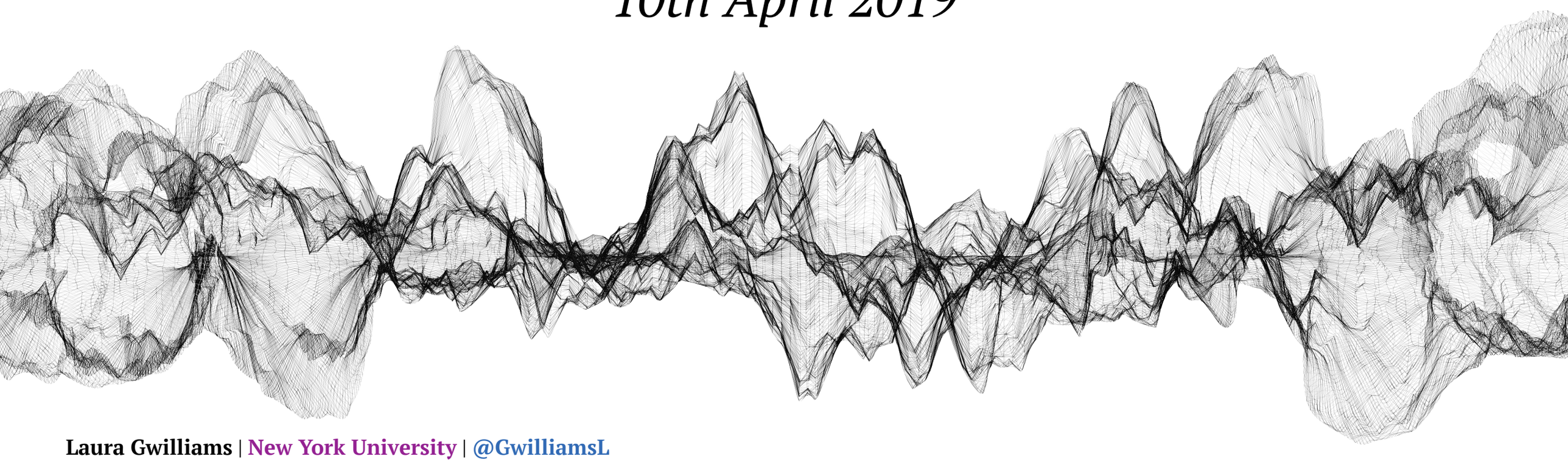
NEW YORK UNIVERSITY

Towards a mechanistic account of speech comprehension in the human brain

Laura Gwilliams

10th April 2019



Laura Gwilliams | **New York University** | **@GwilliamsL**

Collaborators

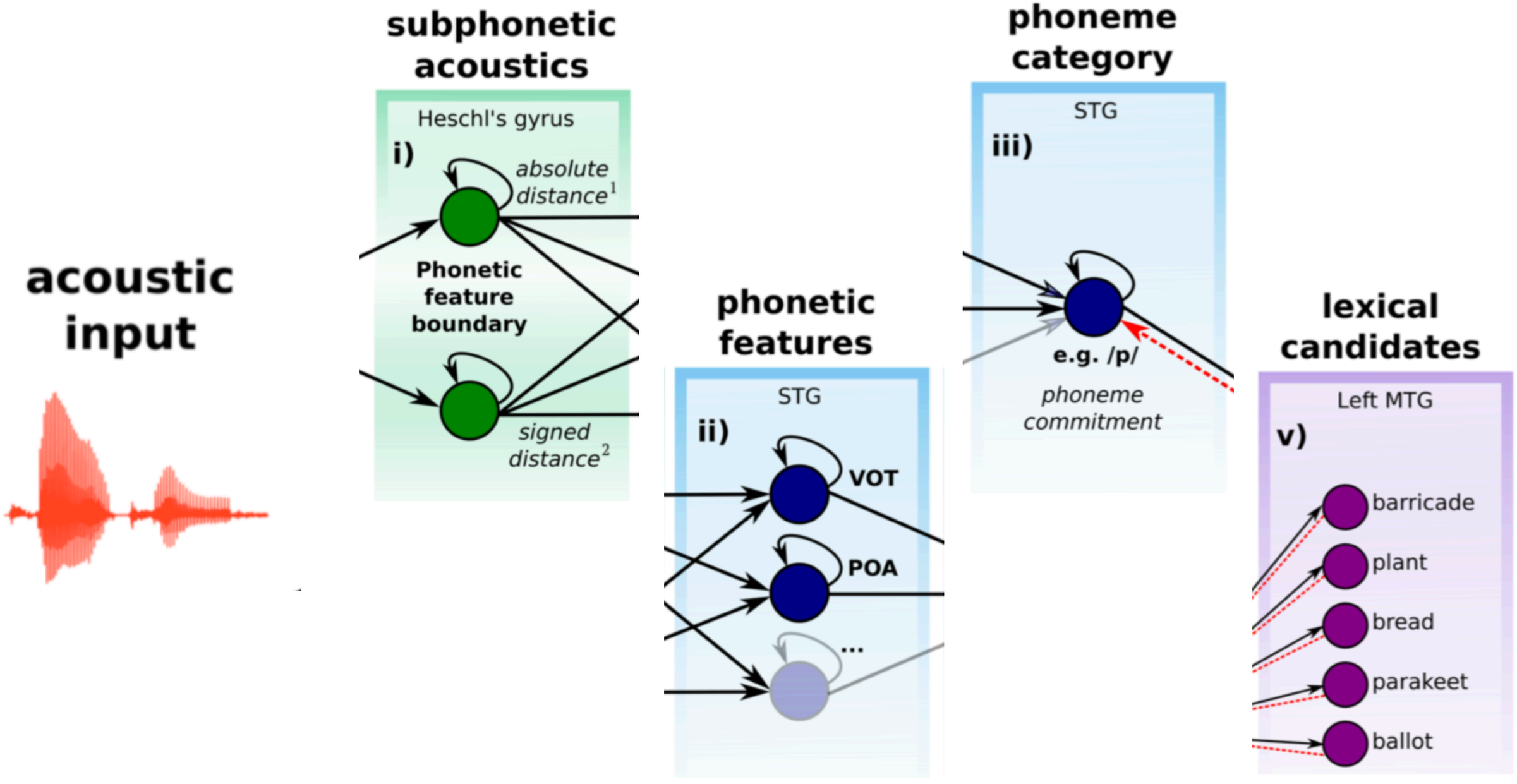
Tal Linzen

David Poeppel

Alec Marantz

Gwilliams, Poeppel, Marantz & Linzen (2018), *CMCL*
Gwilliams, Linzen, Poeppel & Marantz (2018), *JNeuro*

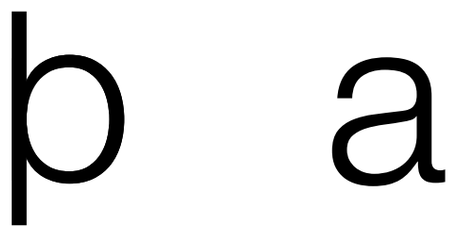
Putting together the processing pieces



Laura Gwilliams | New York University | @GwilliamsL

Roadmap

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



Roadmap

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

