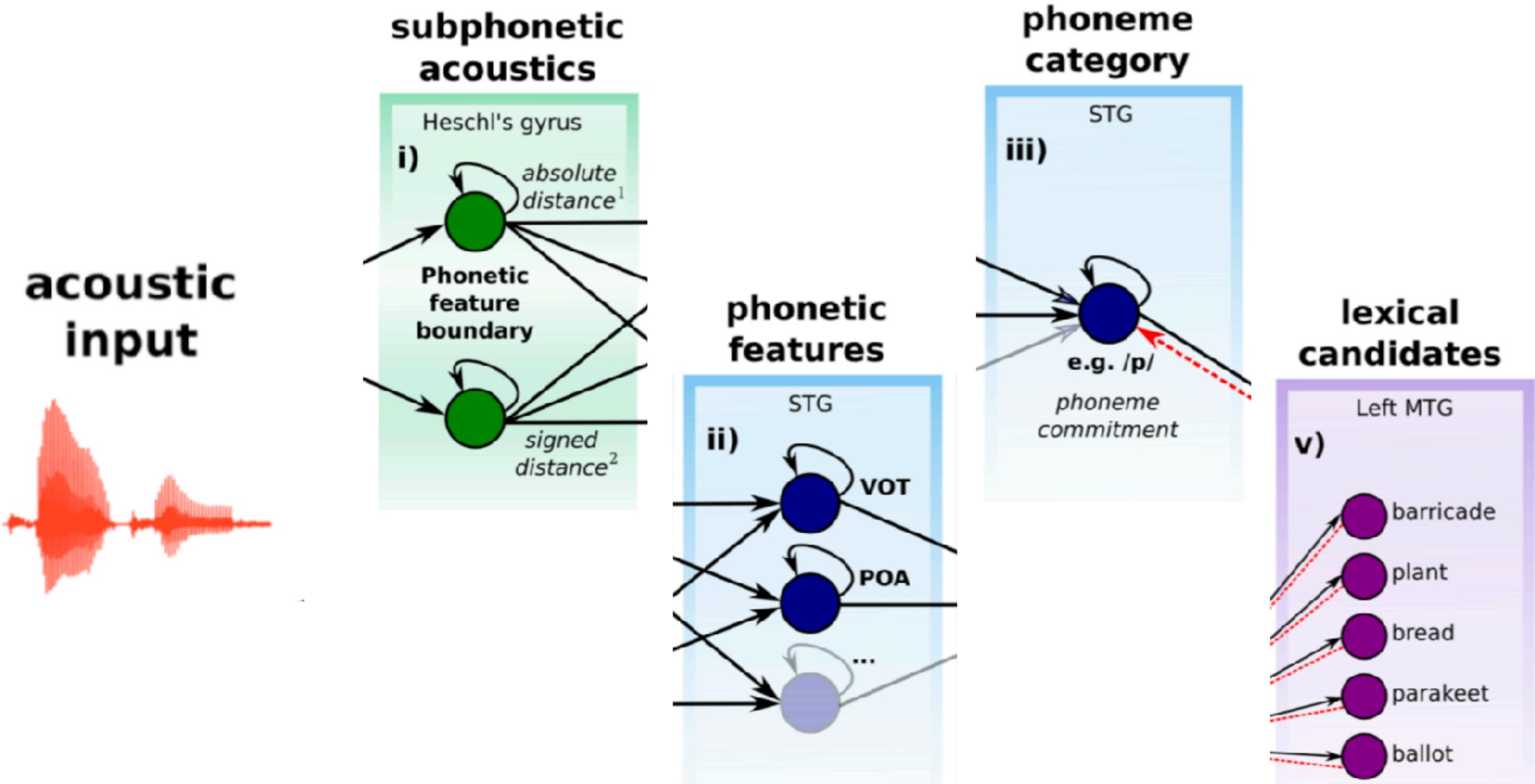
---

Laura Gwilliams

*10th July 2018*



# Putting together the processing pieces



# Roadmap

---

- Completed projects
  - **Bottom-up processes:** Transforming acoustic signal into discrete phonological categories
  - **Top-down processes:** Revising that categorisation based on subsequent context

b a

# Roadmap

---

- Completed projects
  - **Bottom-up processes:** Transforming acoustic signal into discrete phonological categories
  - **Top-down processes:** Revising that categorisation based on subsequent context

b a r a k ee t

# Roadmap

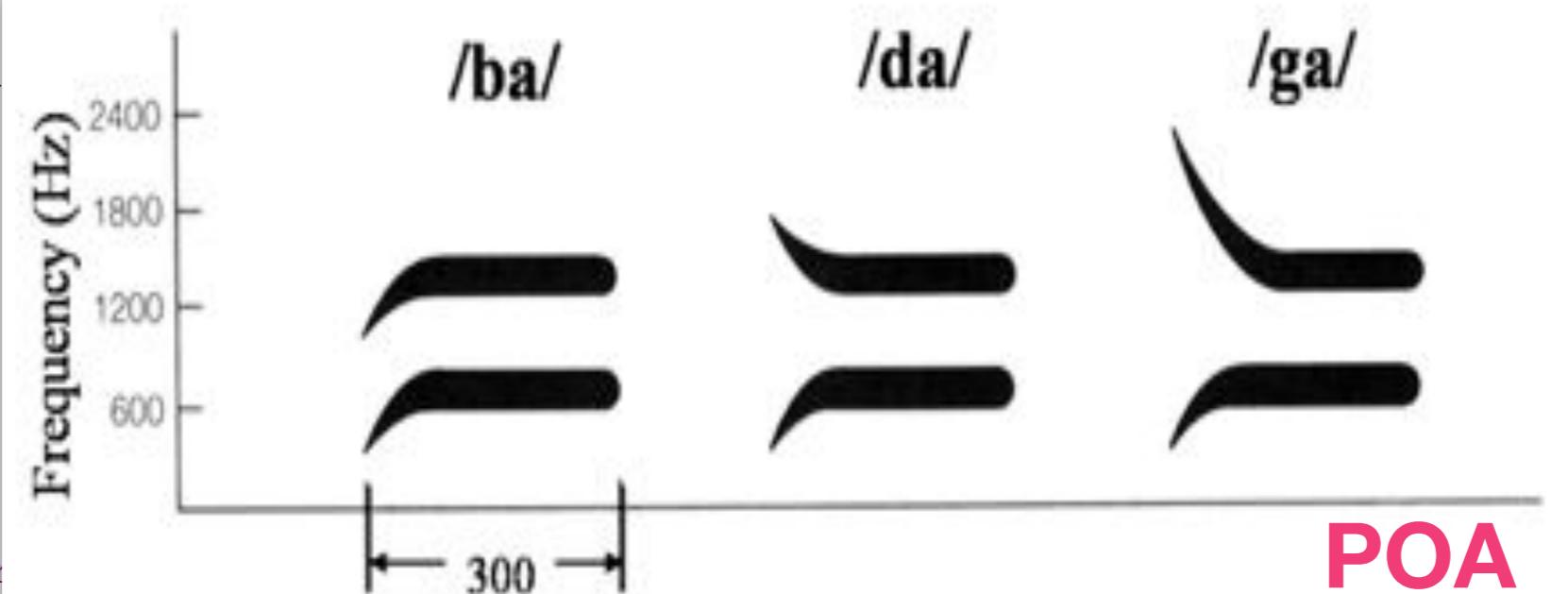
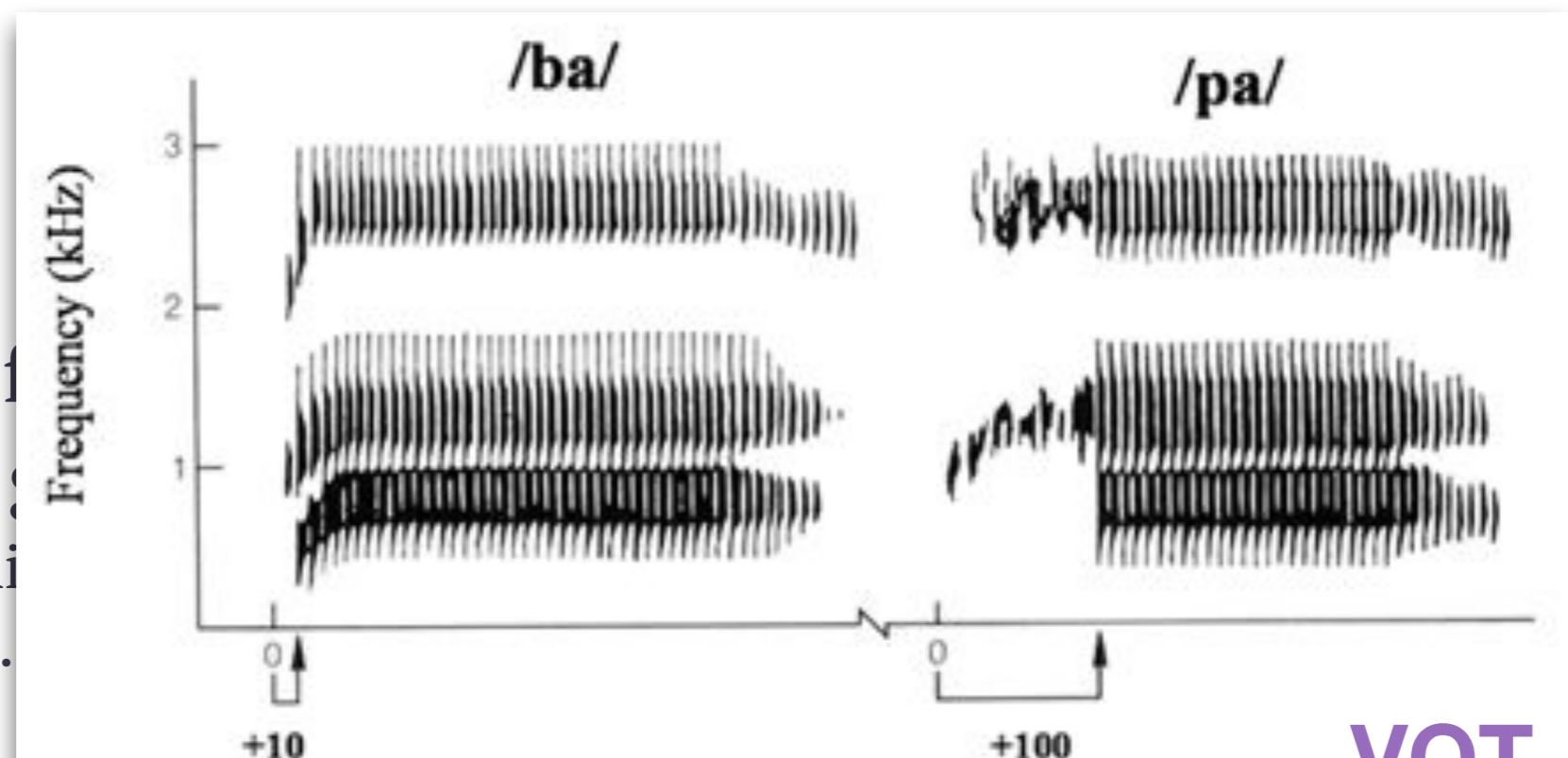
---

- Completed projects
  - **Bottom-up processes:** Transforming acoustic signal into discrete phonological categories
  - **Top-down processes:** Revising that categorisation based on subsequent context
- Future directions
  - Updating the mapping between acoustics to phonemes online through repeated **exposure** to accented speech
  - Testing re-activation processes in **natural speech**

# Bottom-up processing of phonemes

- Phonetic features are temporal features (Papen et al., 2003; Papanikos et al., 2014; Liberto et al., 2014)

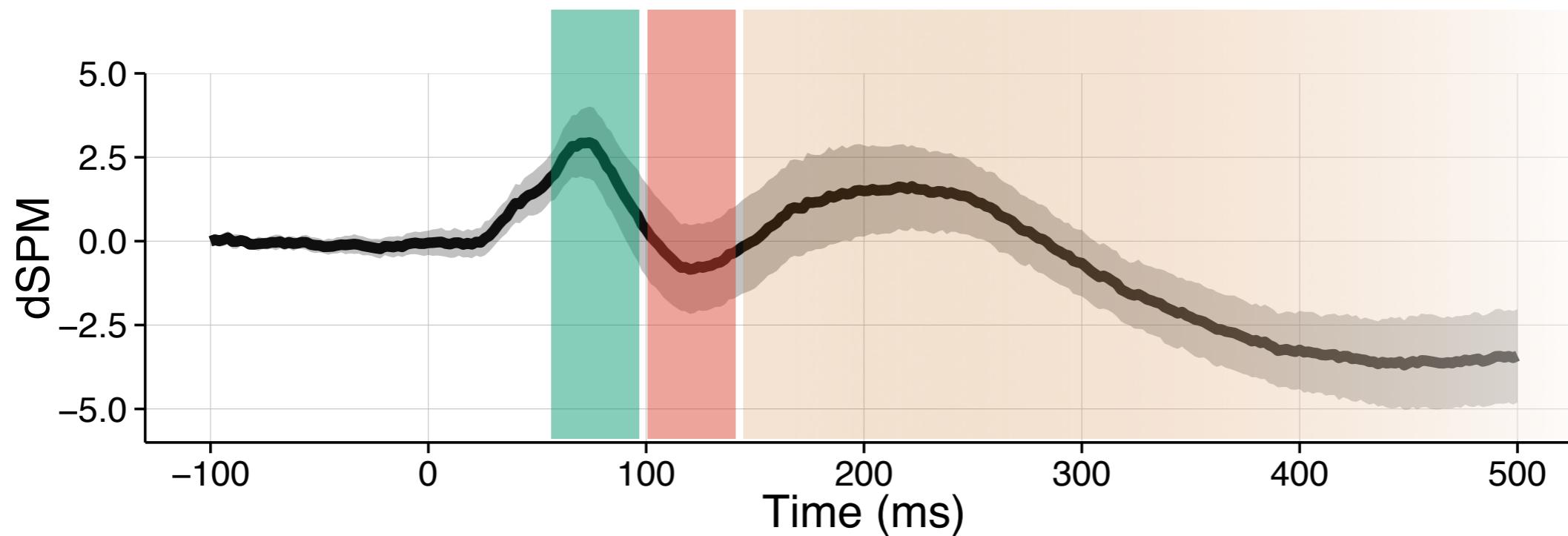
- This process is called bottom-up processing (Chang et al., 2014)



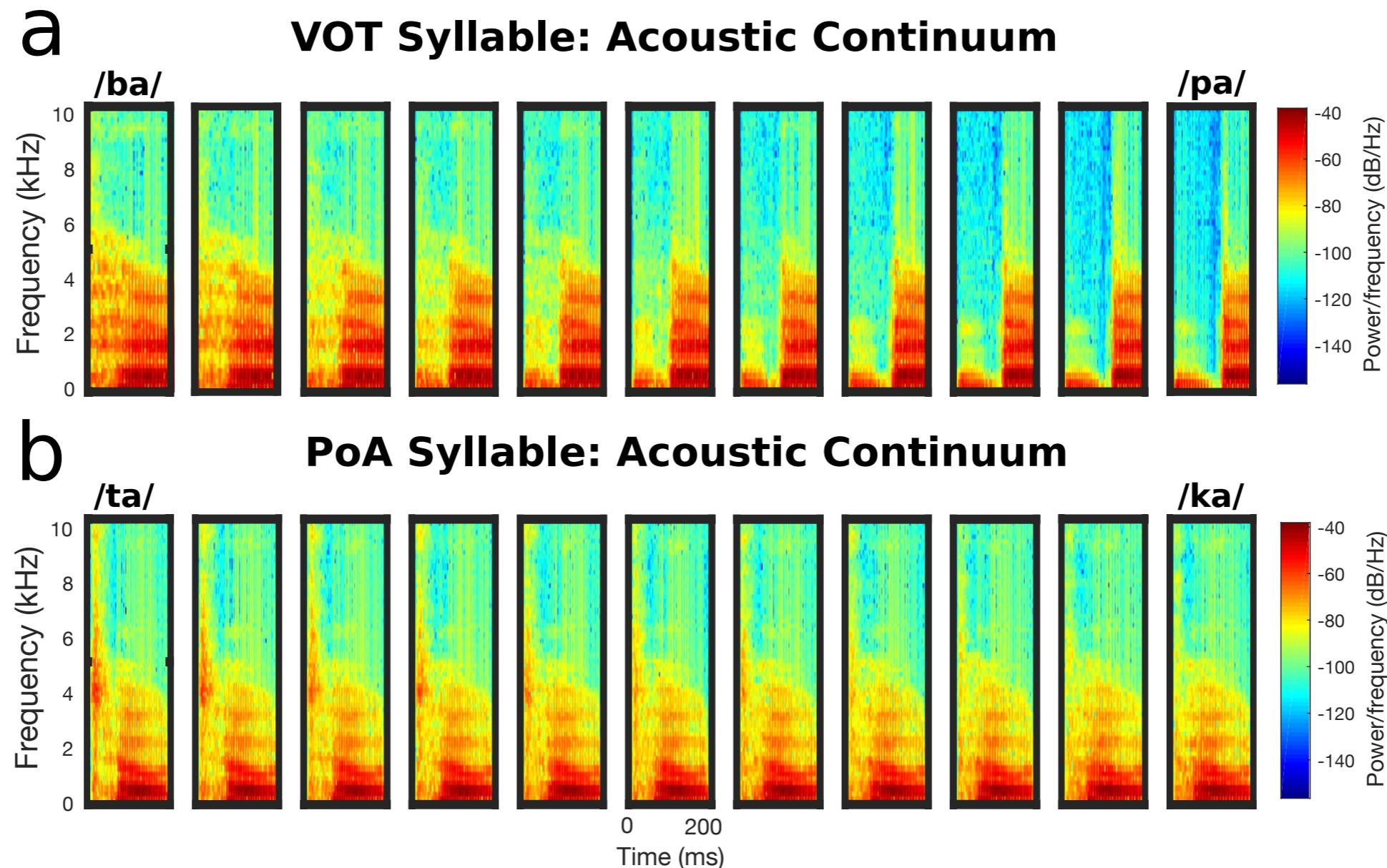
# Neutralising ambiguity

---

At what stage of processing is phonological ambiguity alleviated?



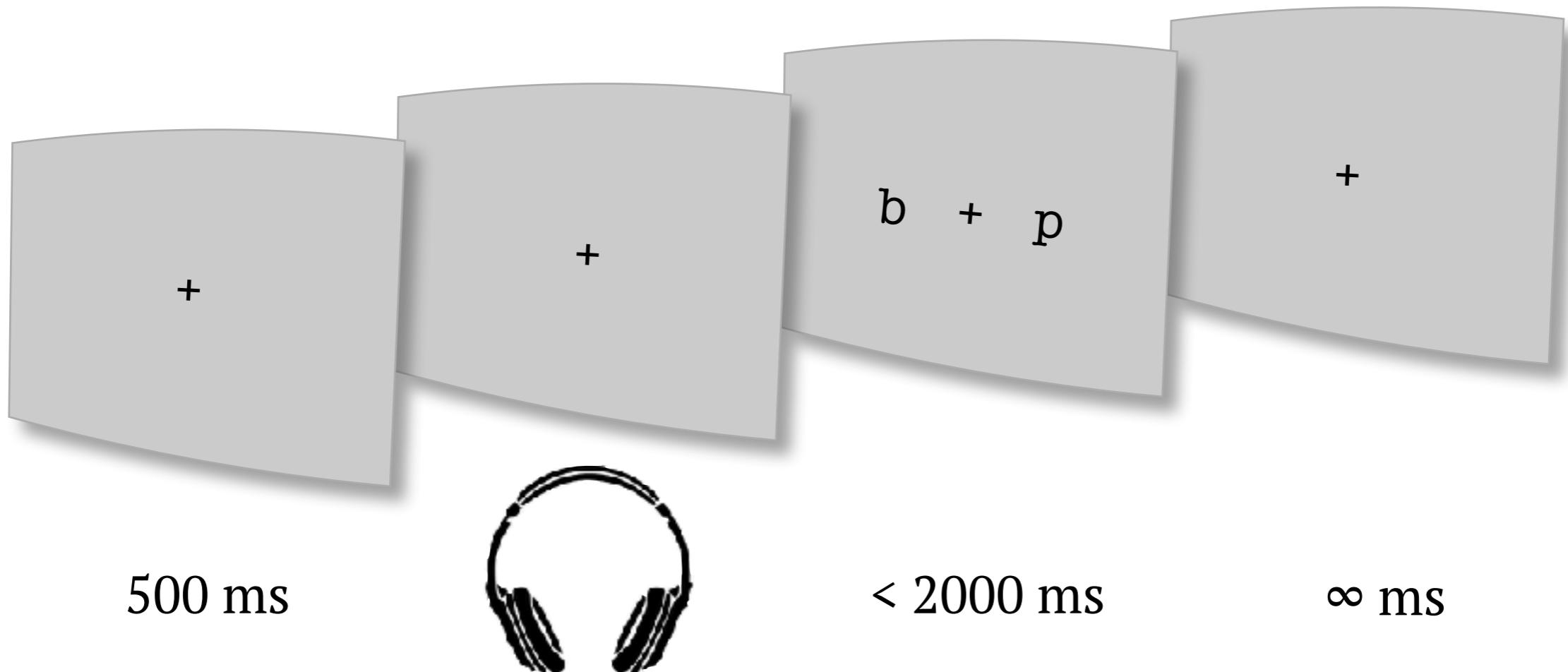
# Materials



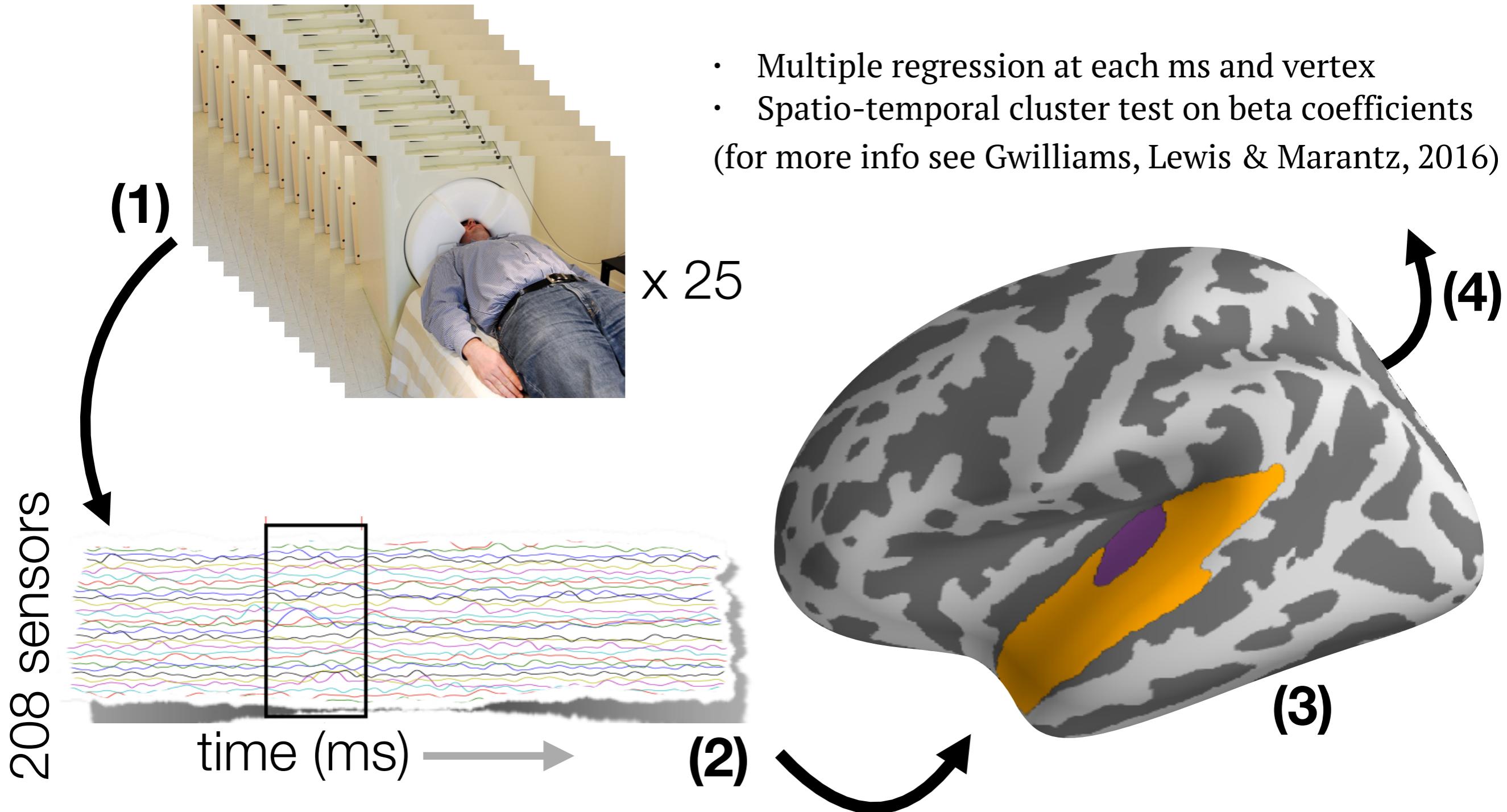
- VOT (31 pairs) {p-b, t-d, k-g} and POA (22 pairs) {t-k, p-t}

# Design

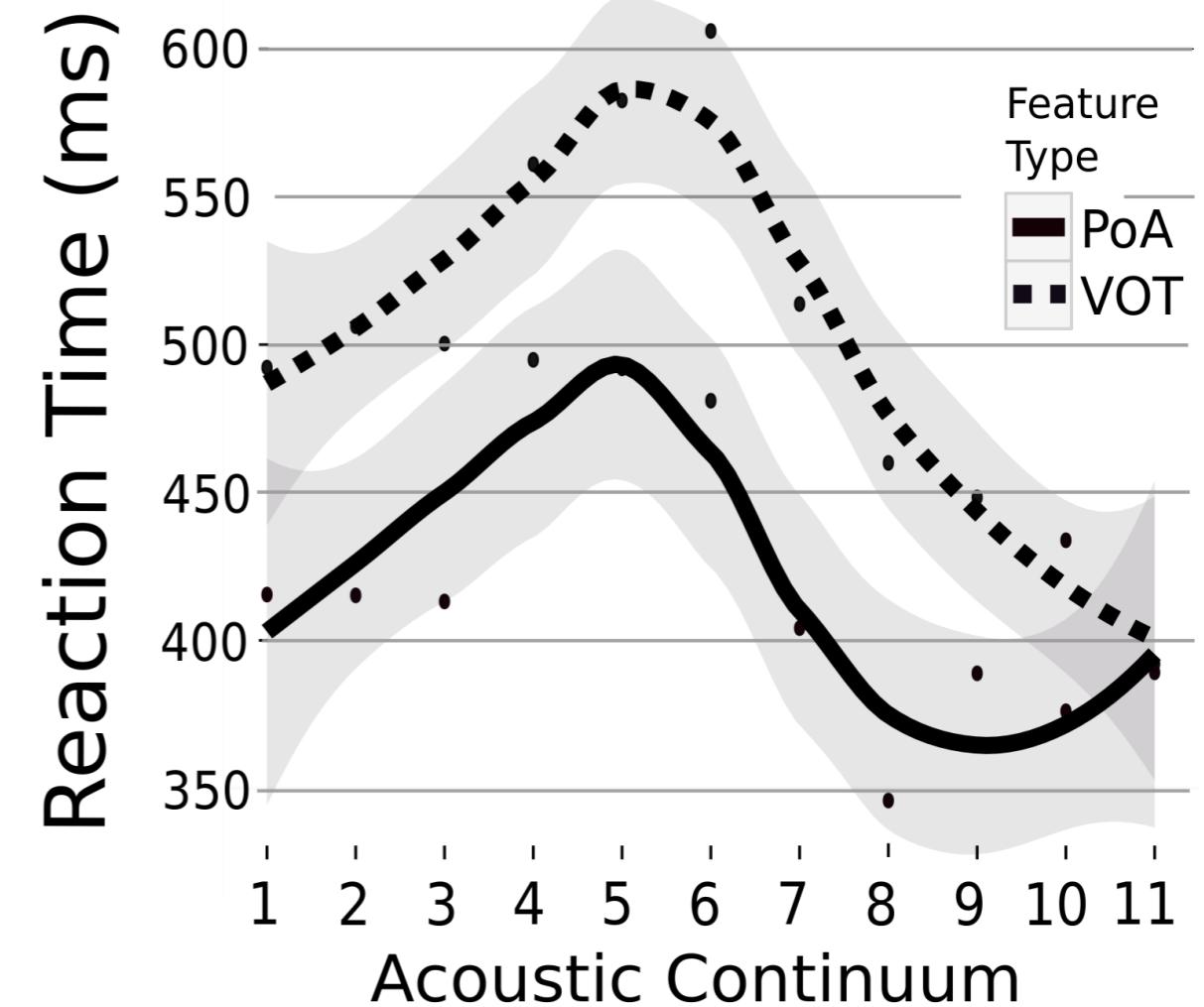
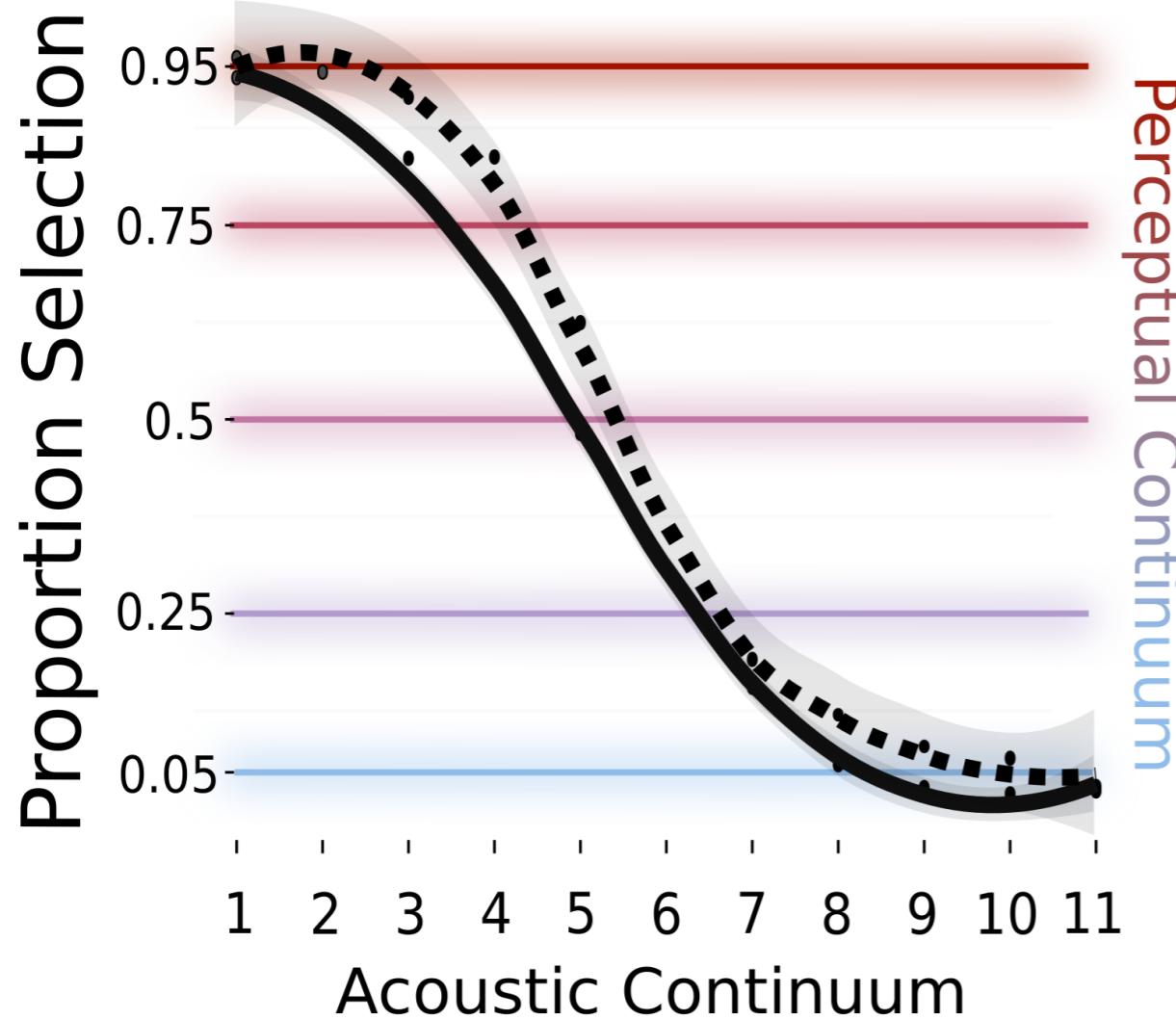
---



# Procedure & Analysis



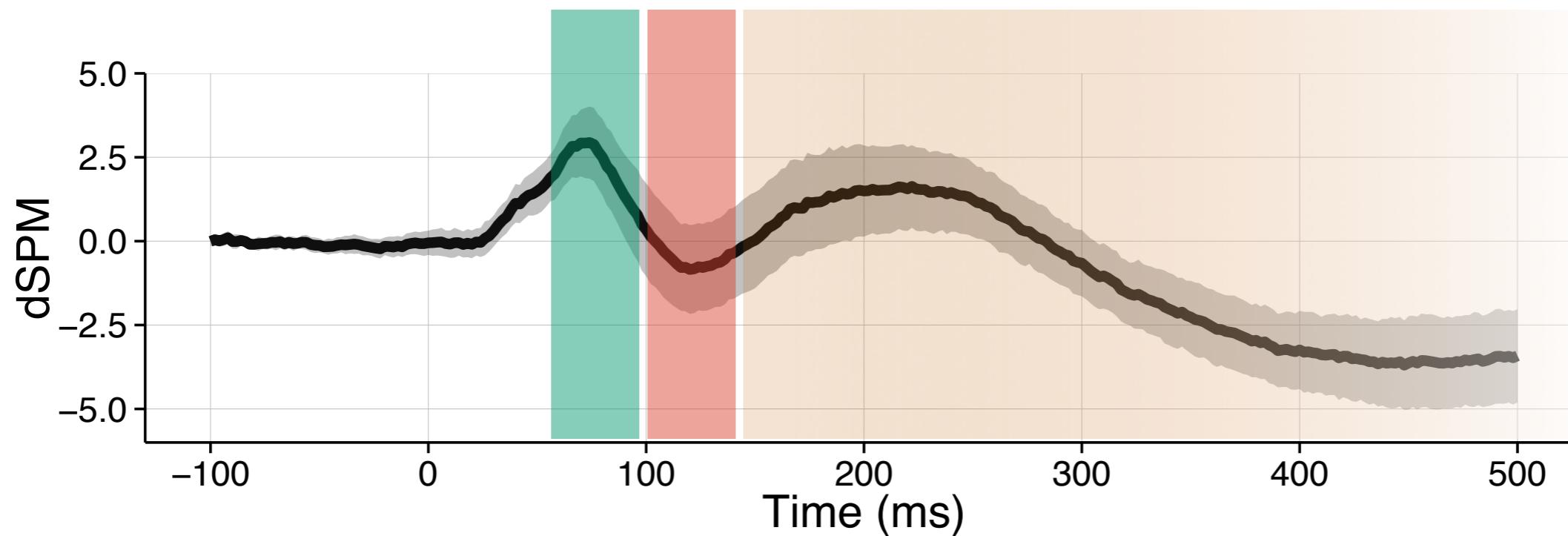
# Behaviour



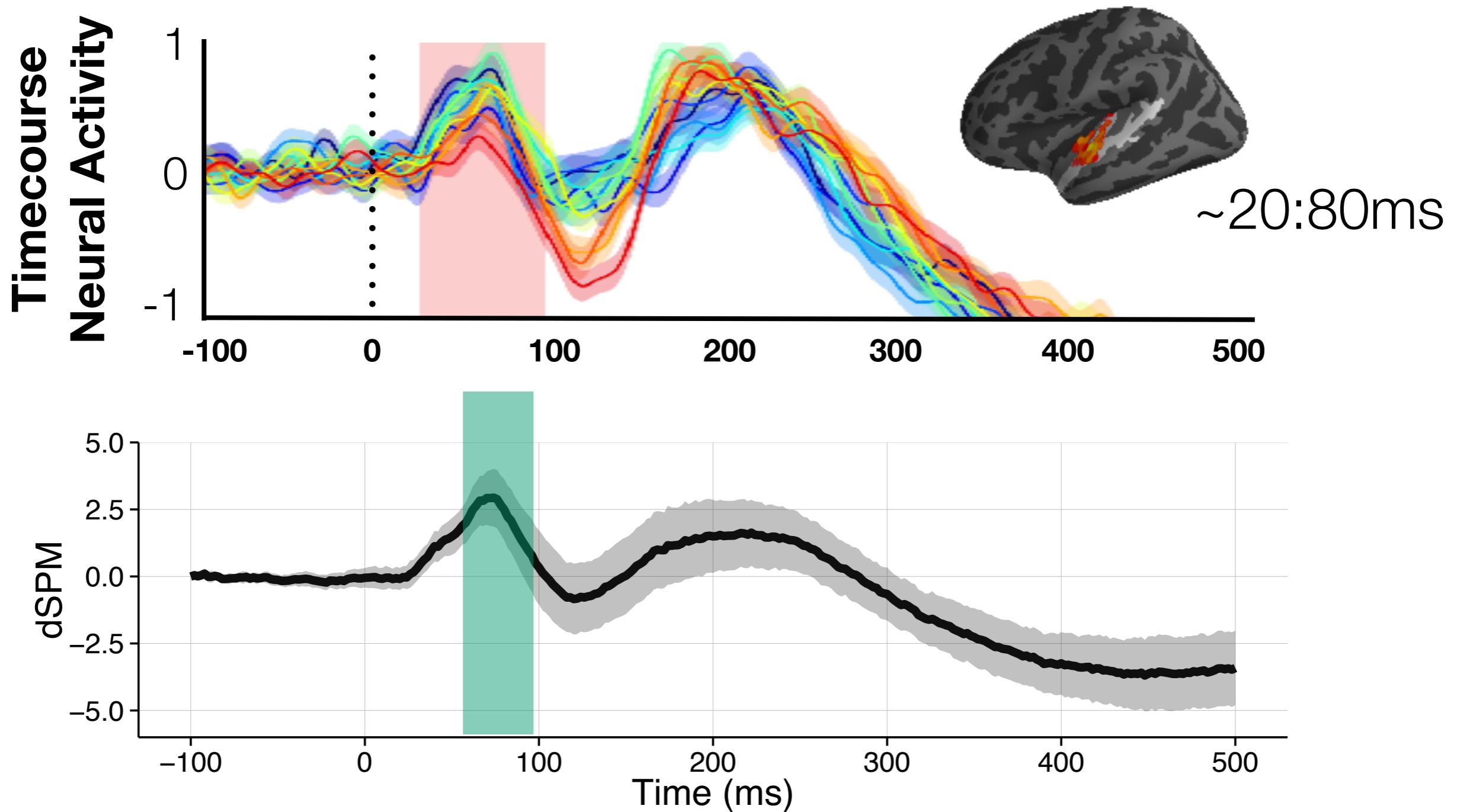
# Neutralising ambiguity

---

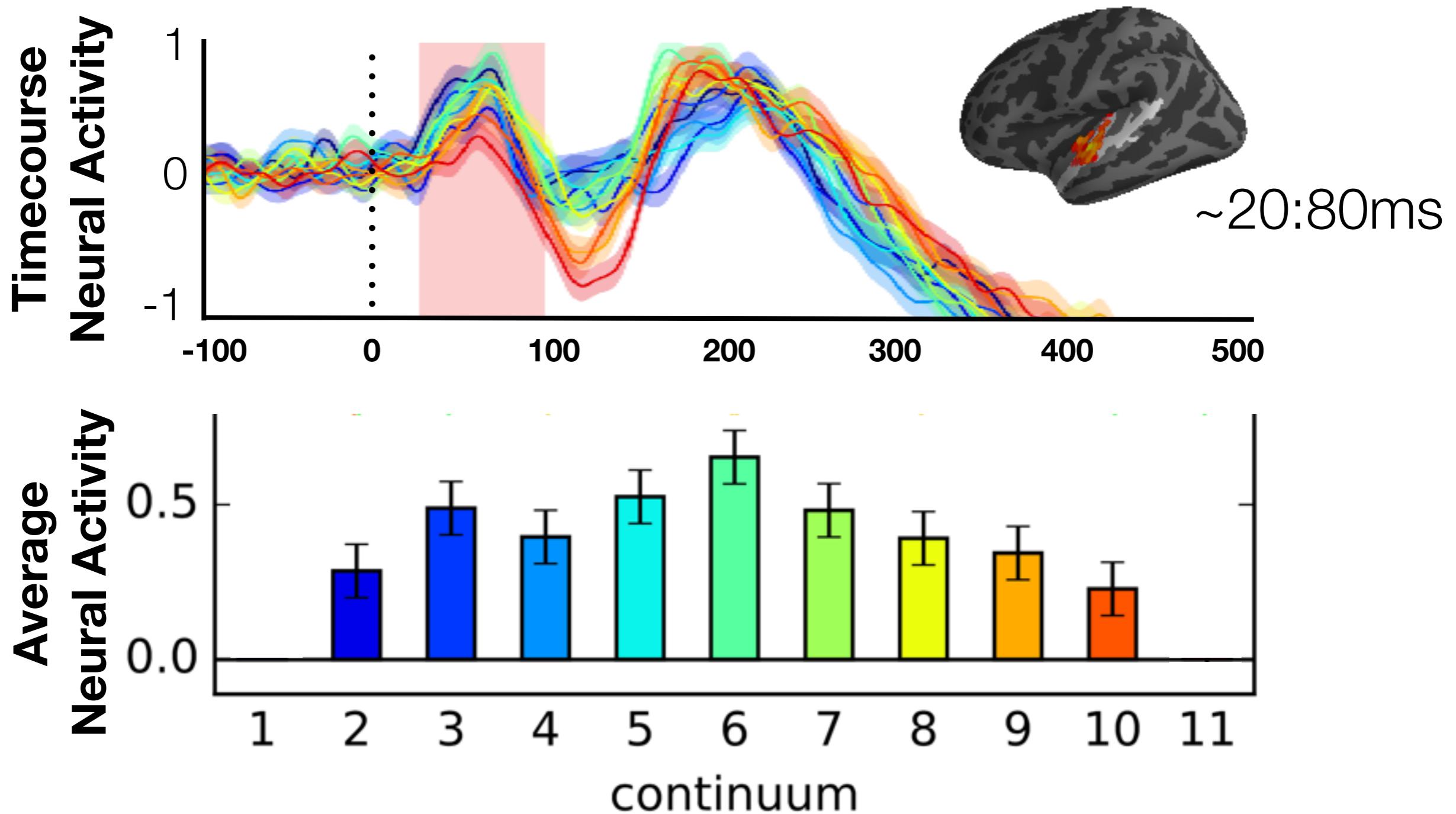
At what stage of processing is phonological ambiguity alleviated?



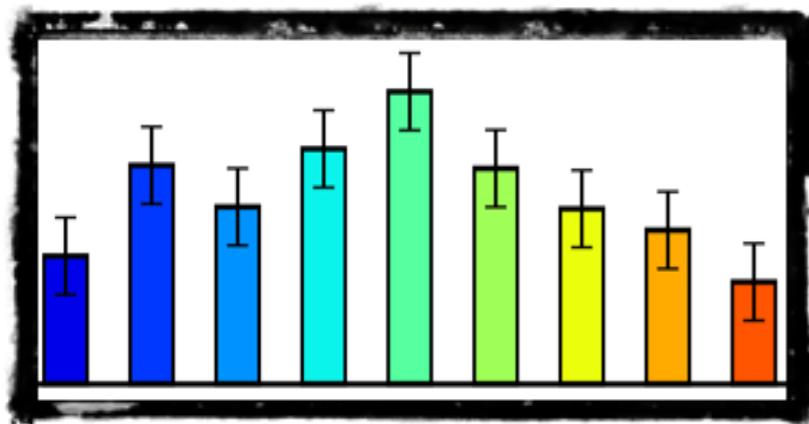
# Early ambiguity responses in Heschl's gyrus



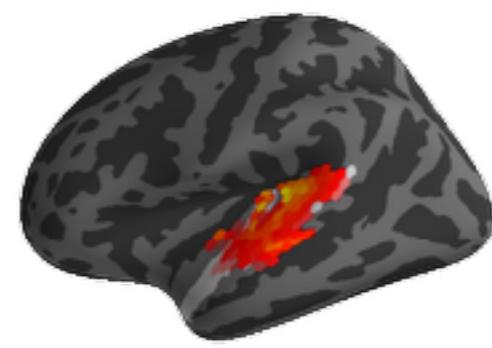
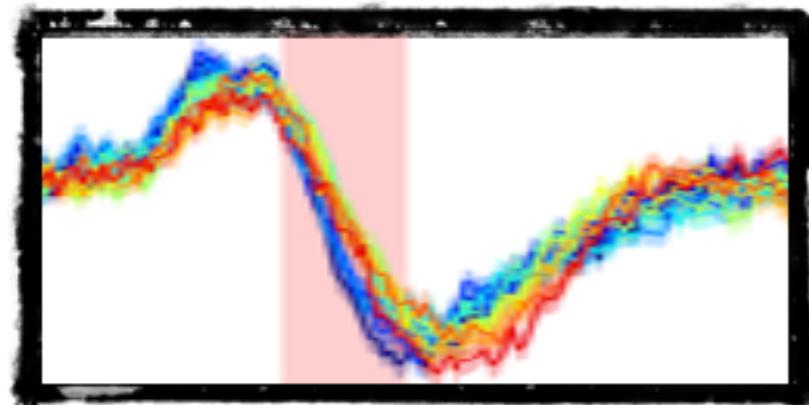
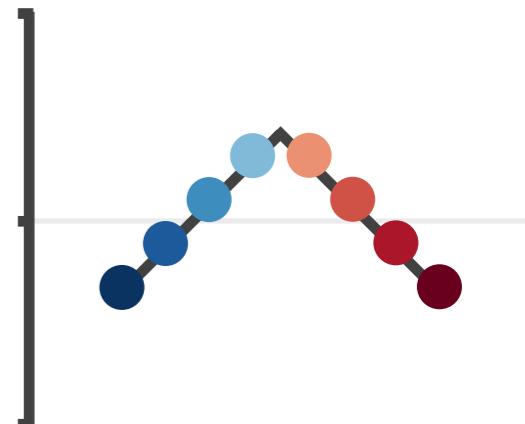
# Early ambiguity responses in Heschl's gyrus



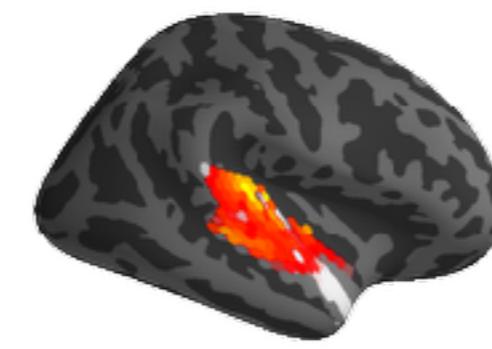
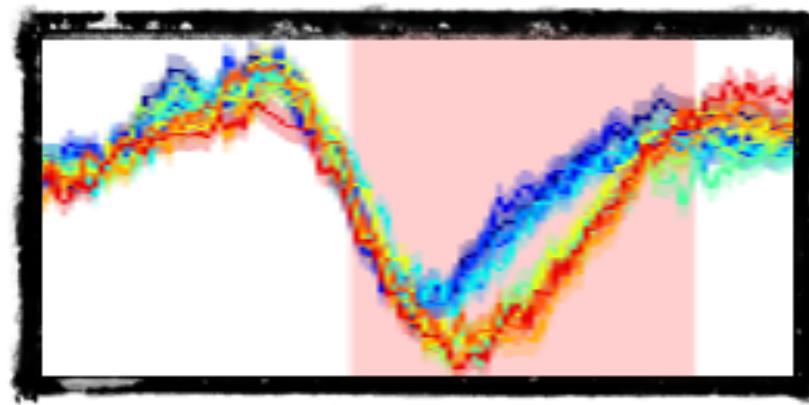
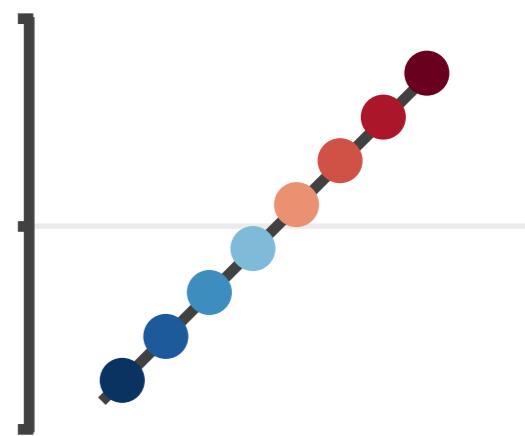
# Replicating the categorical trajectory using MEG



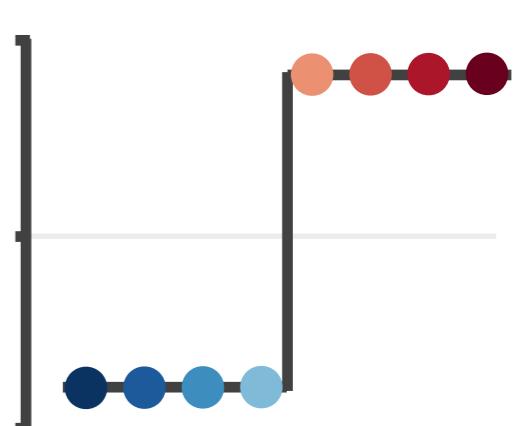
~20:80ms  
ambiguity



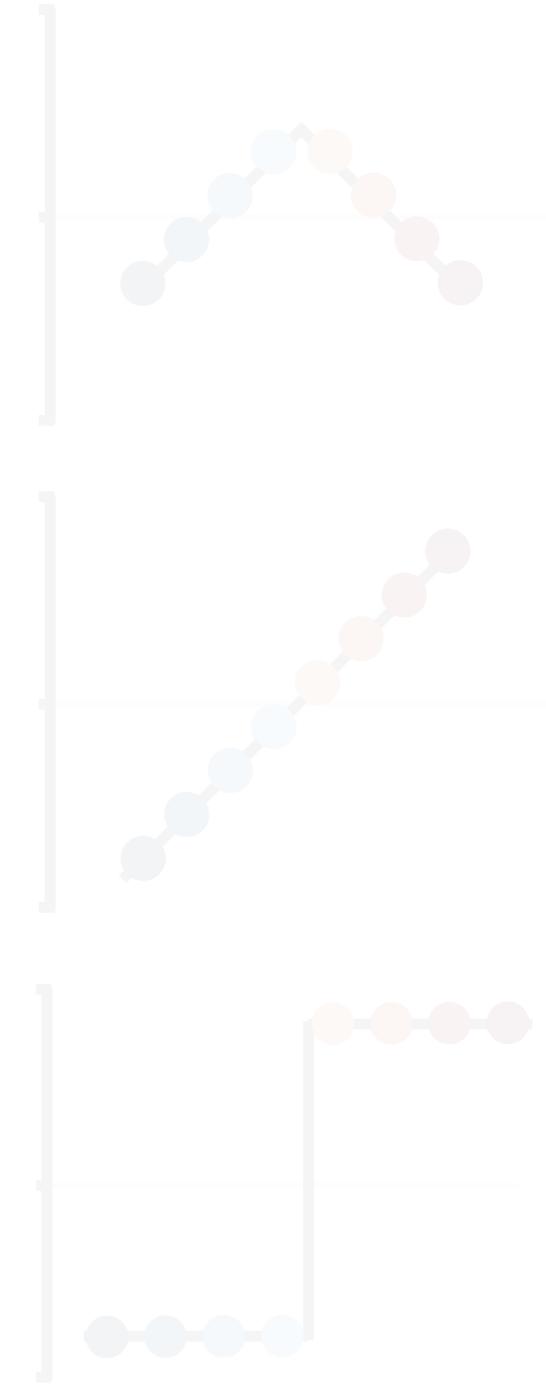
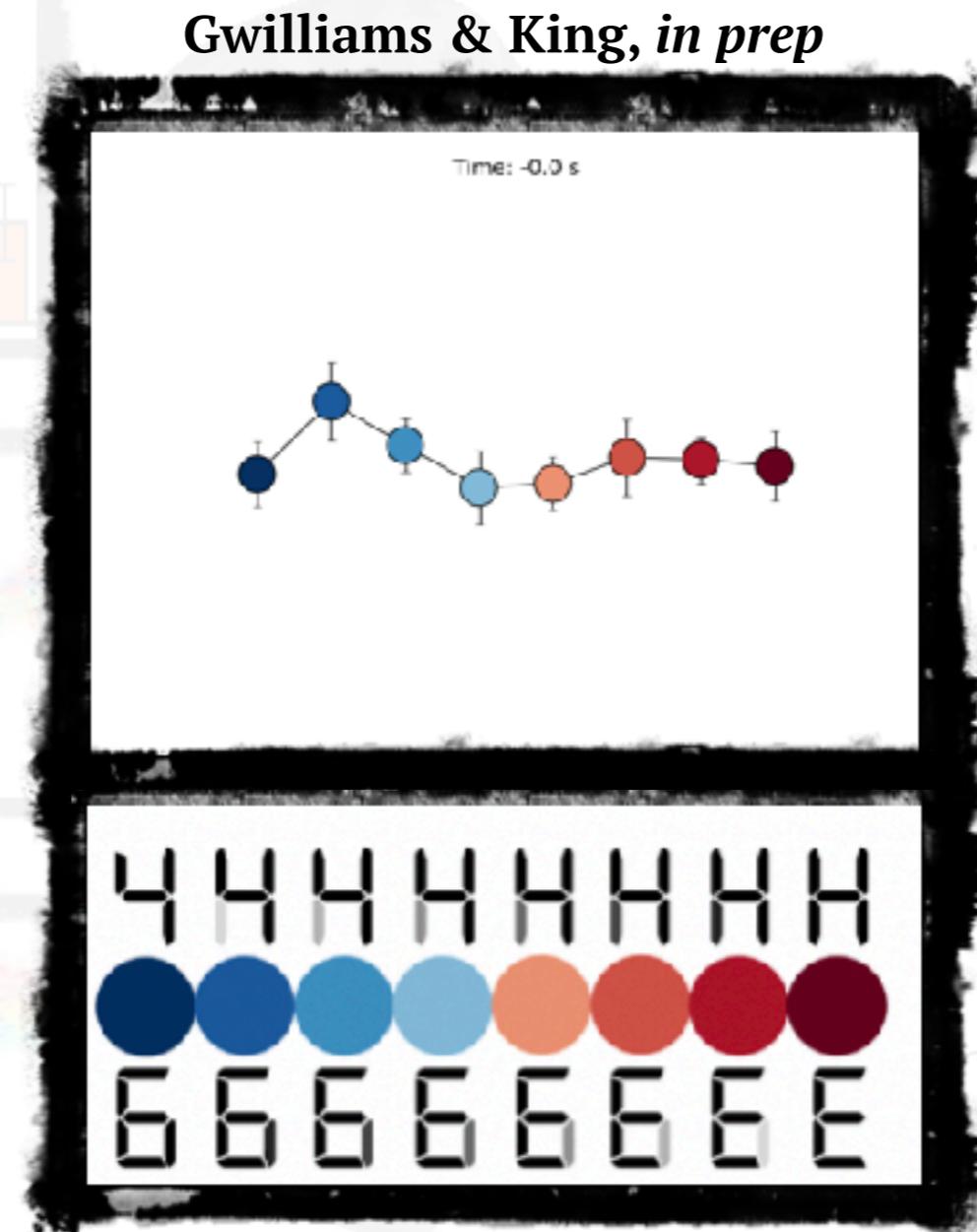
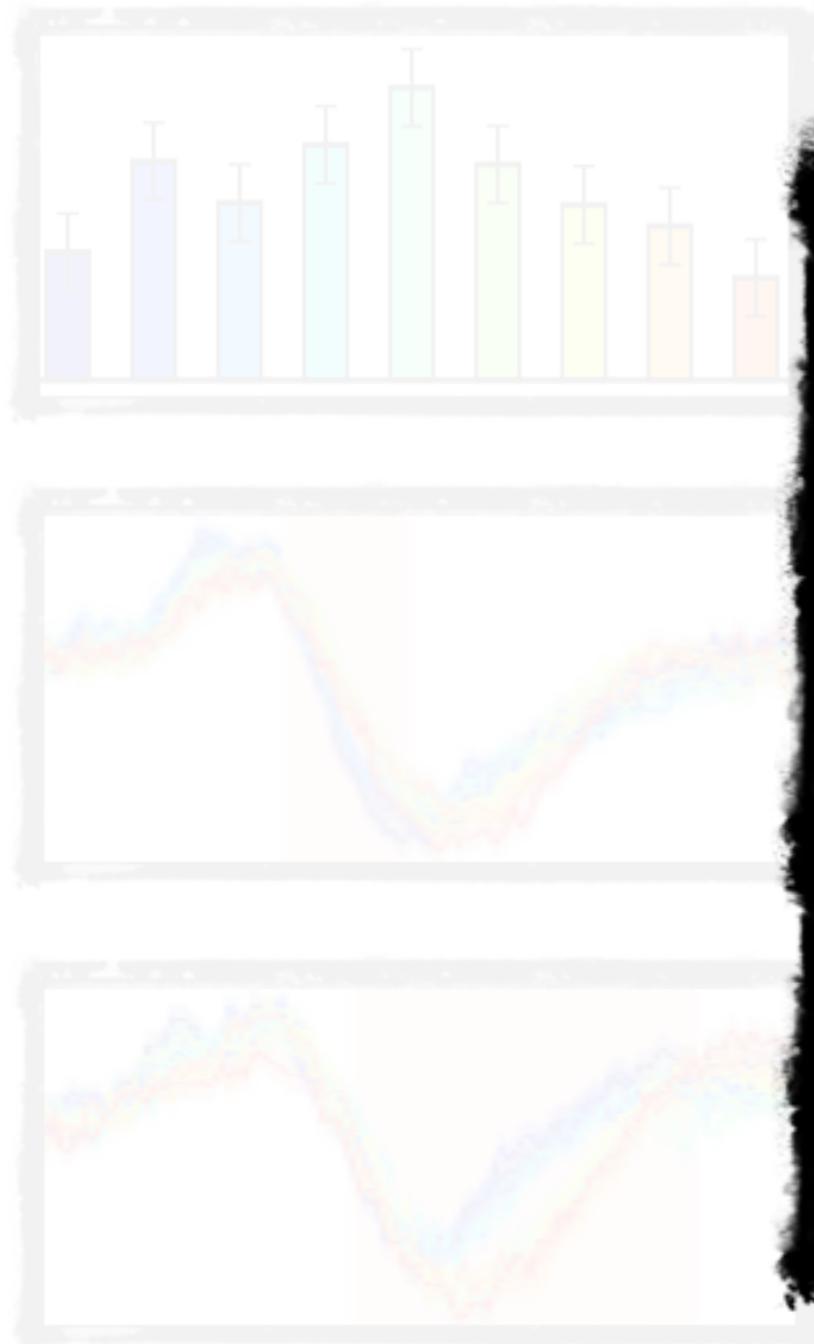
~80:150ms  
linear



~100:200ms  
categorical

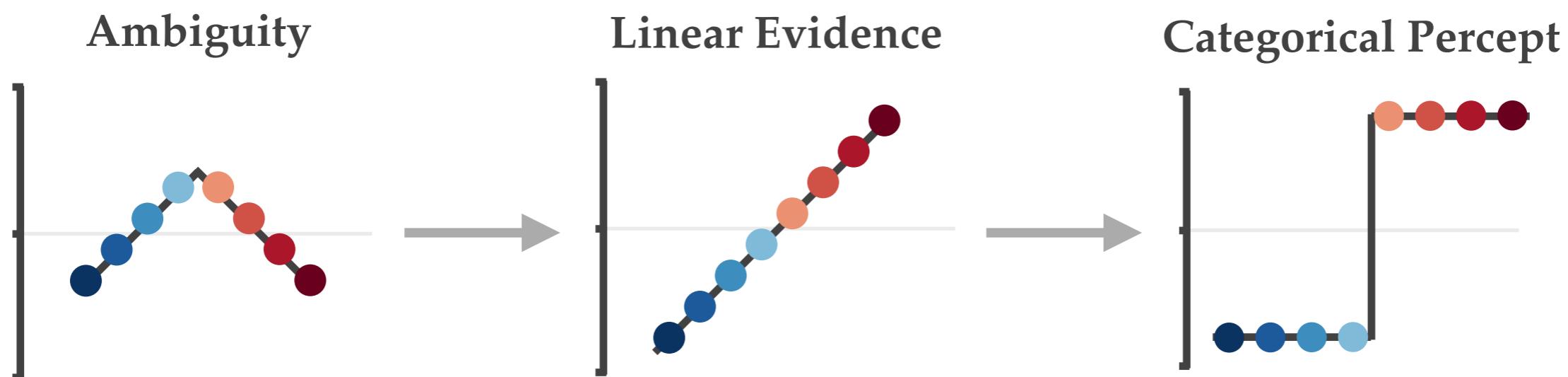


# Interesting links in a different domain

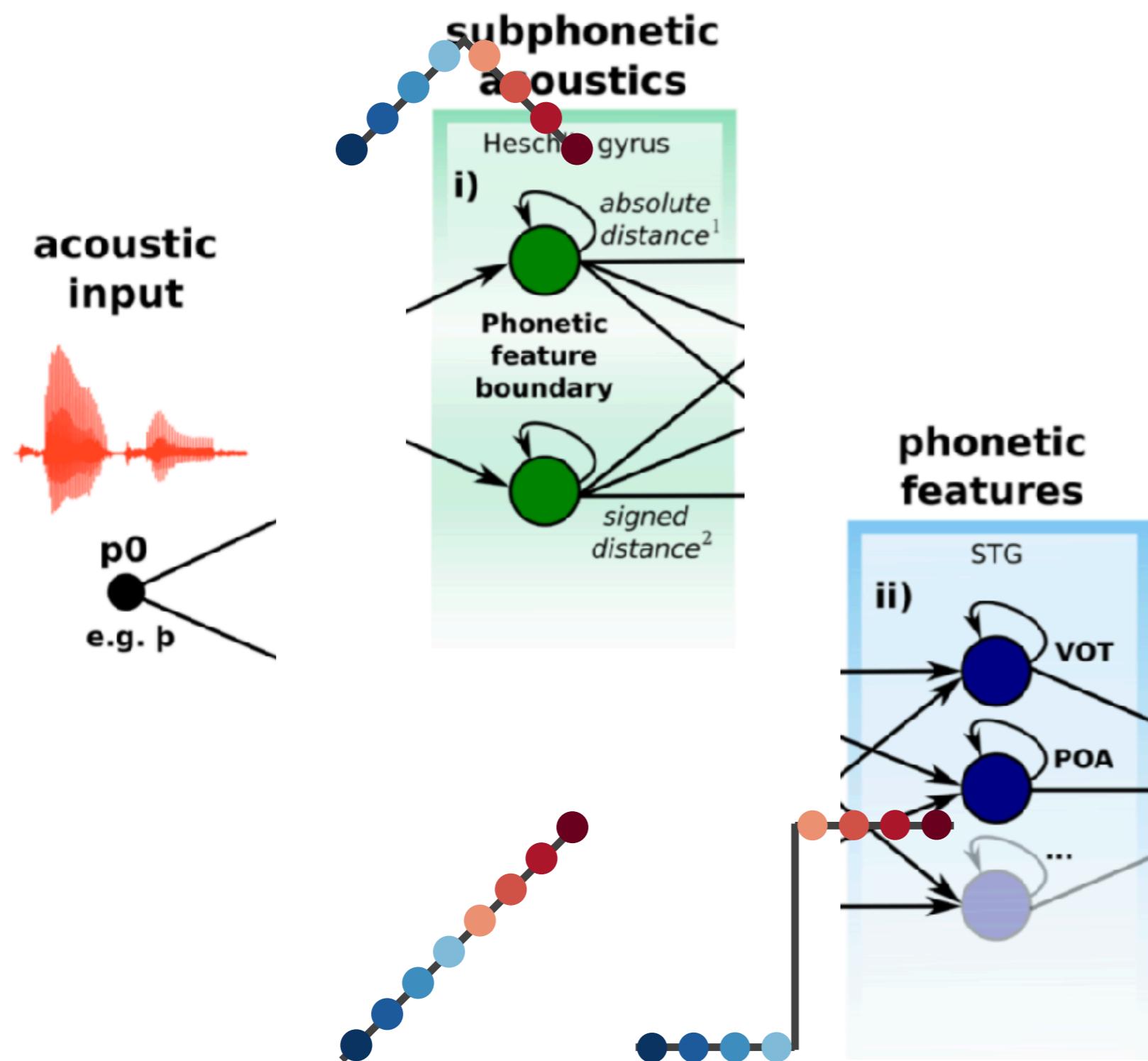


# Experiment 1 Conclusions

- Responses shift from being modulated **linearly** to being modulated **categorically** (domain general?)
- Very **early sensitivity to phonological boundaries** in left Heschl's gyrus – occurs *before* categorisation (speech specific?)



# Putting together the processing pieces



# Roadmap

---

- Completed projects
  - **Bottom-up processes:** Transforming acoustic signal into discrete phonological categories
  - **Top-down processes:** Revising that categorisation based on subsequent context

b a

# Roadmap

---

- Completed projects
  - **Bottom-up processes:** Transforming acoustic signal into discrete phonological categories
  - **Top-down processes:** Revising that categorisation based on subsequent context

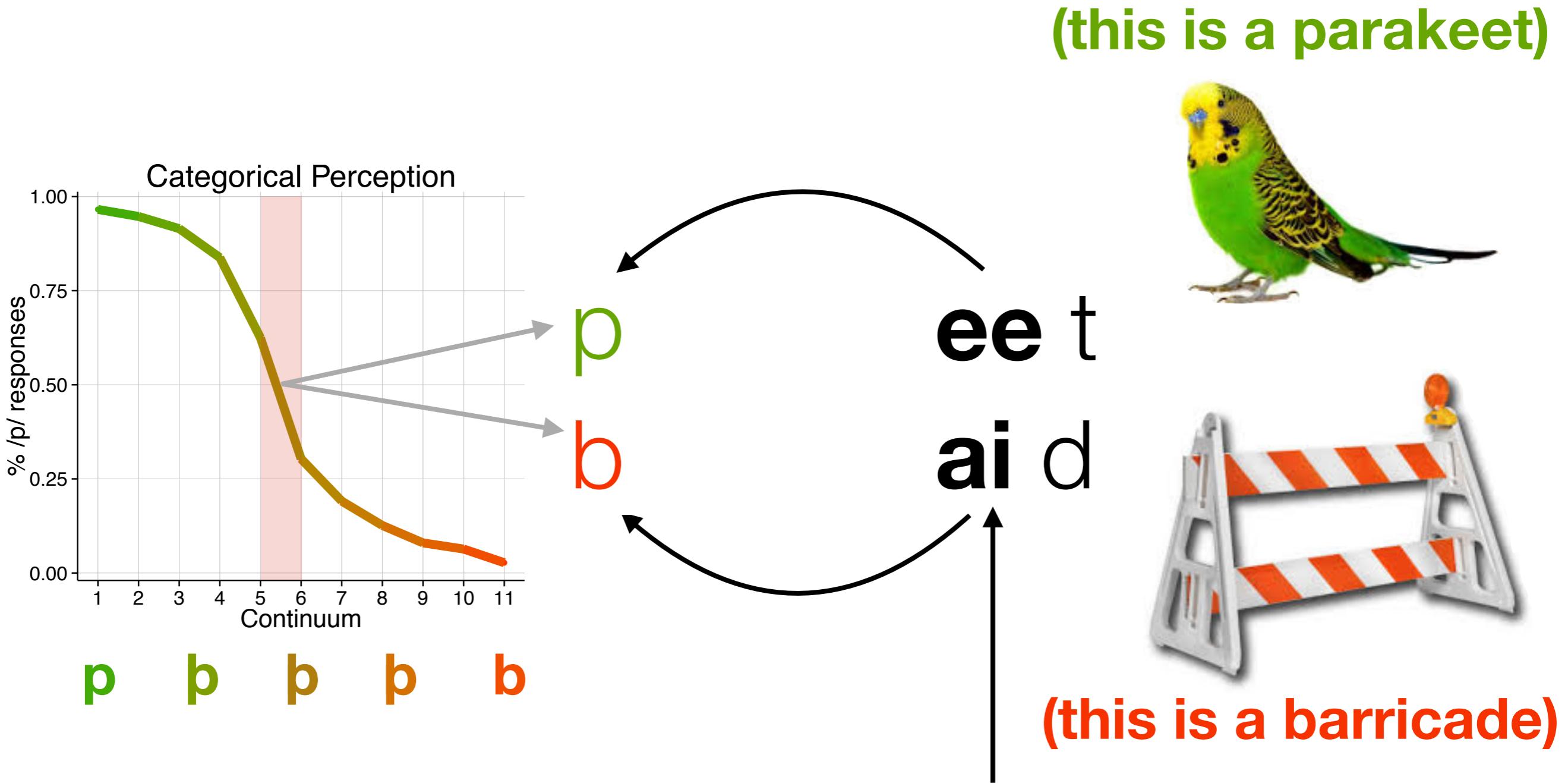
b a r a k ee t

# Top-down Influences on Perception

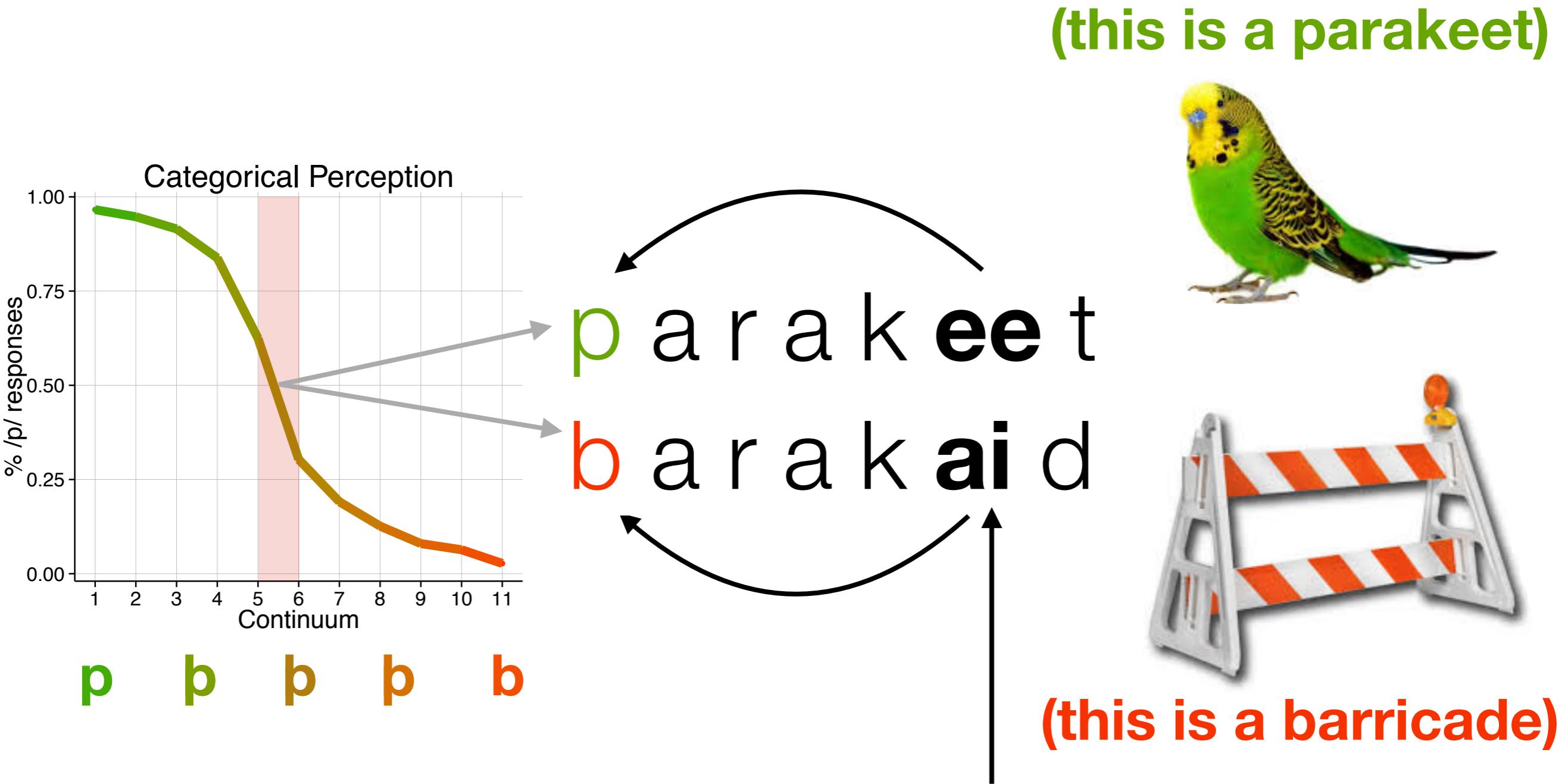
---

- Context occurring *after* an acoustic signal can be integrated to **update 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



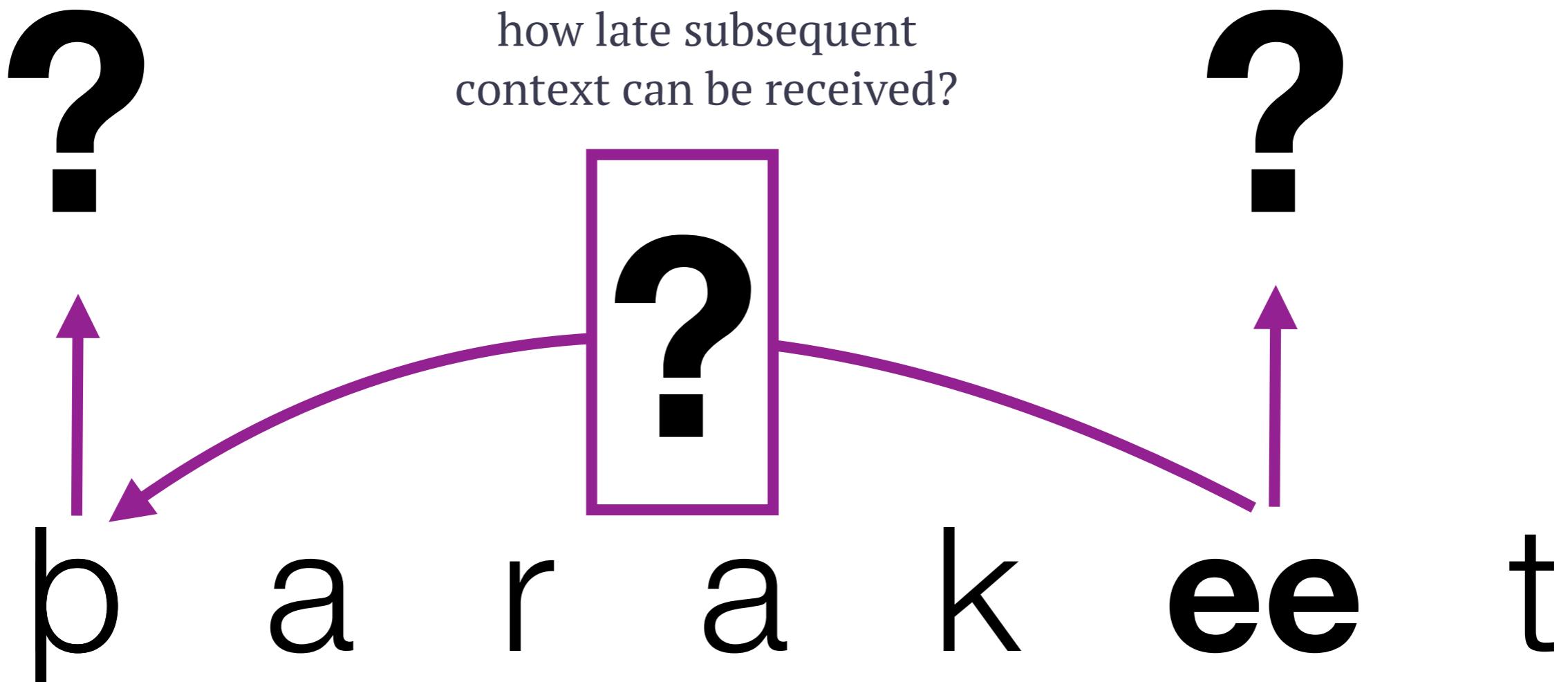
# Today's Questions

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

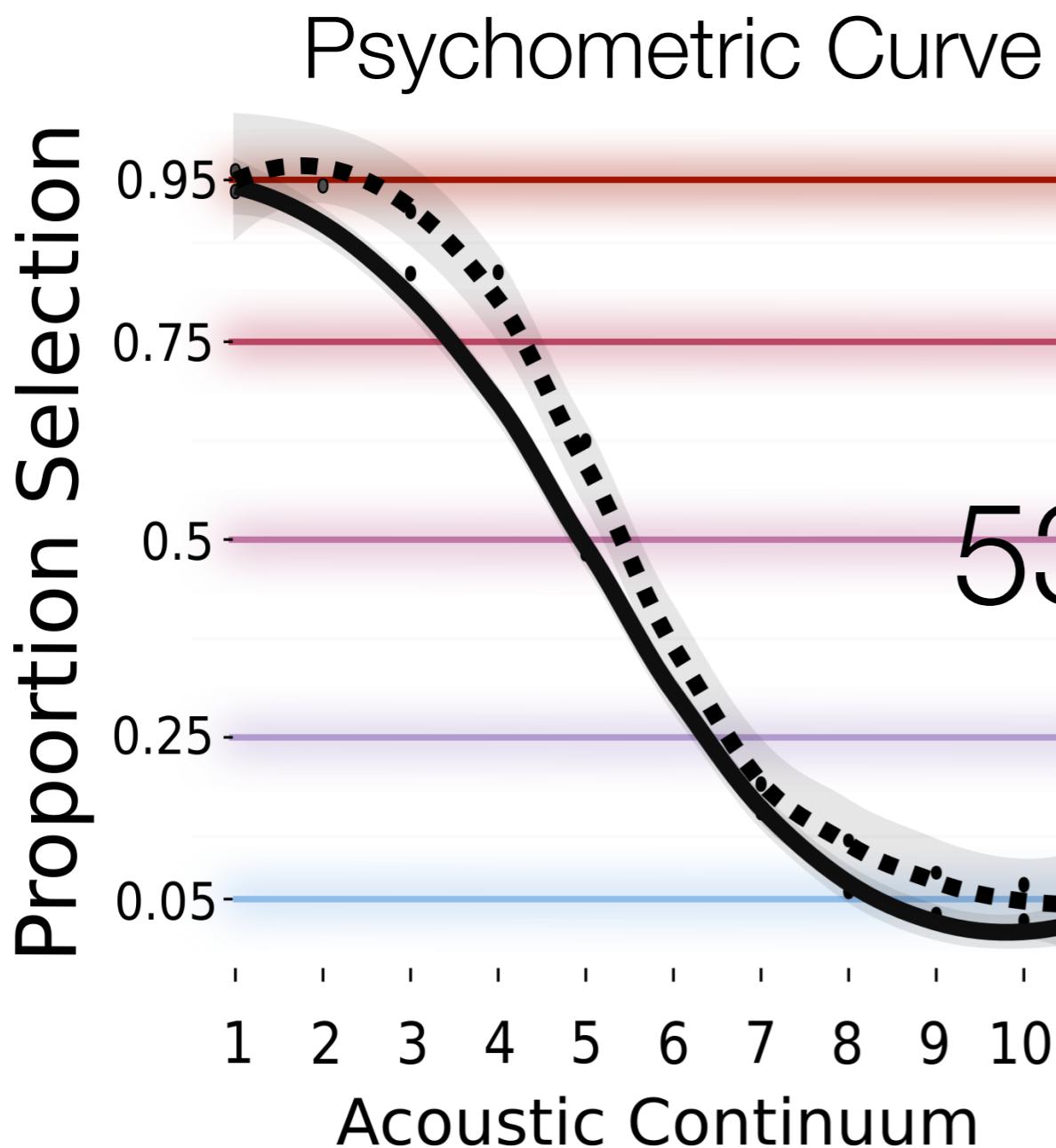
i could focus on the ambiguity resolution part here, rather than the original response to ambiguity. then, tie in the ambiguity response part later, linking it with AI?

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

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

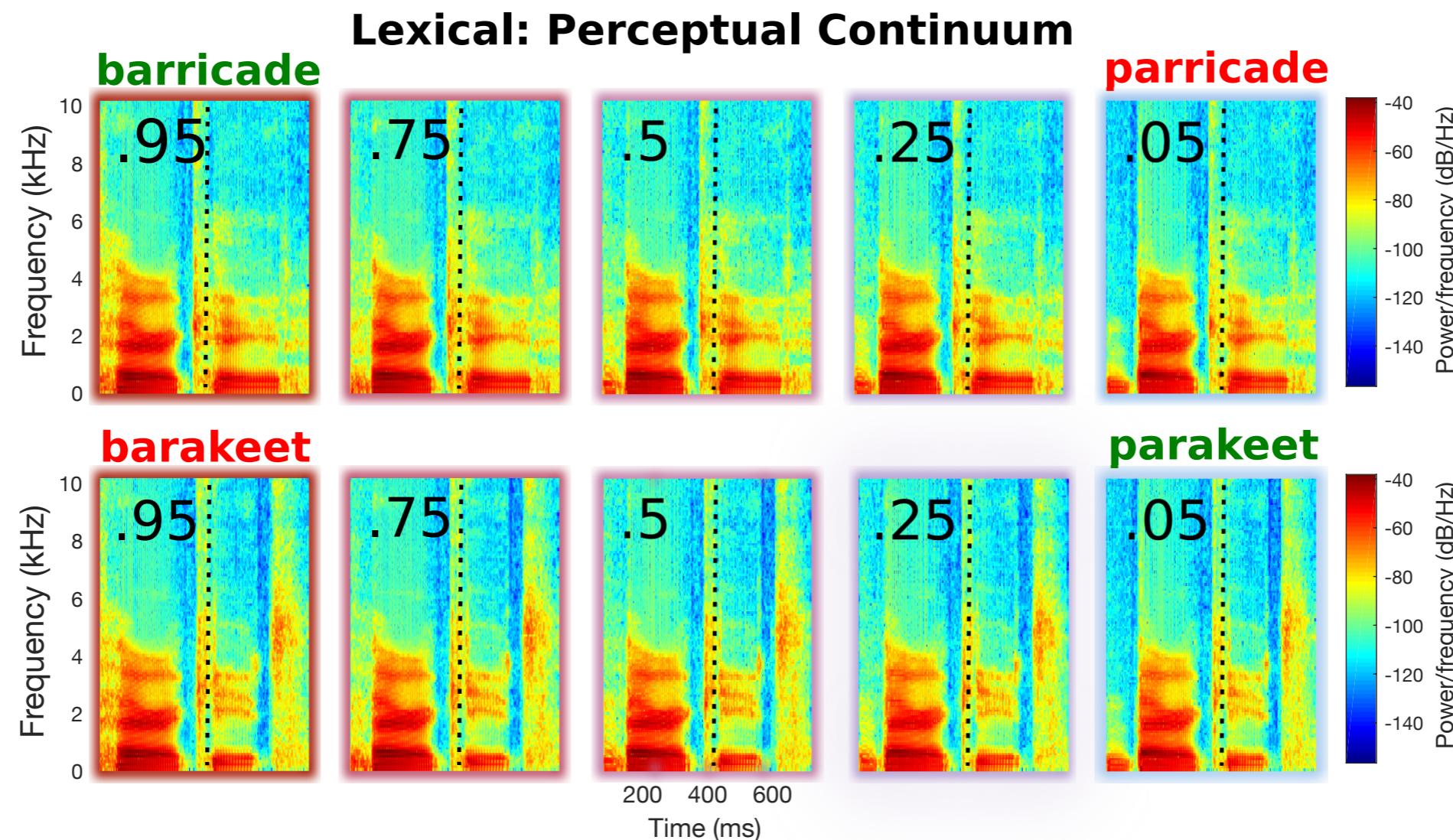


# Design & Materials



p p b b b

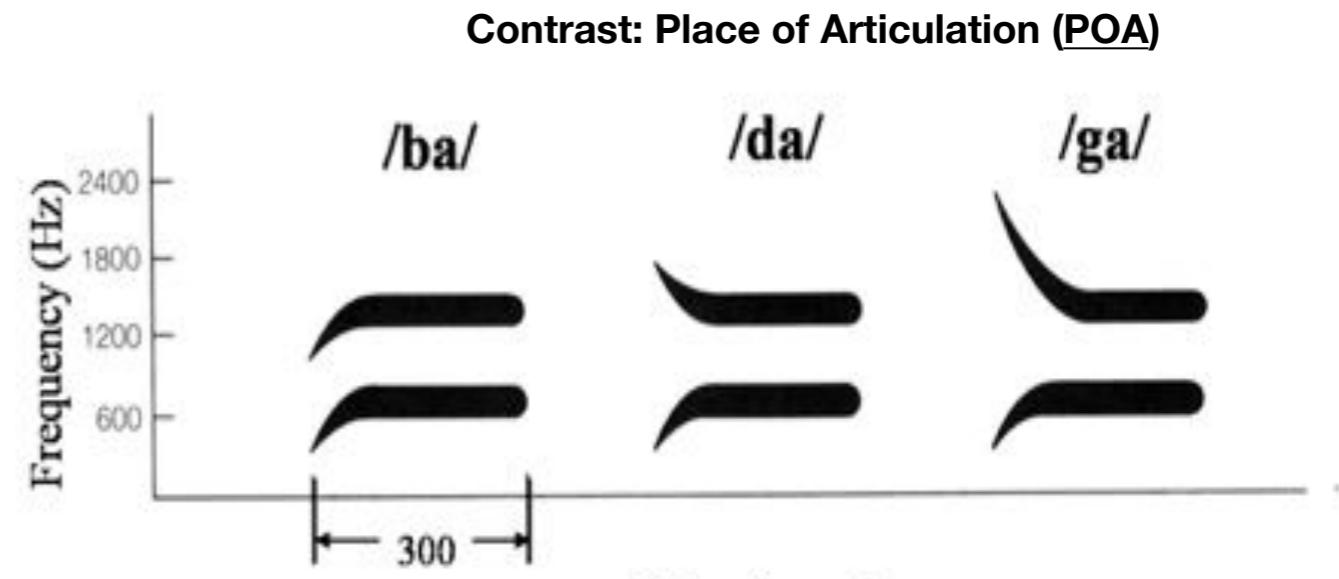
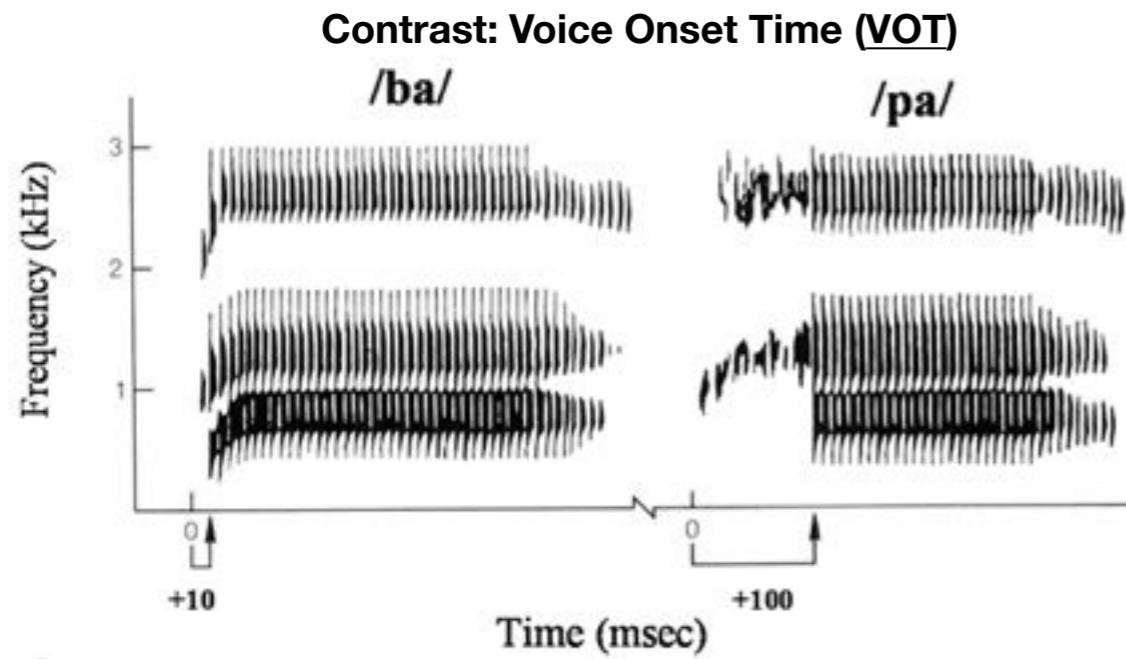
# Design & Materials



- 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}

p b b b b

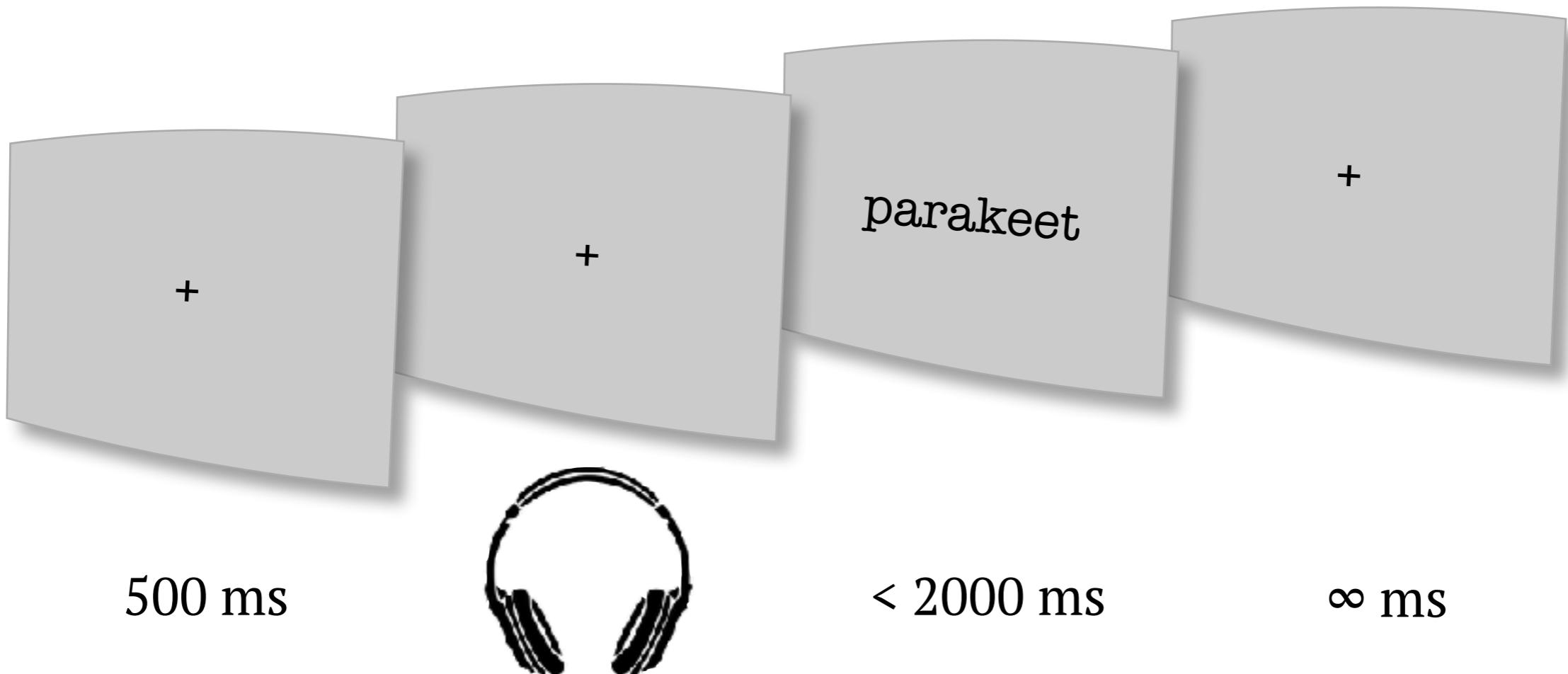
# Design & Materials



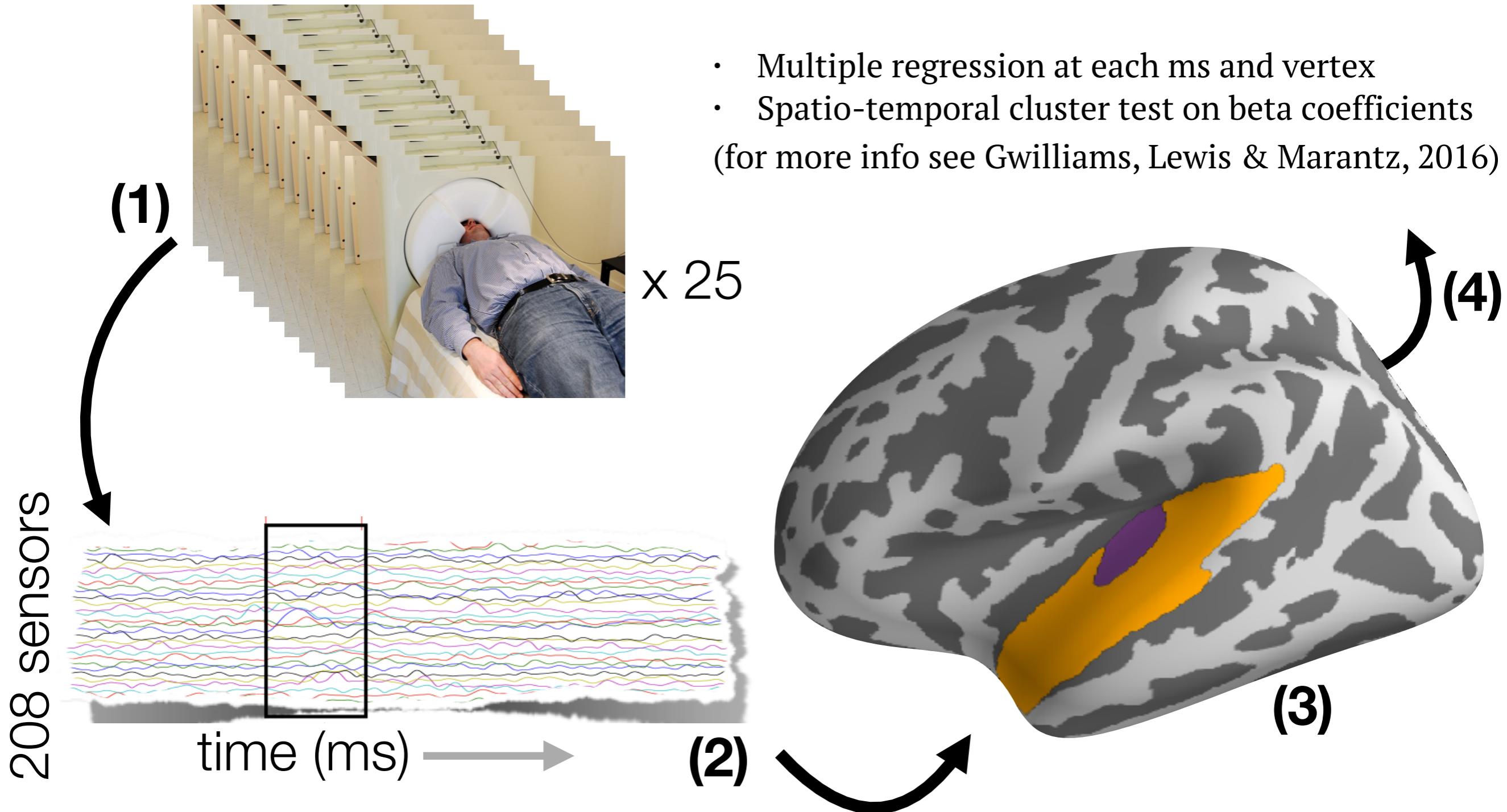
- Point of Disambiguation (POD) ranged 3-8 phonemes / 150-750 ms
- VOT (31 pairs) {p-b, t-d, k-g} and POA (22 pairs) {t-k, p-t}

# Design & Materials

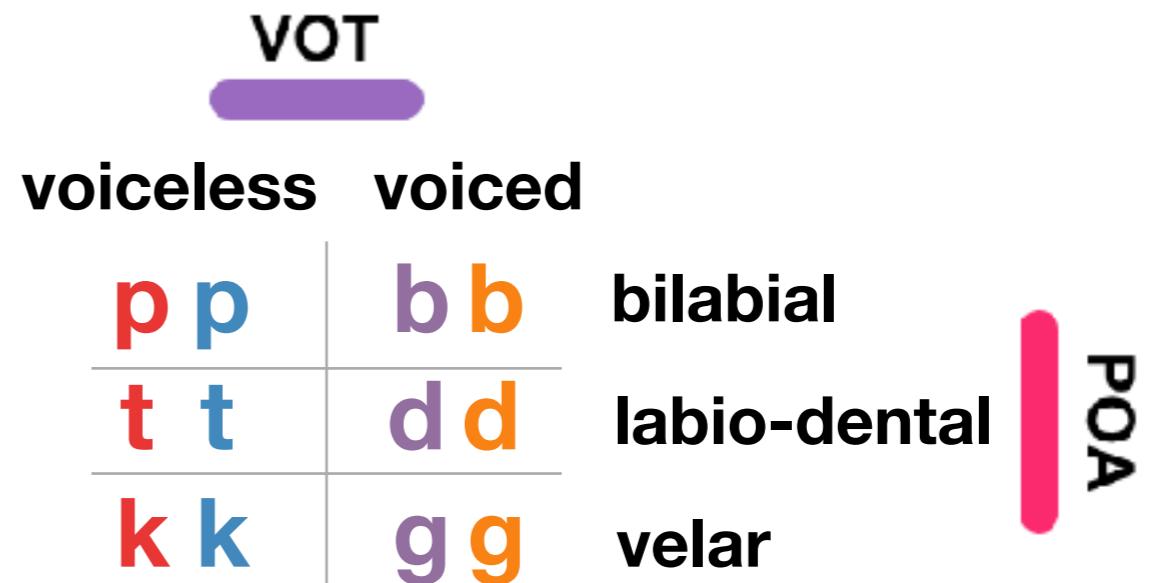
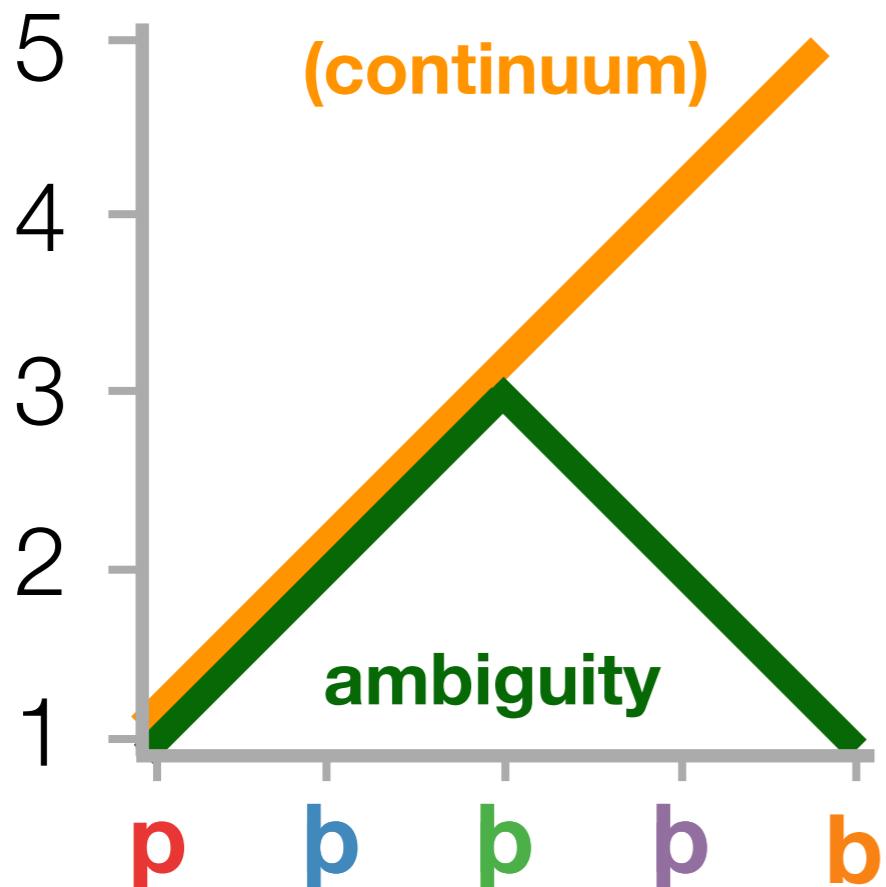
---



# Procedure & Analysis



# Four Experimental Variables



# Today's Questions

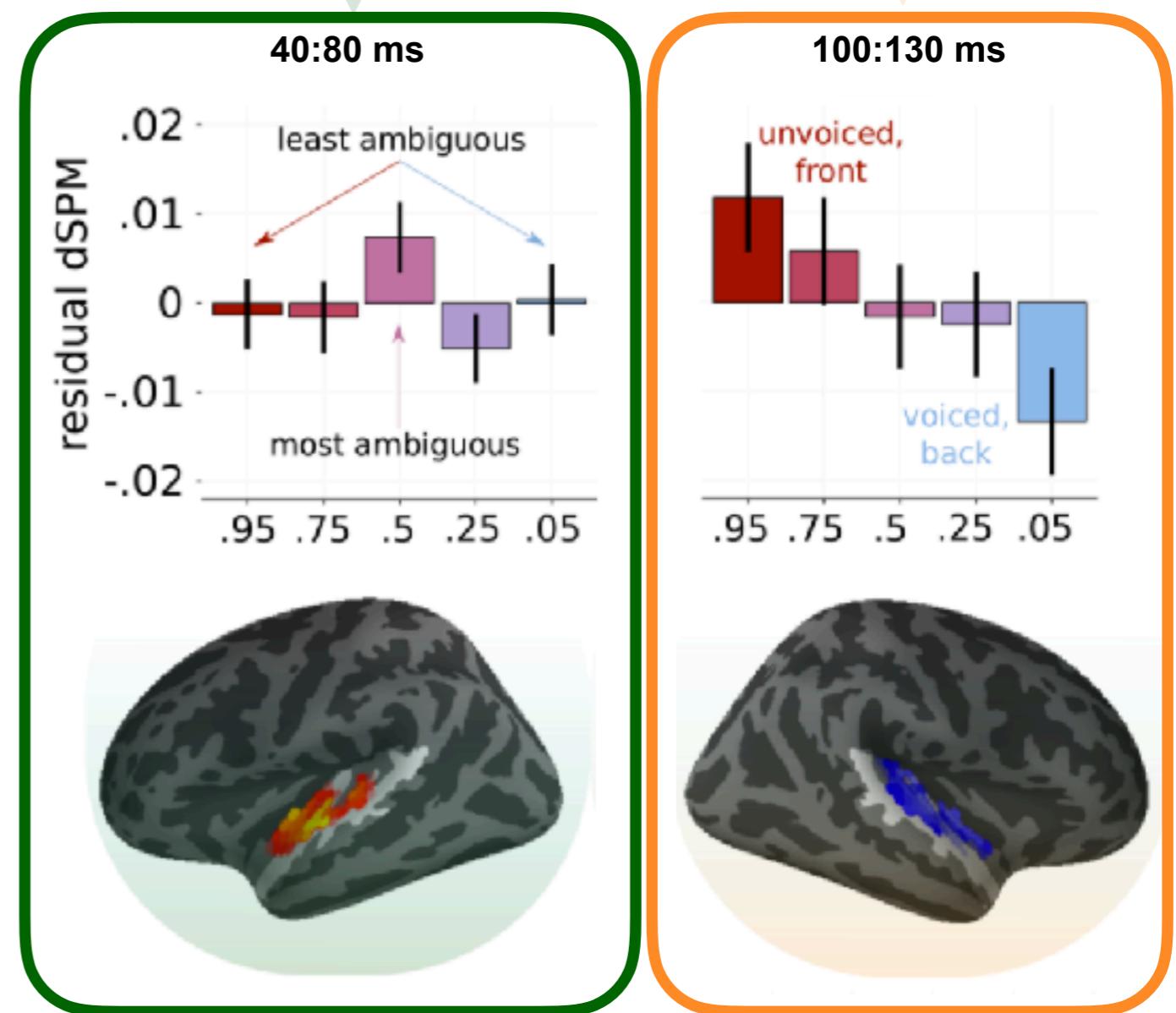
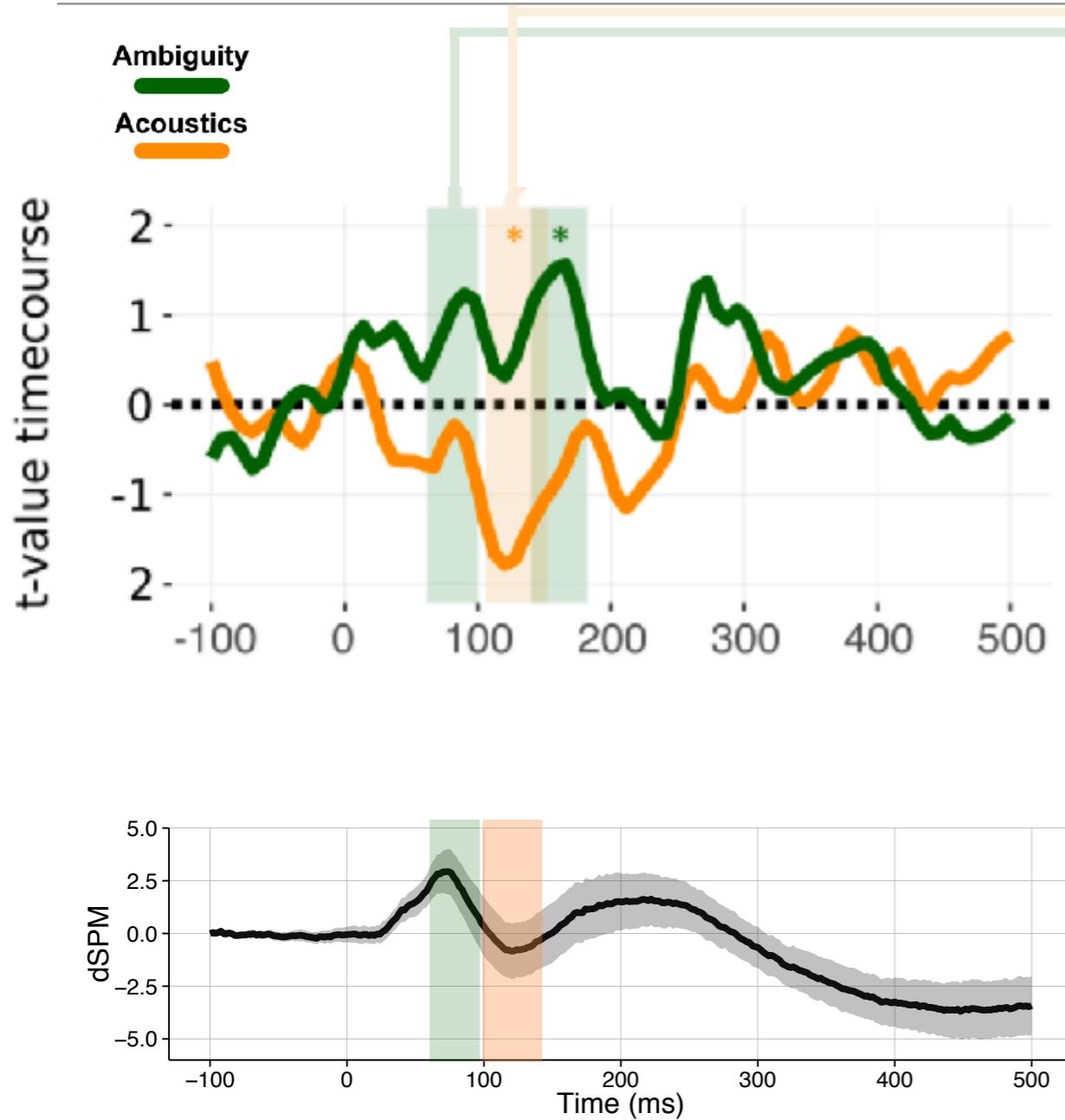
---

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



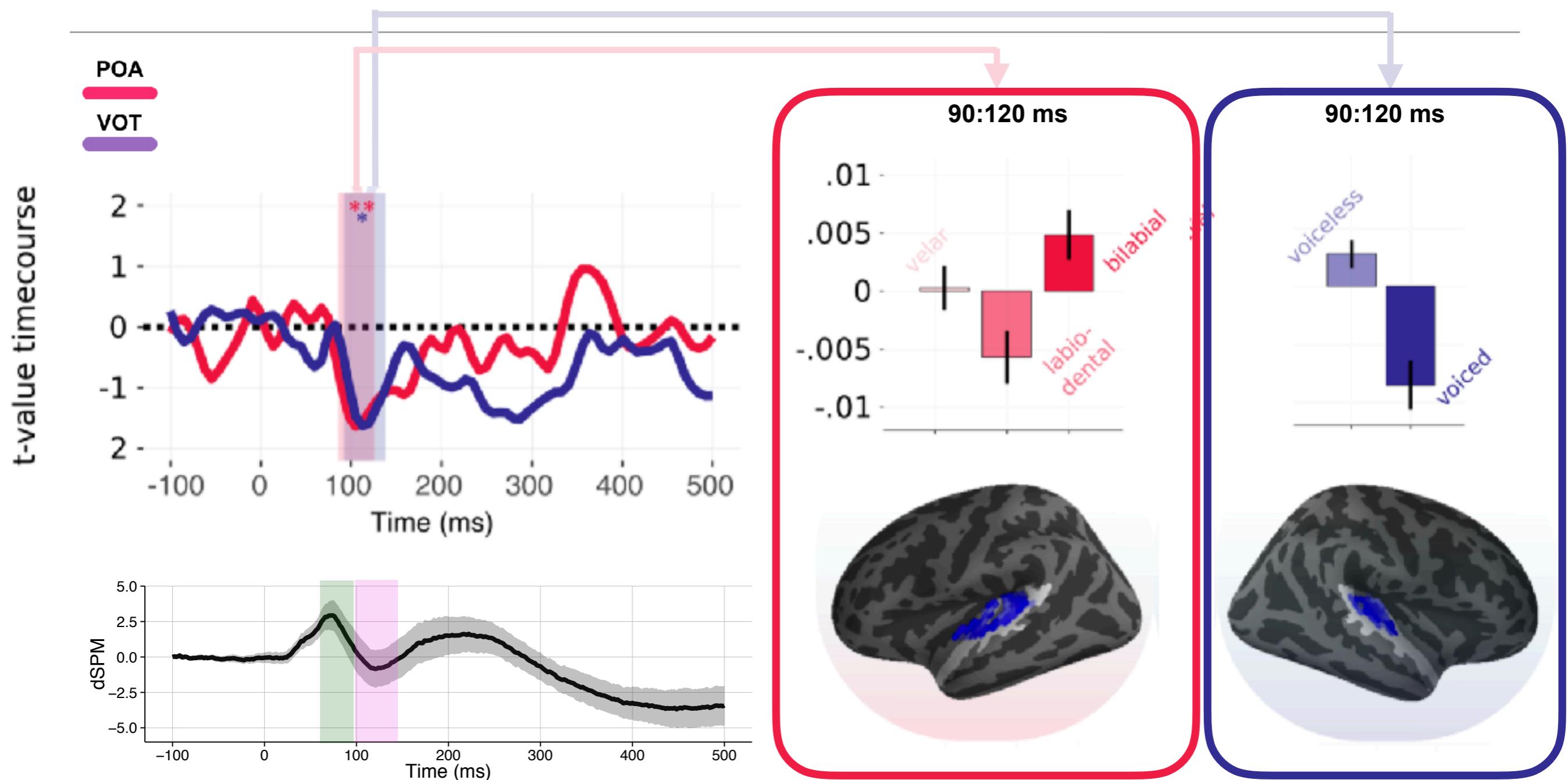
p p b b b

# Subphonetics at Onset

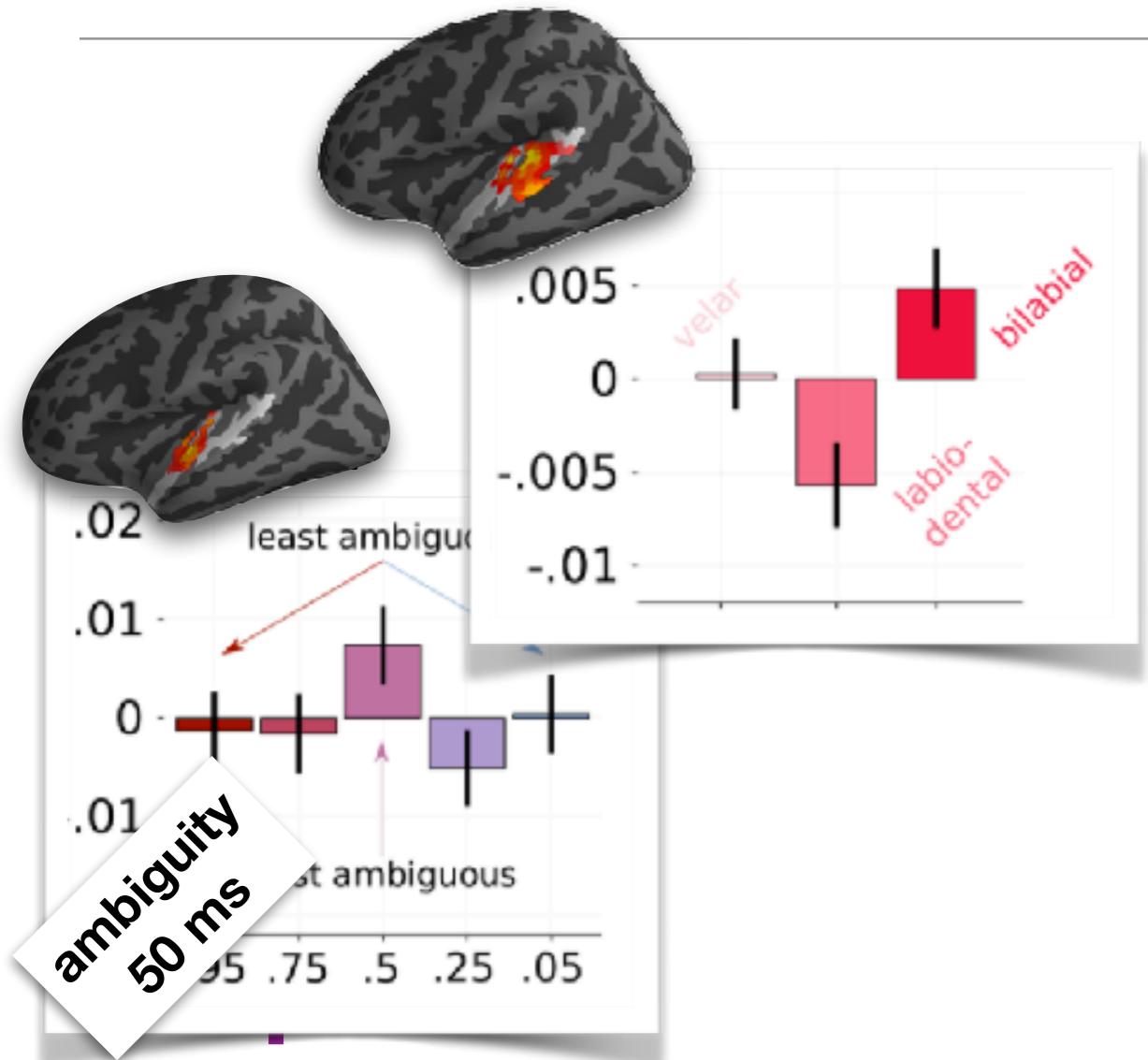


p b b b b

# Phonetic Features at Onset

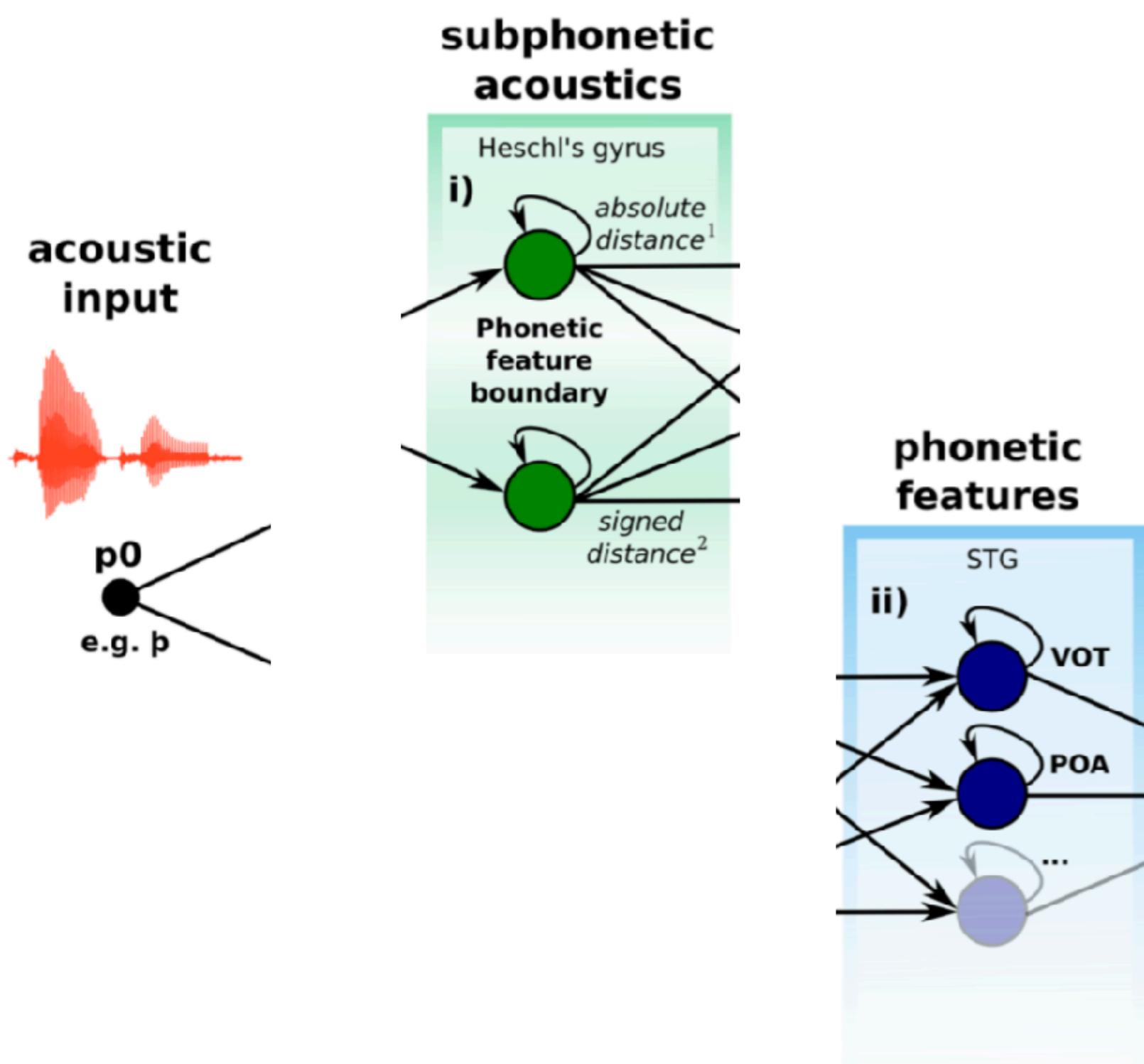


# Interim Conclusion

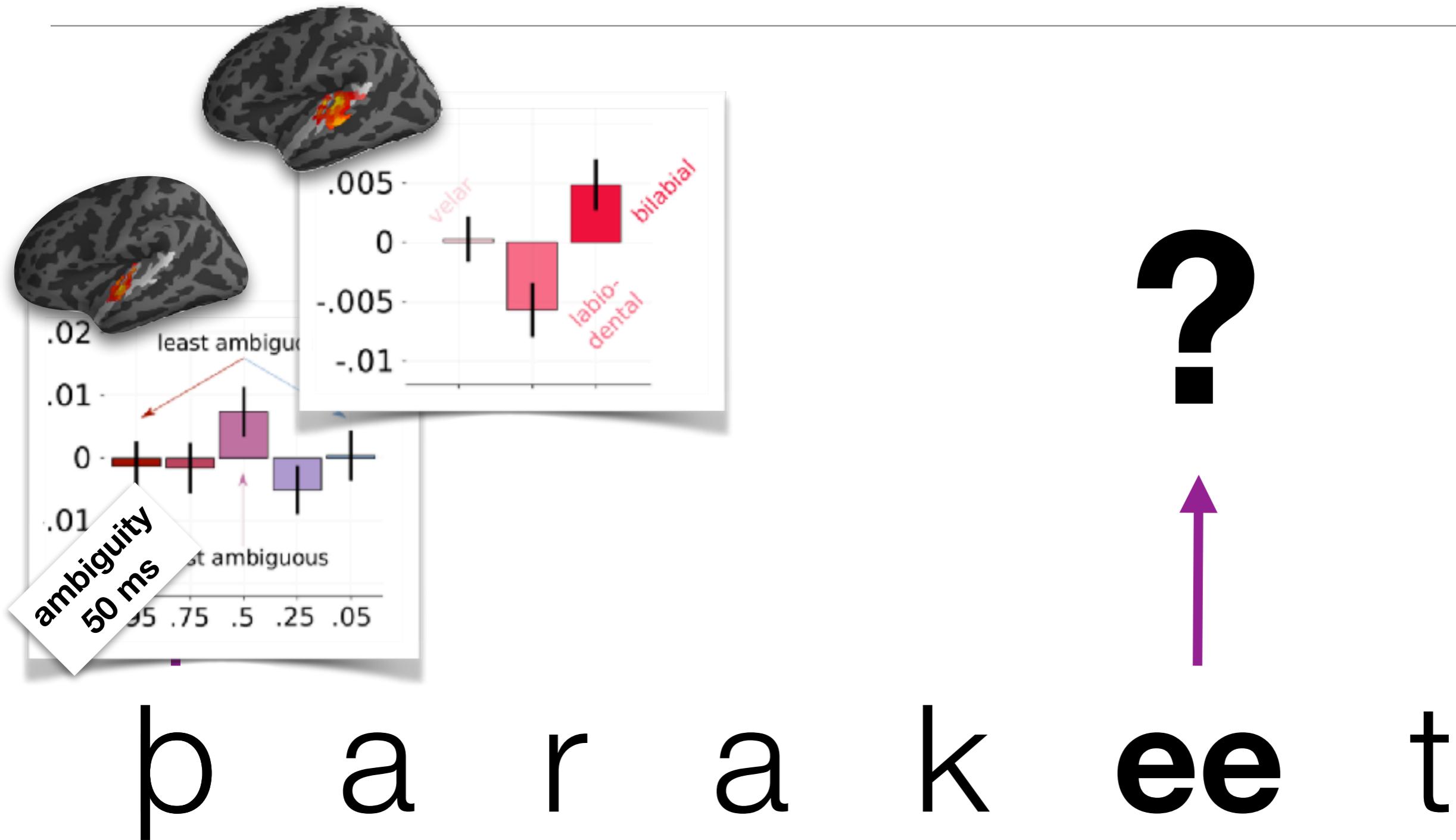


b a r a k ee t

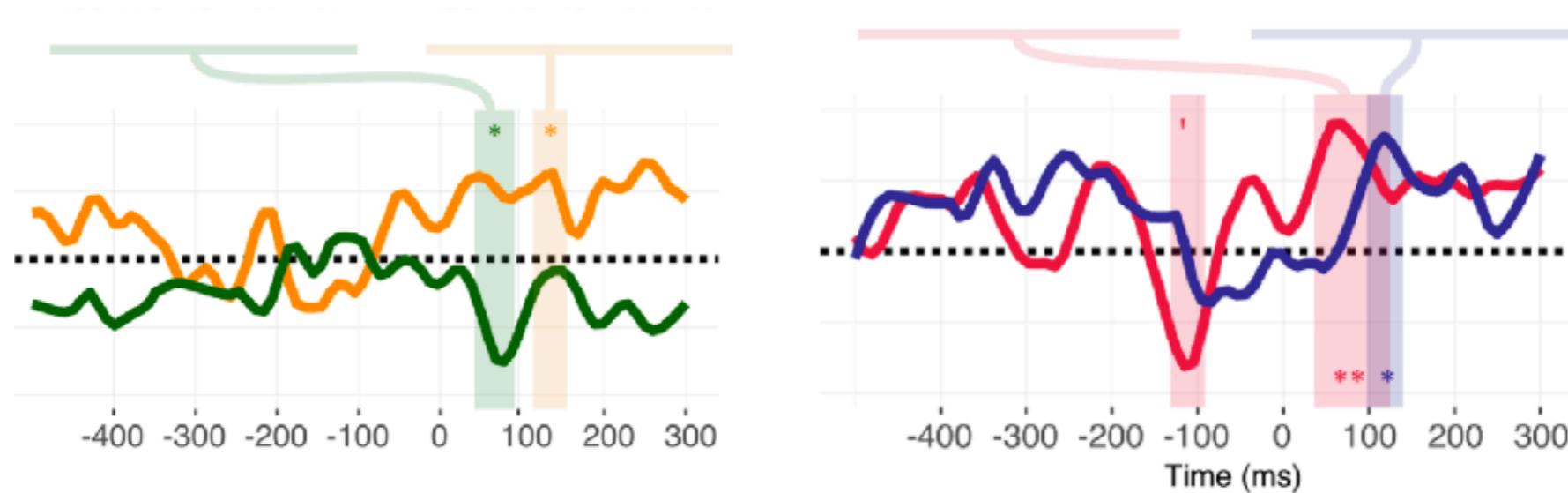
# Putting together the processing pieces



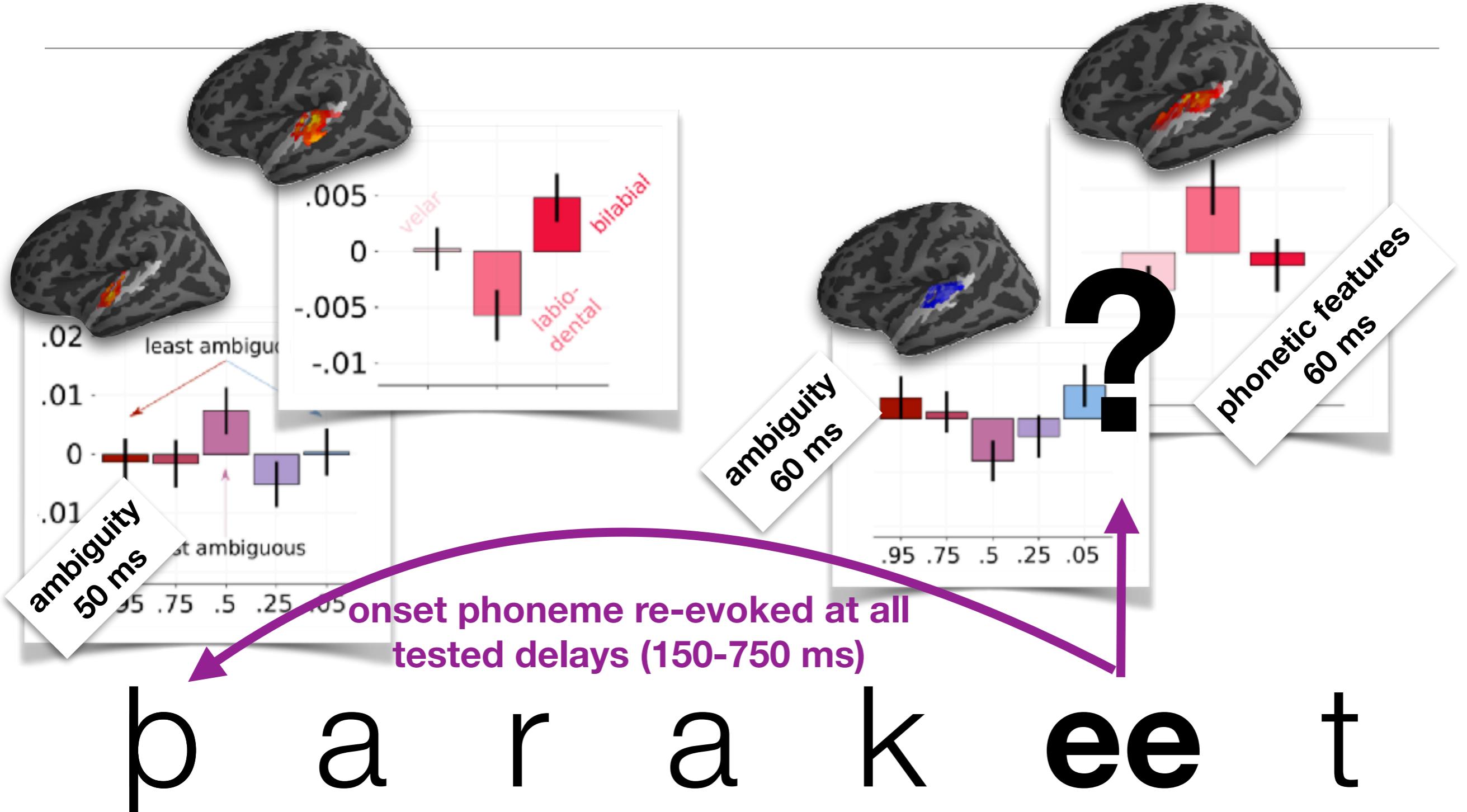
# Interim Conclusion



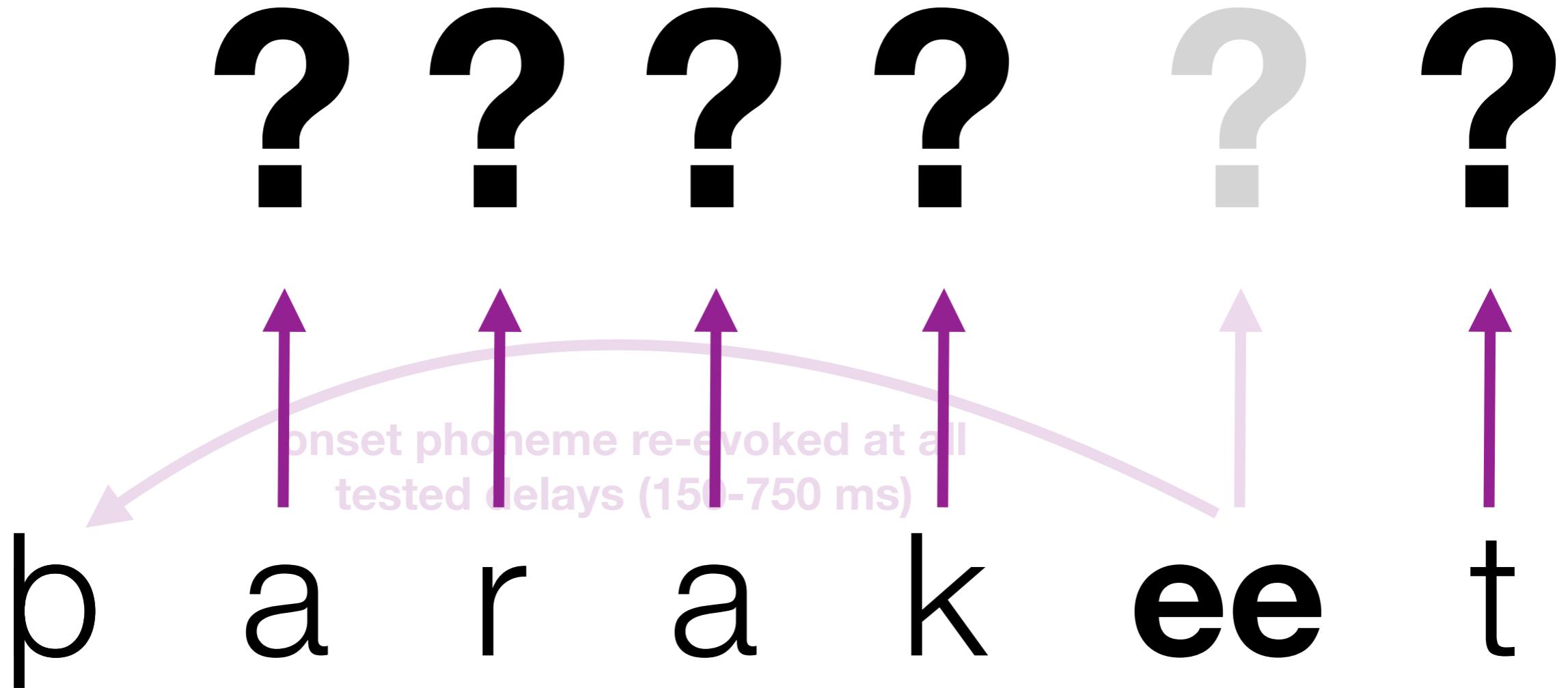
# Ambiguity at POD



# Interim Conclusion

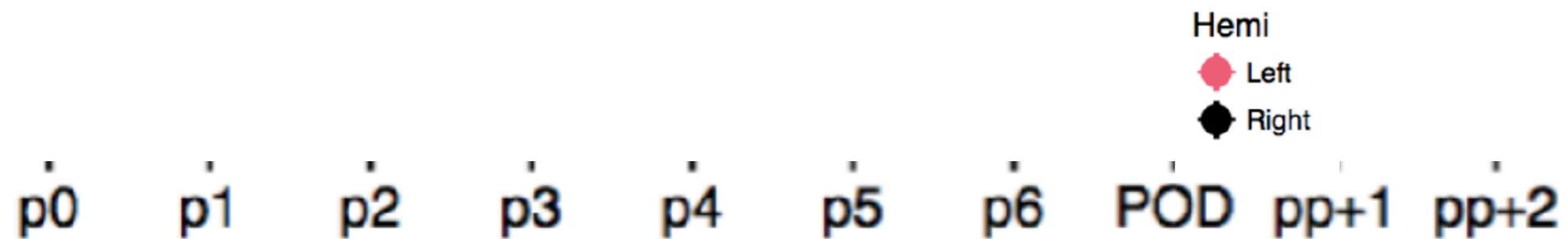


# Interim Conclusion

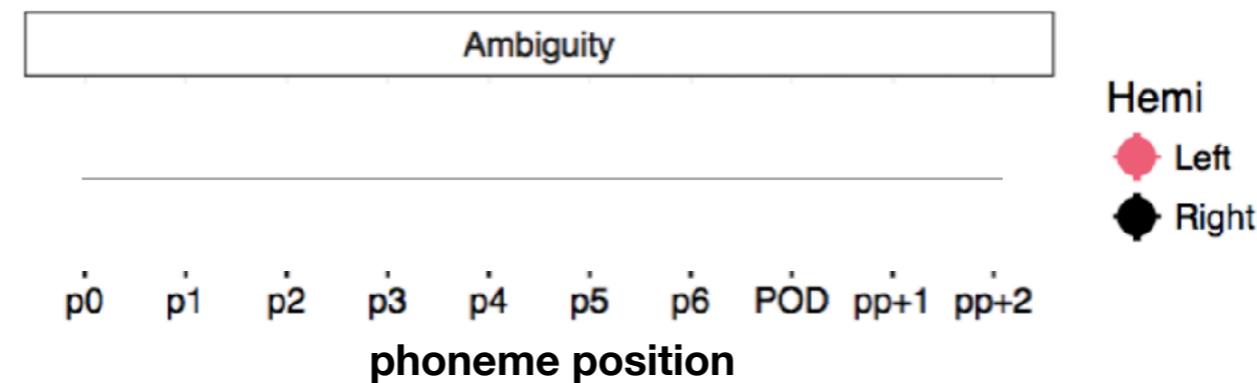


# Reactivation in Intermediate Positions

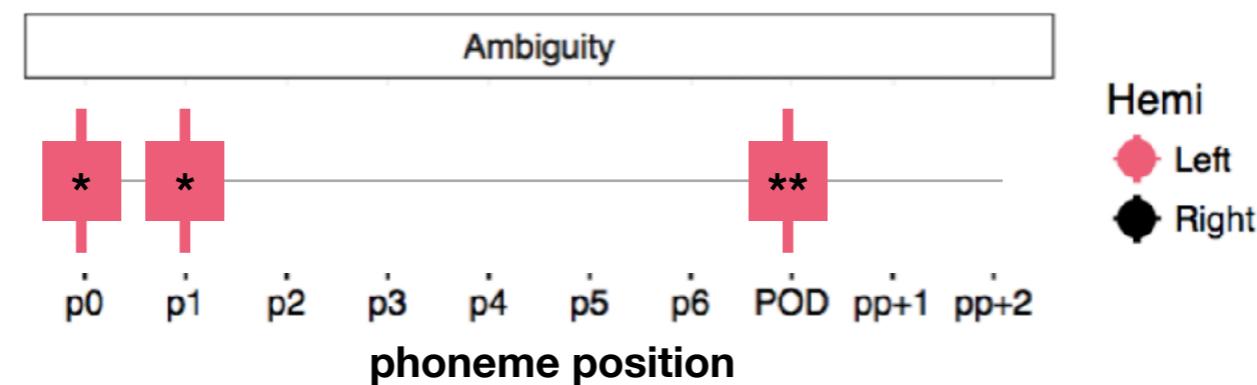
---



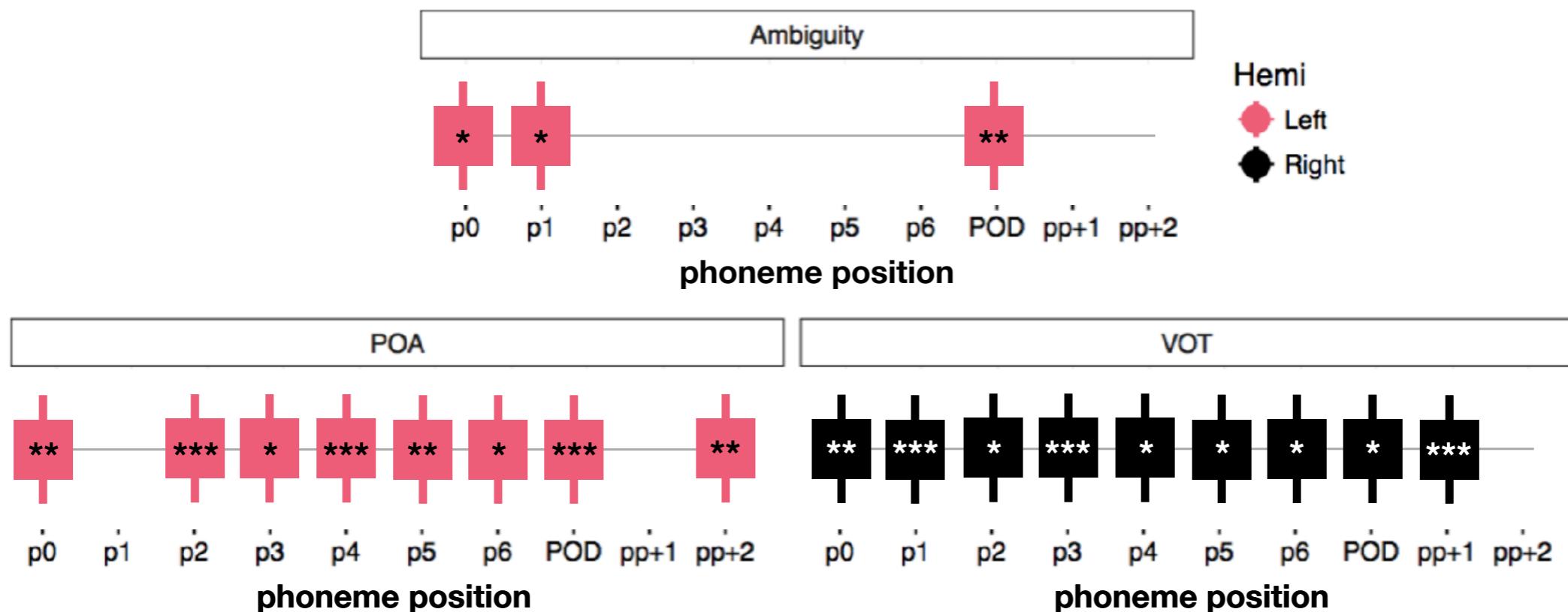
# Reactivation in Intermediate Positions



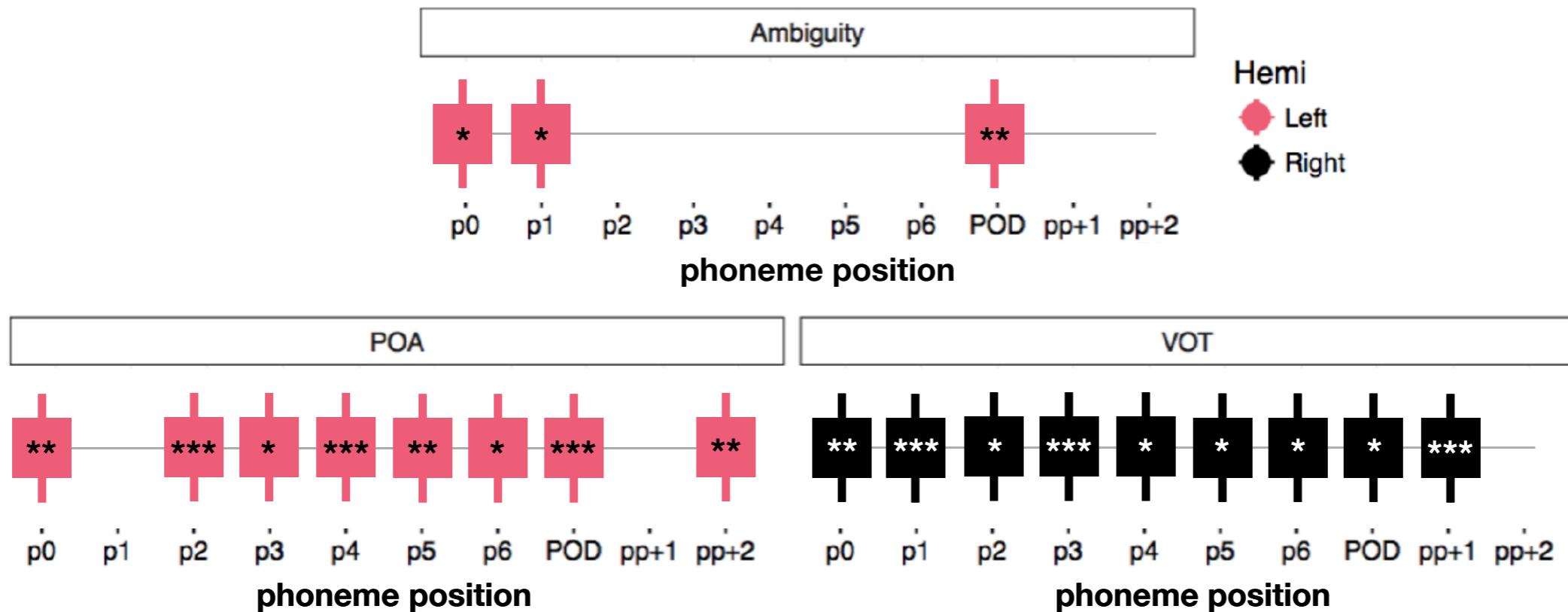
# Reactivation in Intermediate Positions



# Reactivation in Intermediate Positions

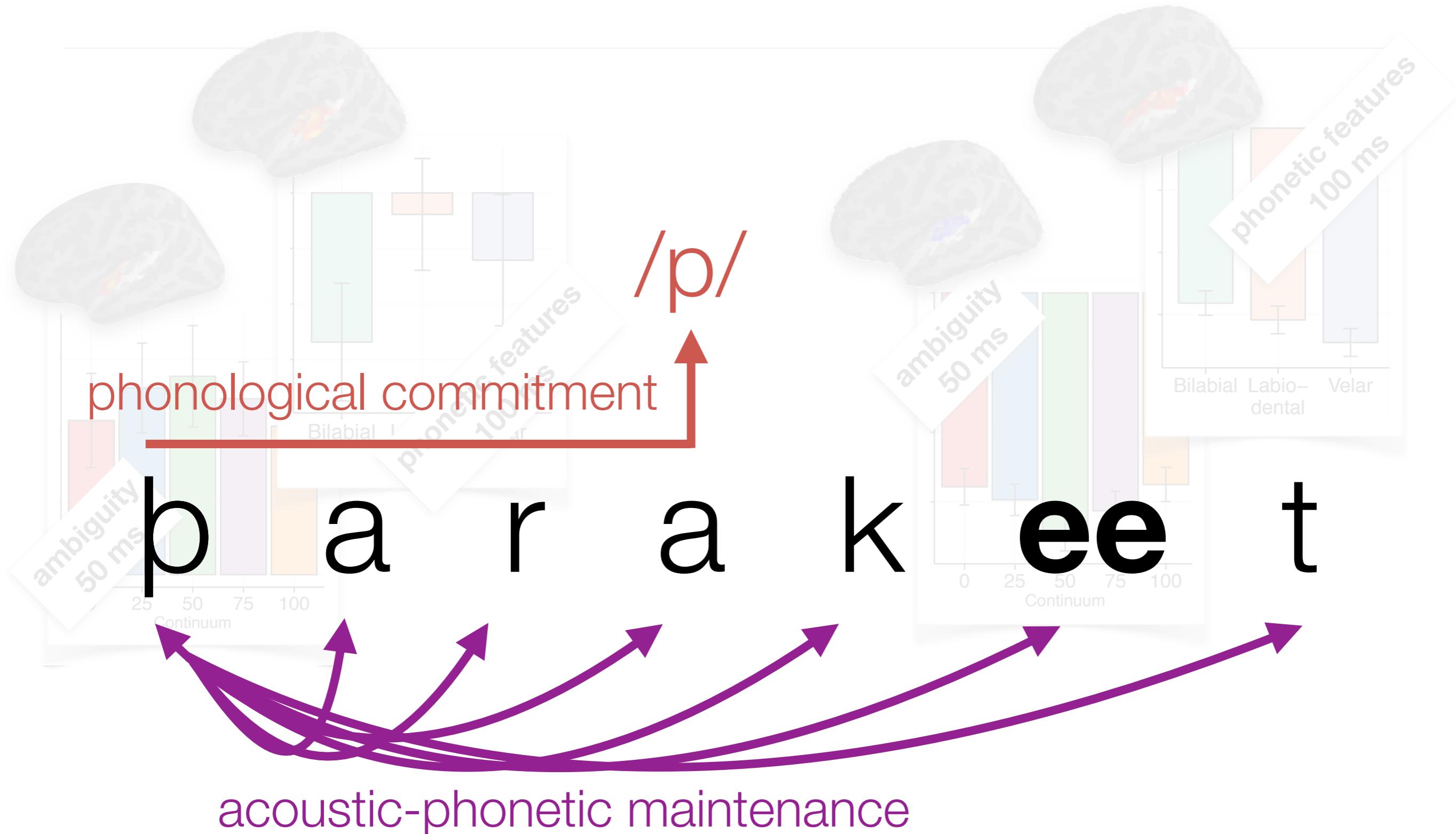


# Reactivation in Intermediate Positions



- Information is re-evoked in auditory cortex
- Specifically time-locked to the onset of subsequent phonemes
- Not specific to the ambiguous tokens – general to language processing

# Interim Conclusion

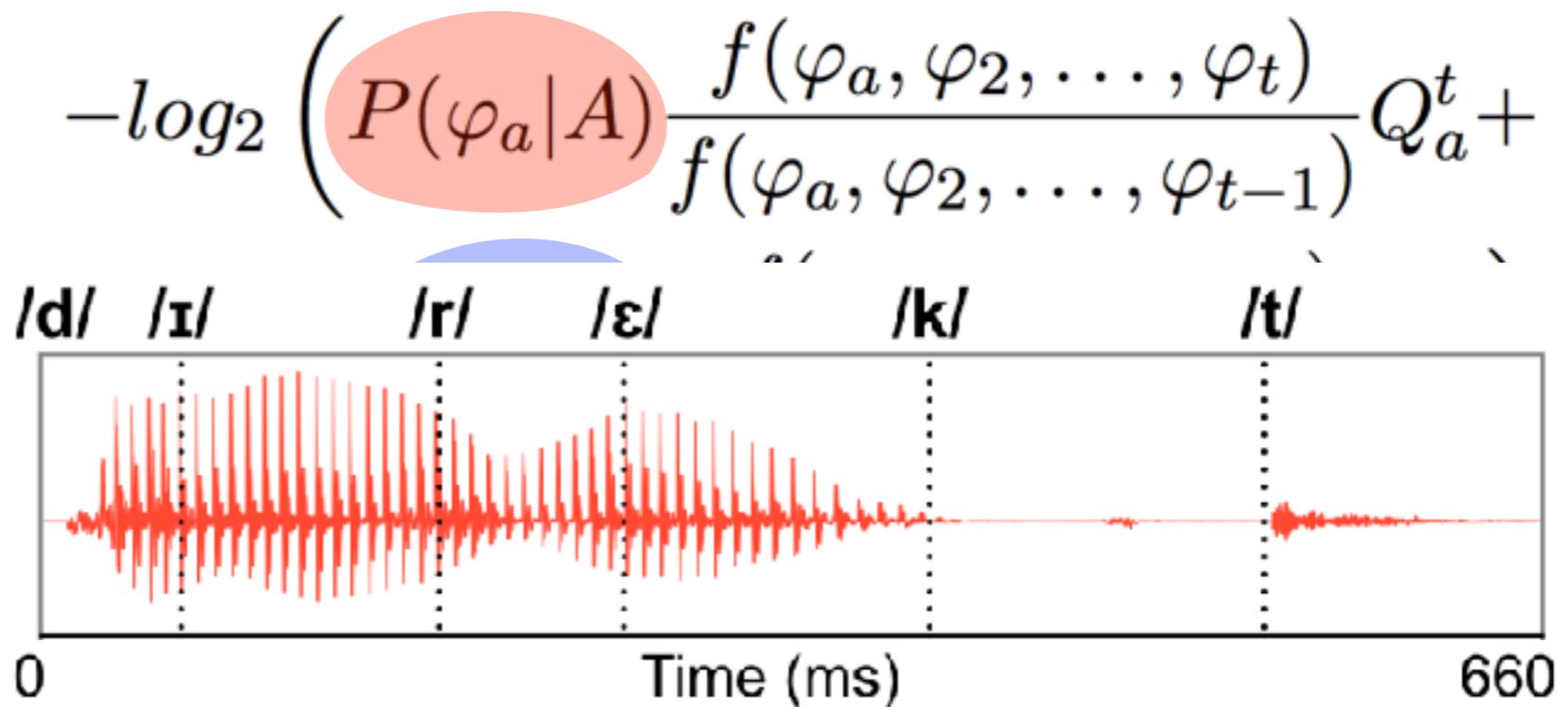


# Testing for phonological commitment

- **Surprisal:**

No commitment

Commitment



- **Entropy:**

No commitment

Commitment

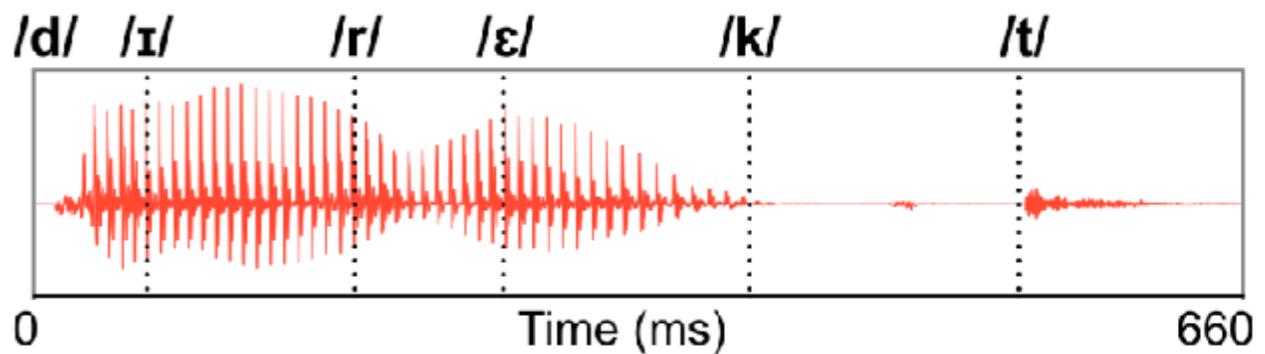
$$P(w|C, A) = P(w|C_a) P(\varphi_a | A) + P(w|C_b) P(\varphi_b | A)$$

# Model Setup

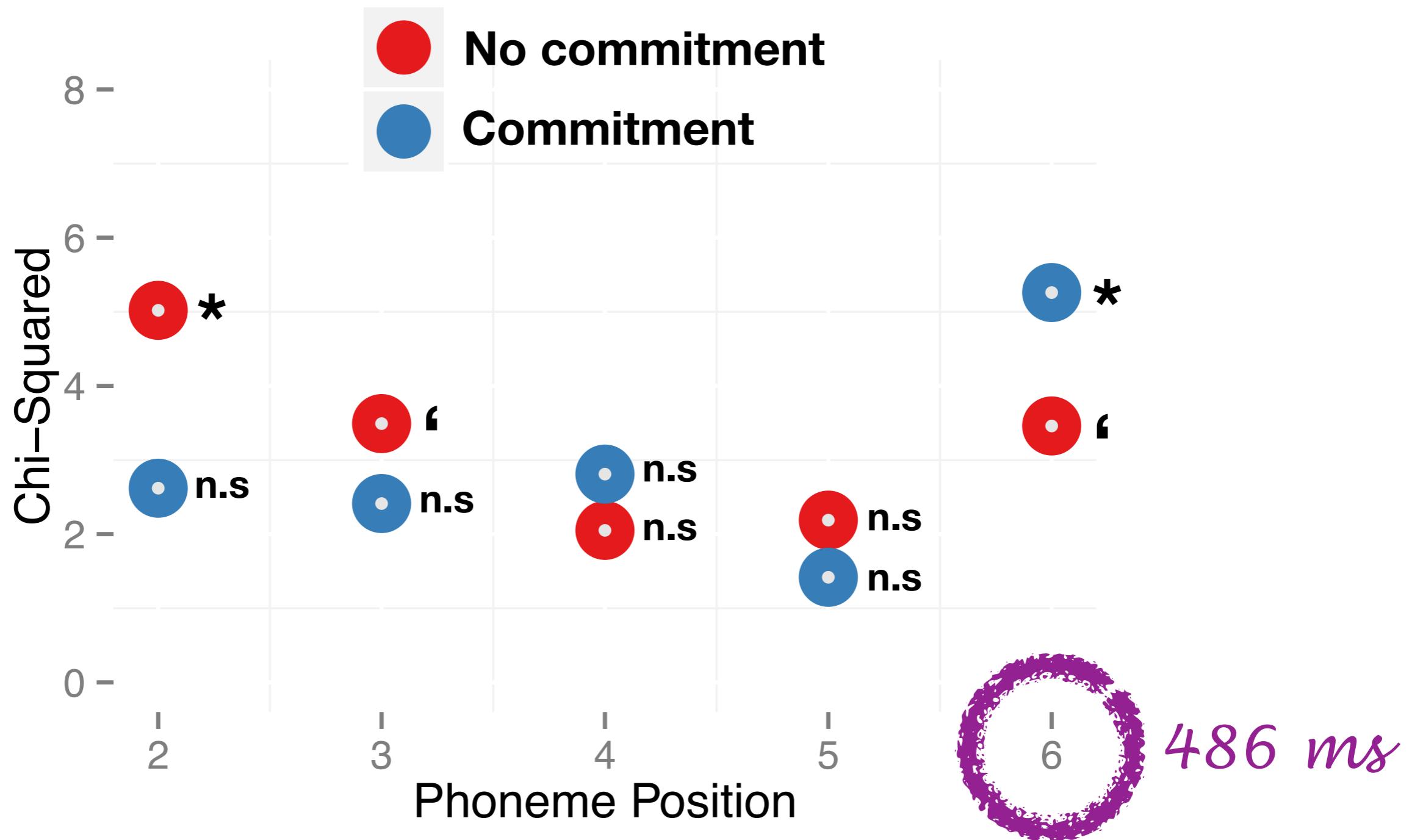
---

- **Critical variables:**  
no commitment entropy  
no commitment surprisal  
commitment entropy  
commitment surprisal

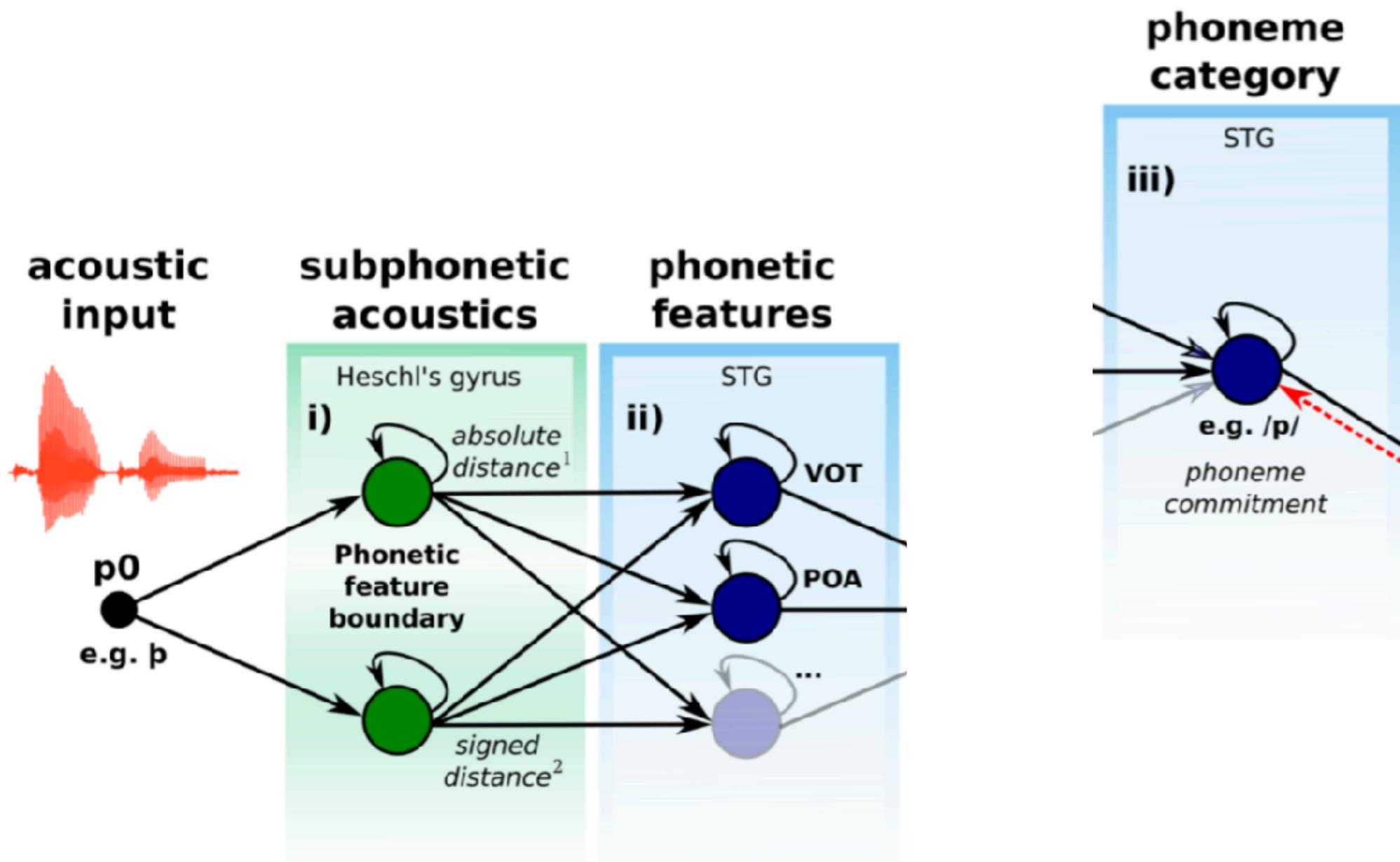
- **Control variables:**  
phoneme latency (ms)  
phoneme latency (number of phonemes)  
trial number  
block number  
stimulus amplitude  
phoneme pair  
ambiguity



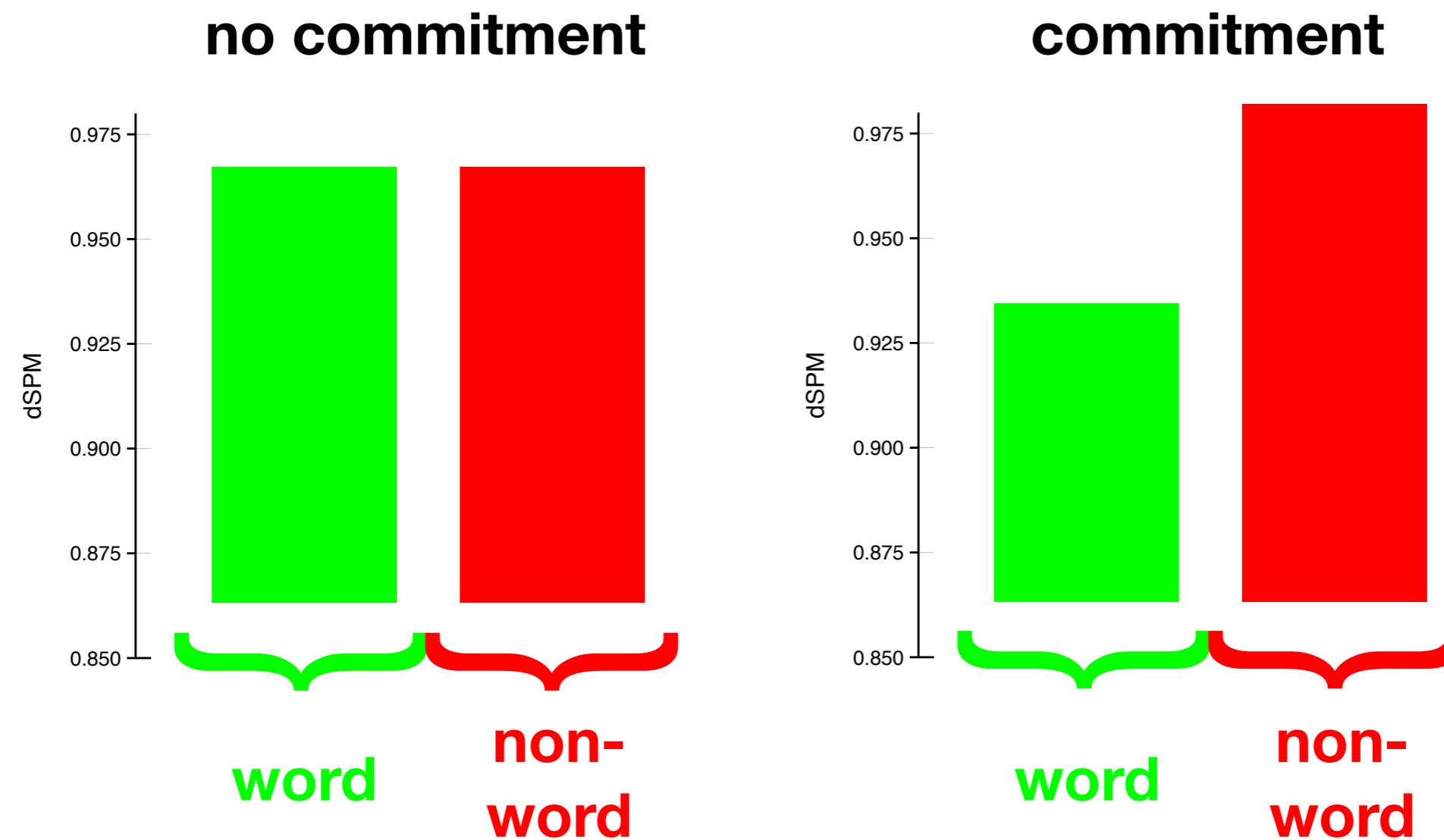
# Results



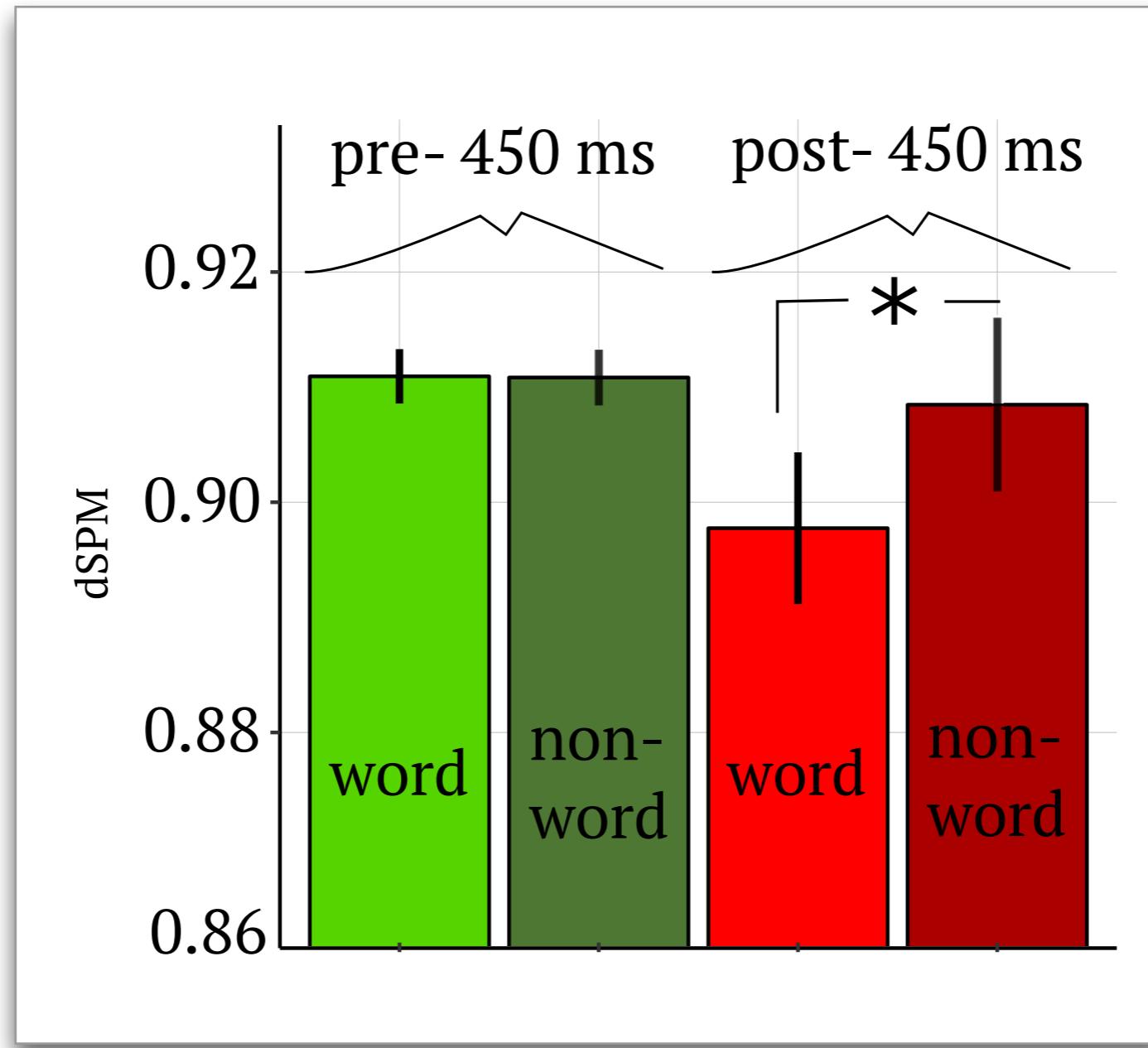
# Putting together the processing pieces



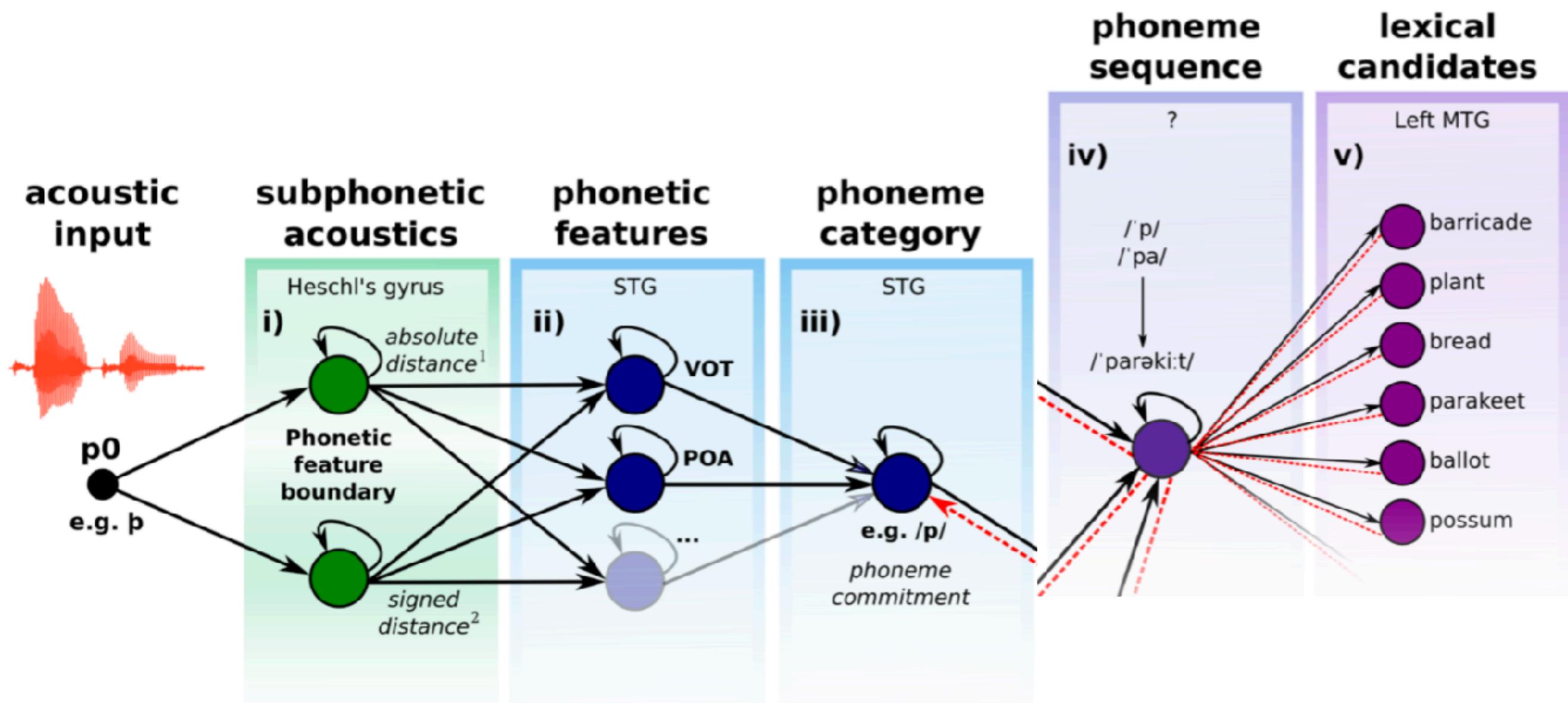
# Further test of commitment



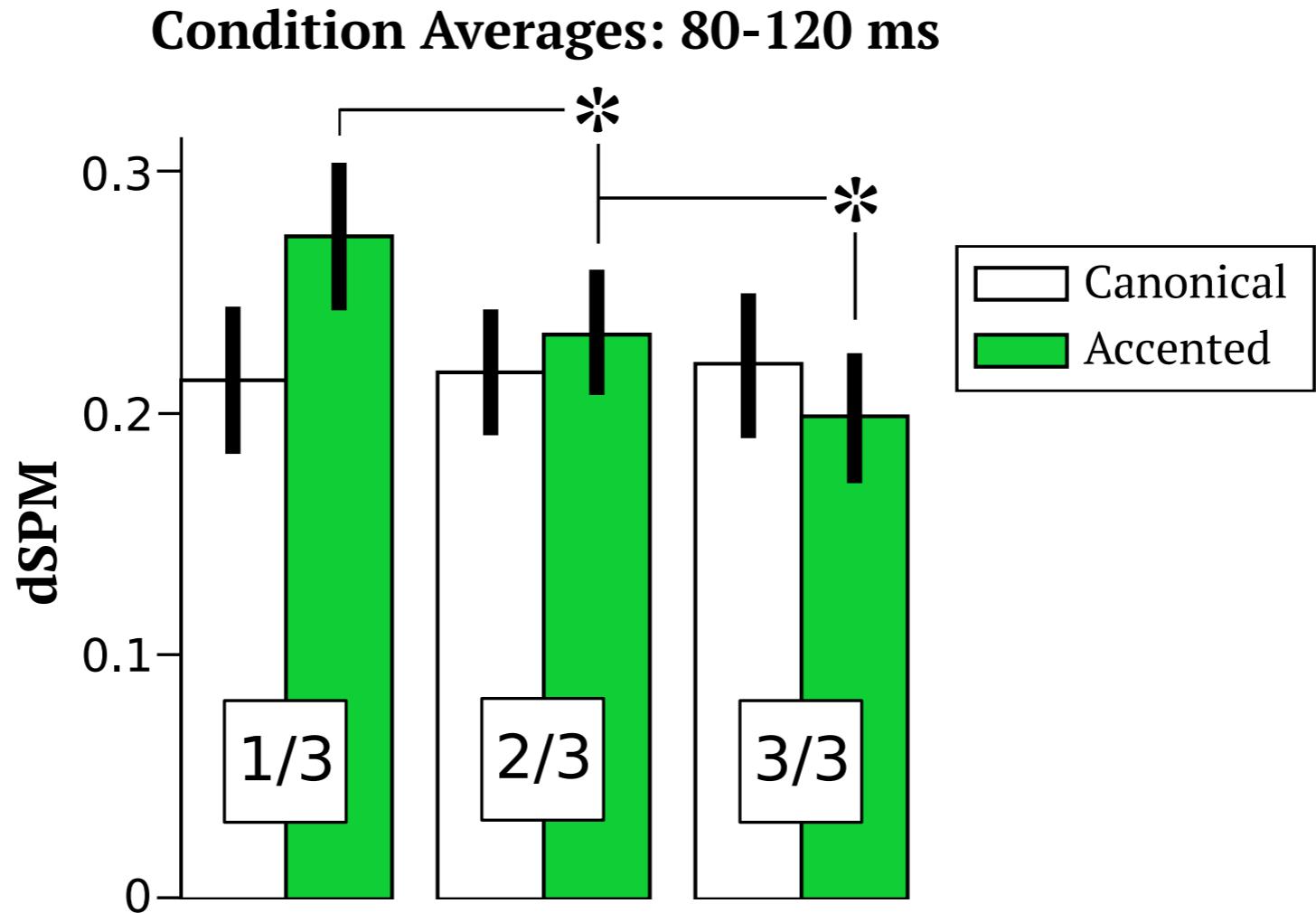
# Further test of commitment



# Putting together the processing pieces



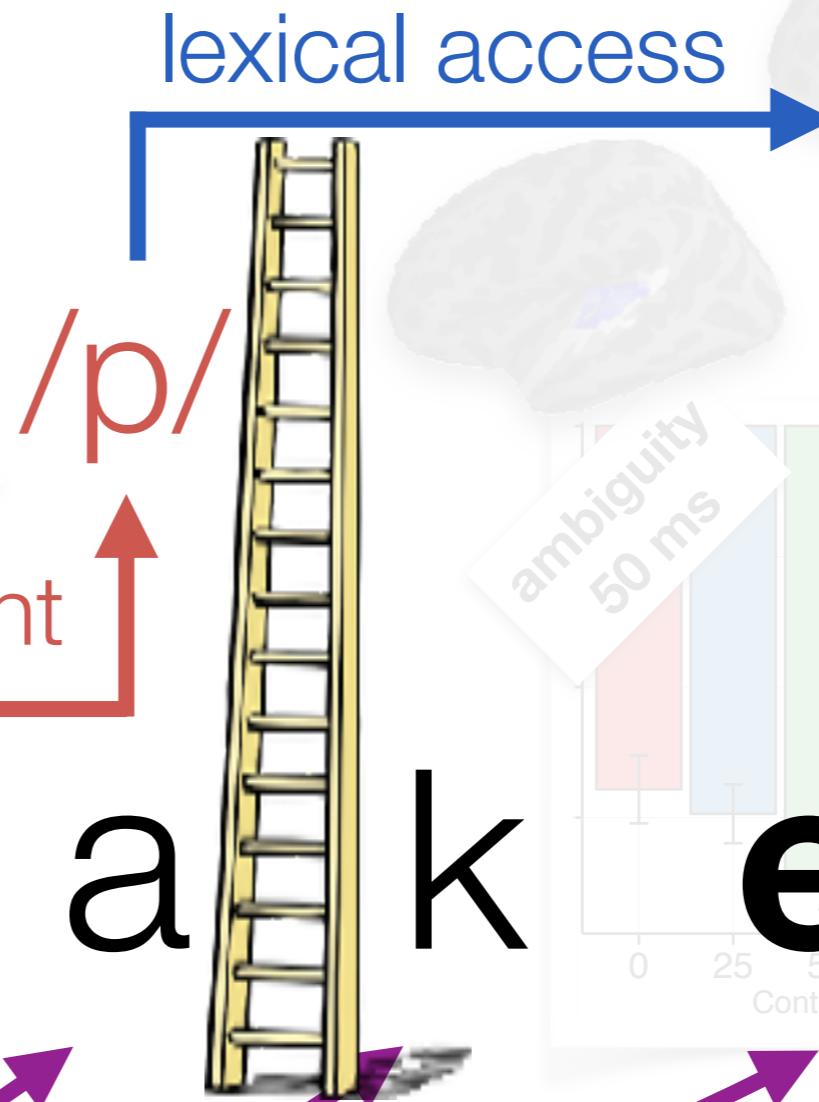
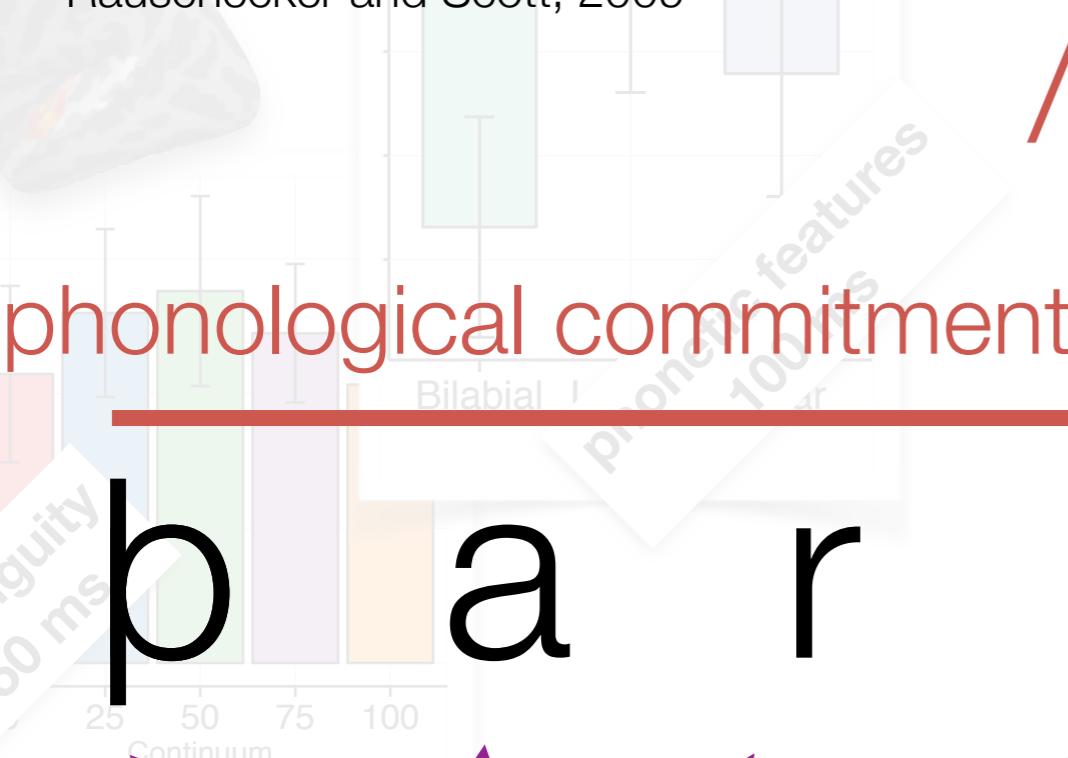
# Lexical influence



- Attunement is proposed to involve **re-tuning perceptual boundaries** between phonological categories (Norris et al., 2003; Kraljic and Samuel, 2005, 2006, 2007; Maye et al., 2008; see Samuel & Kraljic, 2009 for a review)

# Interpretation

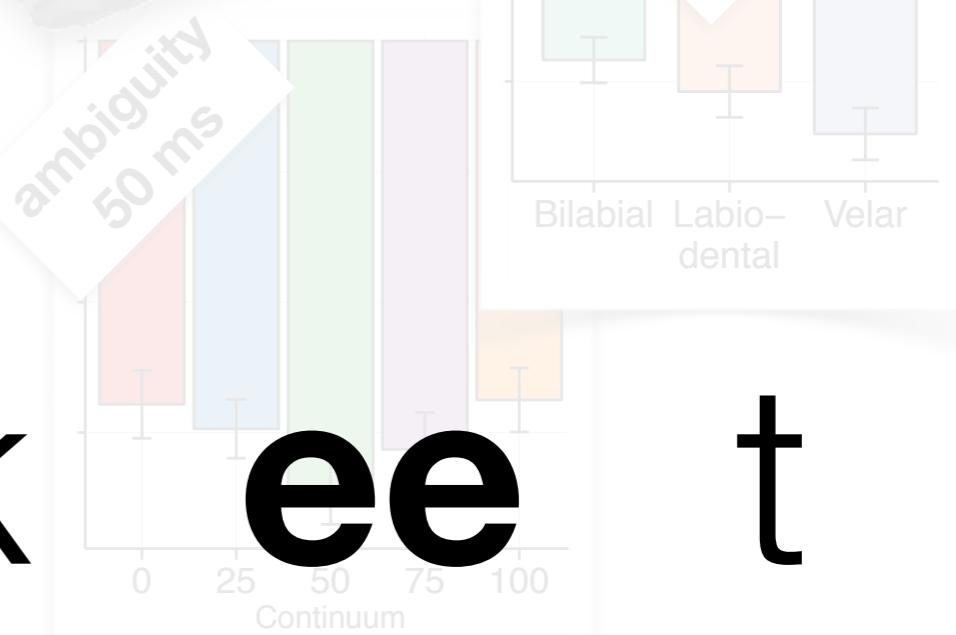
**Processing hierarchy:** Scott and Johnsrude, 2003; Hickock and Poeppel, 2004; Liebenthal et al., 2005; Rauschecker and Scott, 2009



b a r a k ee t

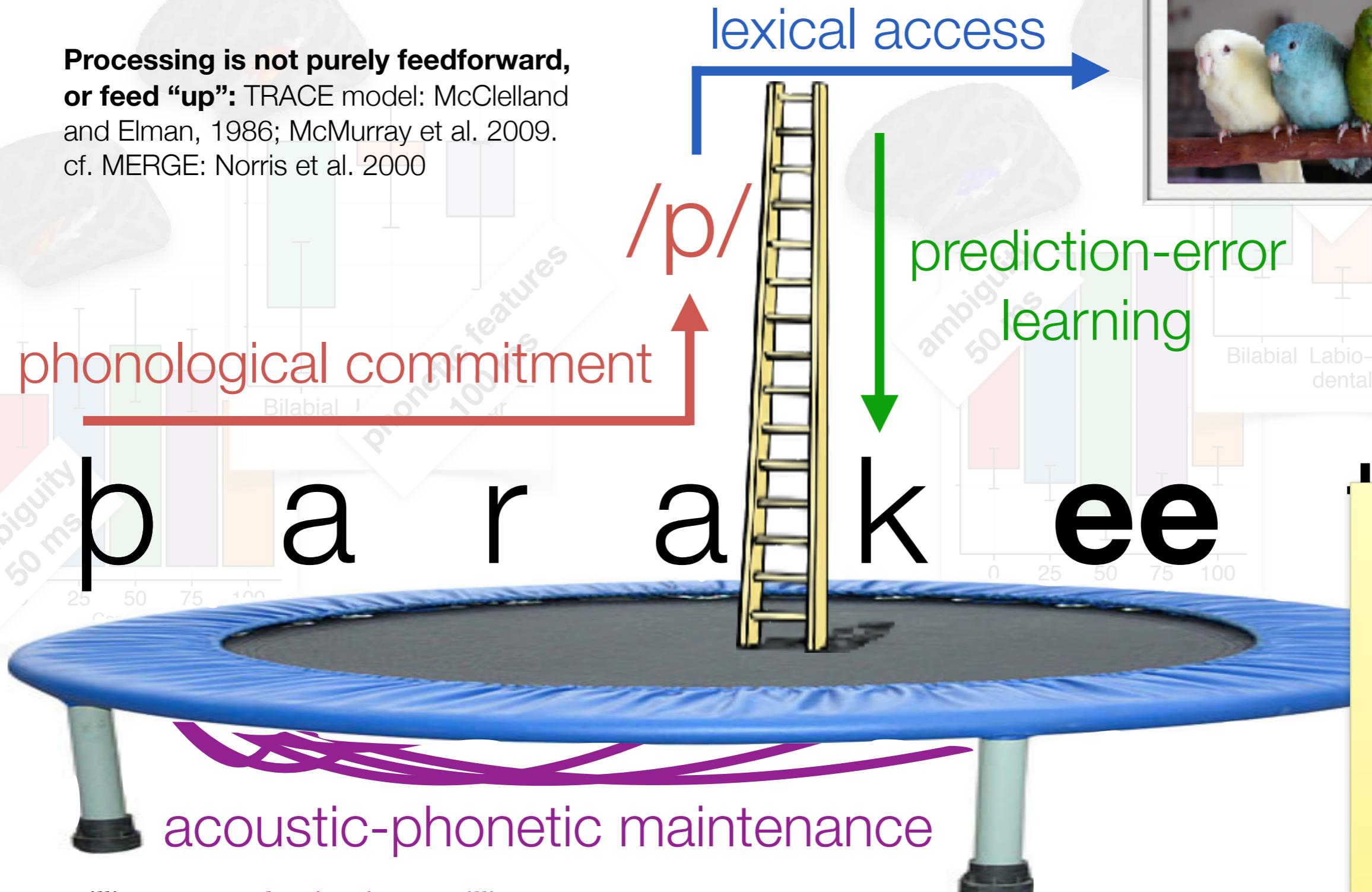


acoustic-phonetic maintenance



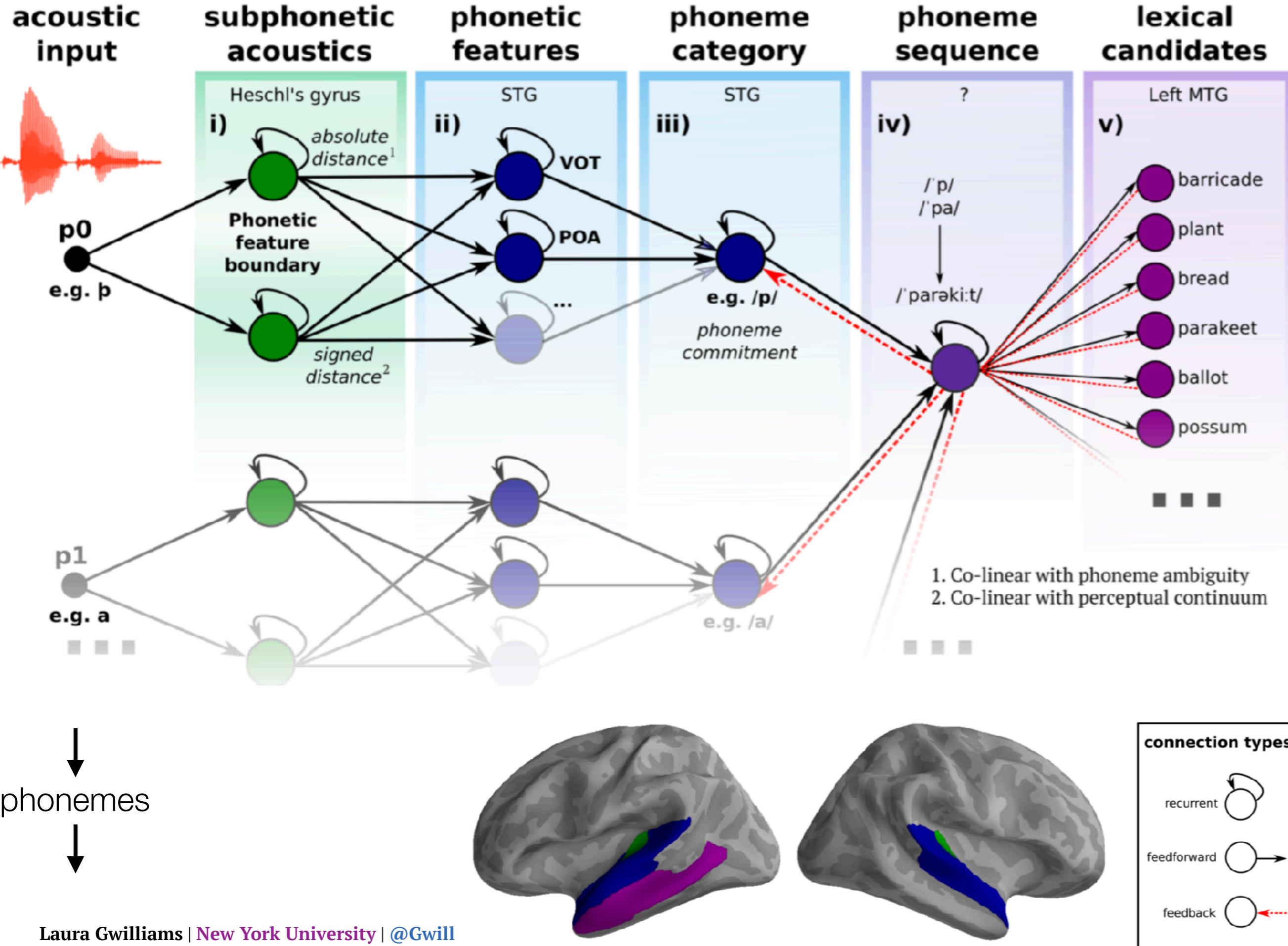
# Interpretation

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



making commitment to category does not cost — the system can flexibly avoid committing to category; avoiding the function of exposure

but what happens if you continuously jump on a trampoline because you made a wrong choice? well, that's what happens when talking to someone with an accent.



 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 | New York University | @GwilliamsL





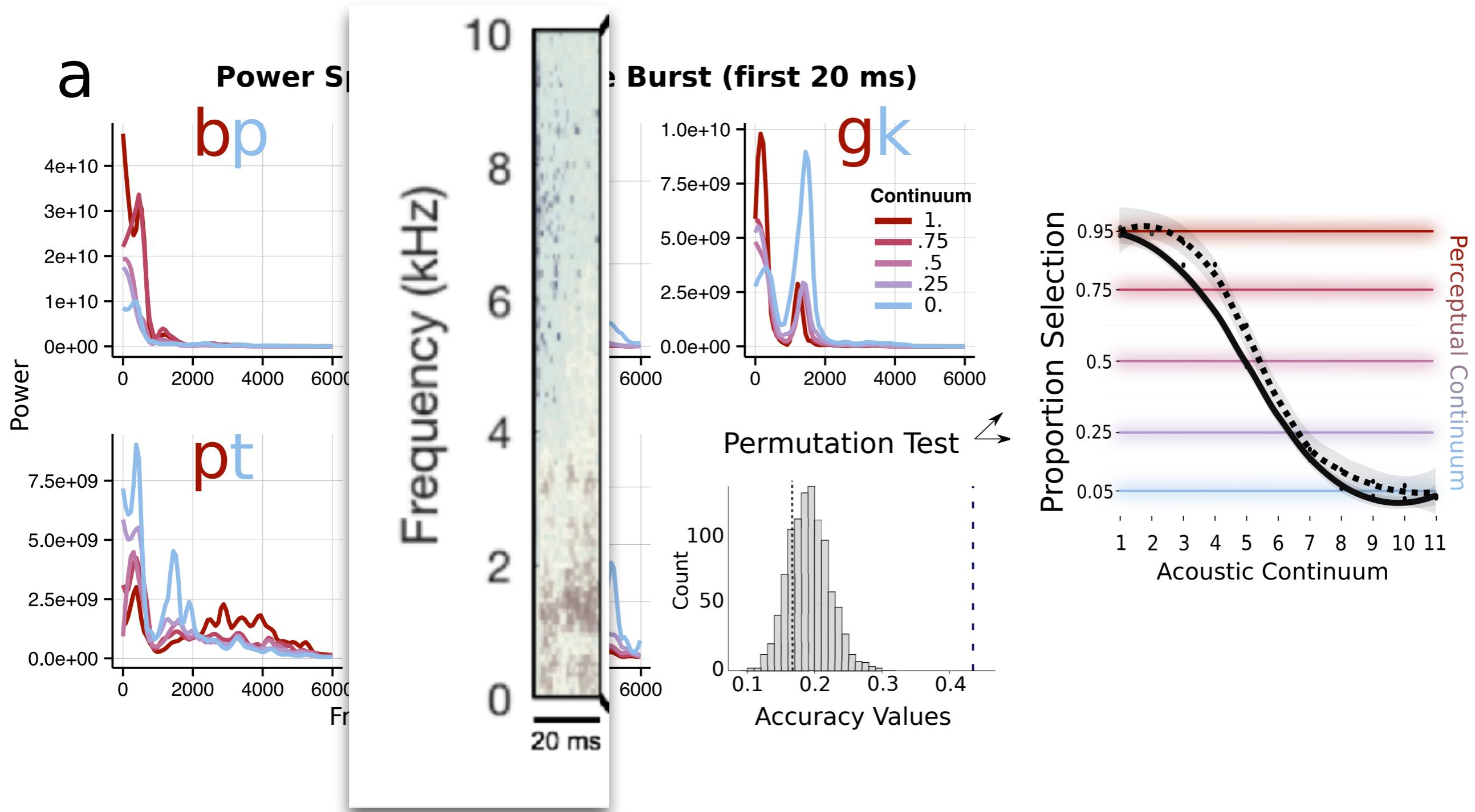
NEW YORK UNIVERSITY

✉ laura.gwilliams@nyu.edu  
🐦 @GwilliamsL

# Thank you!

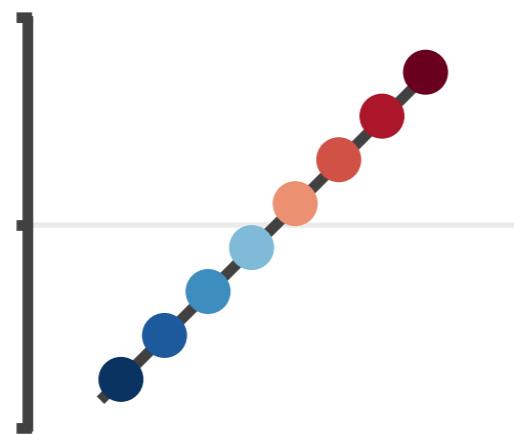


# Is ambiguity correlated with acoustic properties?

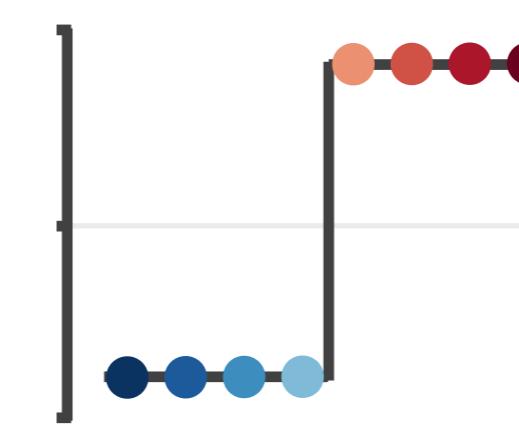


## Predictive Coding

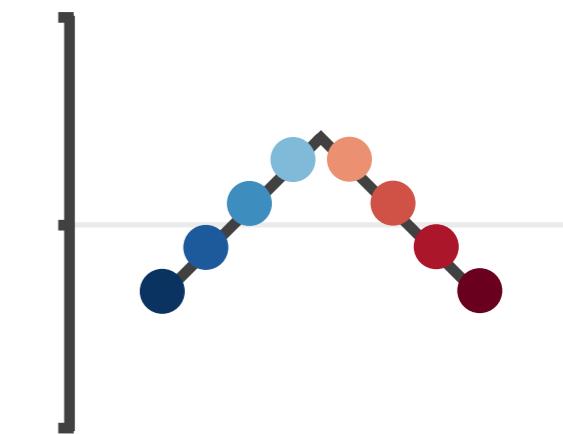
Linear Evidence



Categorical Percept

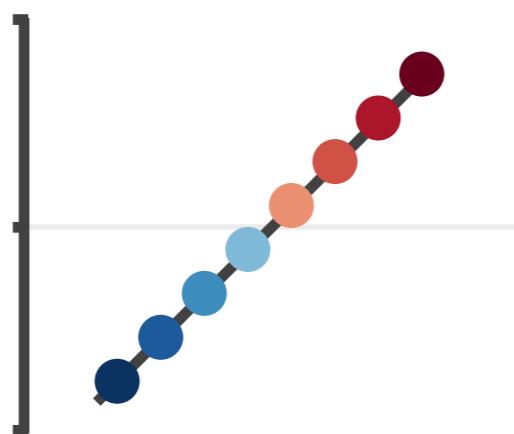


Ambiguity

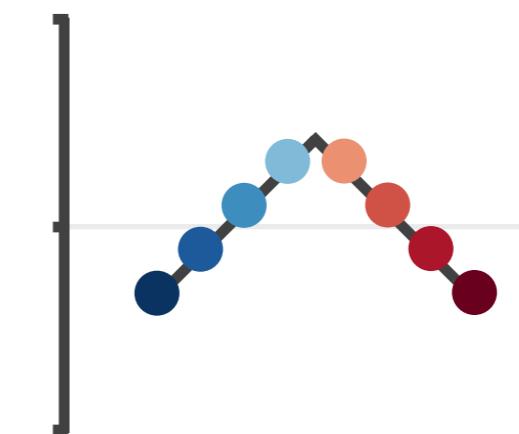


## Neutralisation

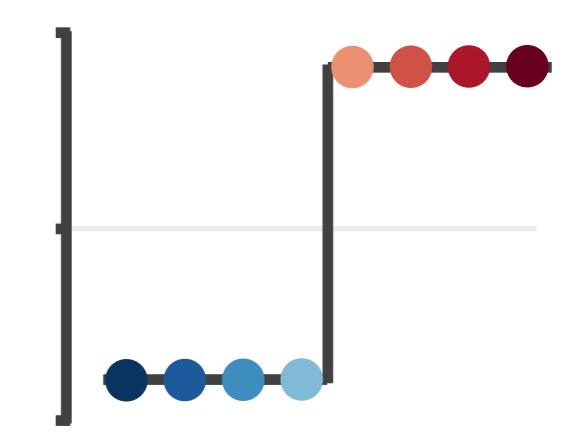
Linear Evidence



Ambiguity

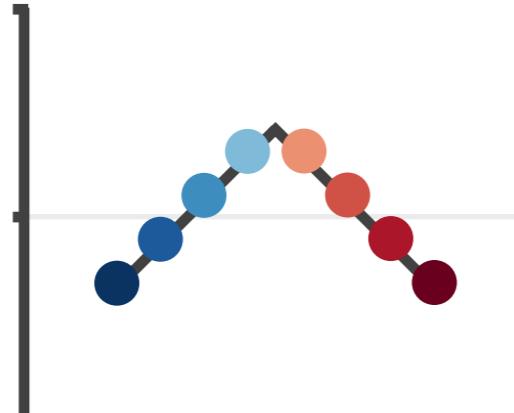


Categorical Percept

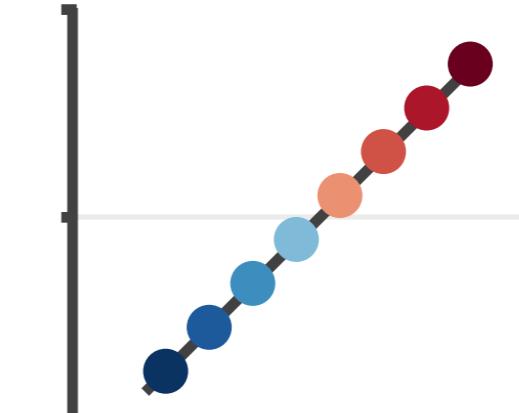


## Cut-through connection

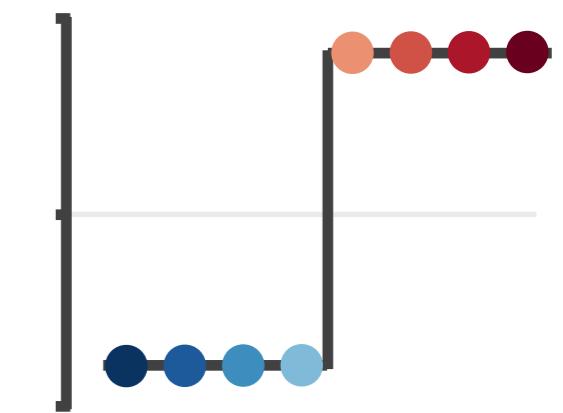
Ambiguity



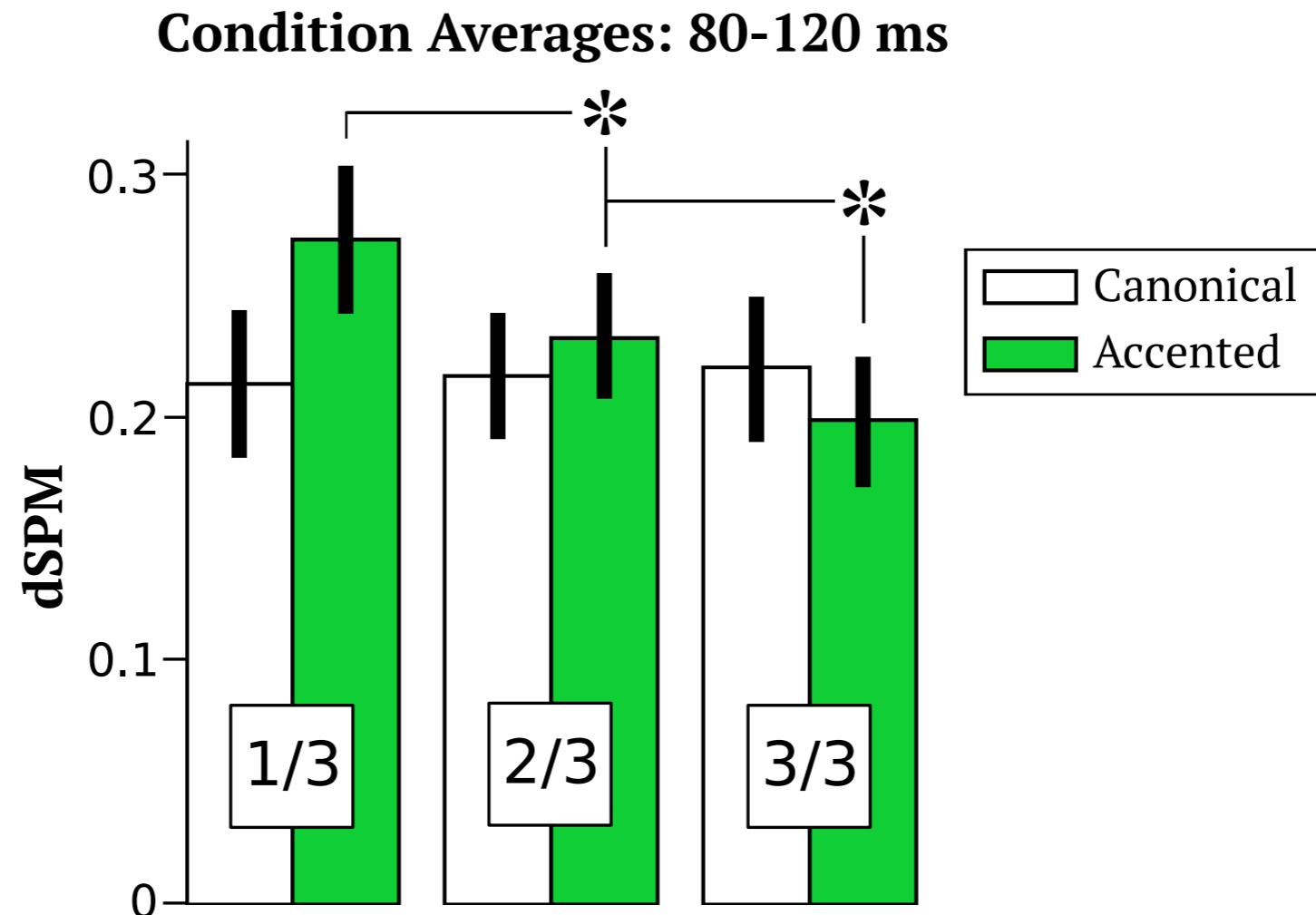
Linear Evidence



Categorical Percept



# Interpretation



- Attunement is proposed to involve **re-tuning perceptual boundaries** between phonological categories (Norris et al., 2003; Kraljic and Samuel, 2005, 2006, 2007; Maye et al., 2008; see Samuel & Kraljic, 2009 for a review)

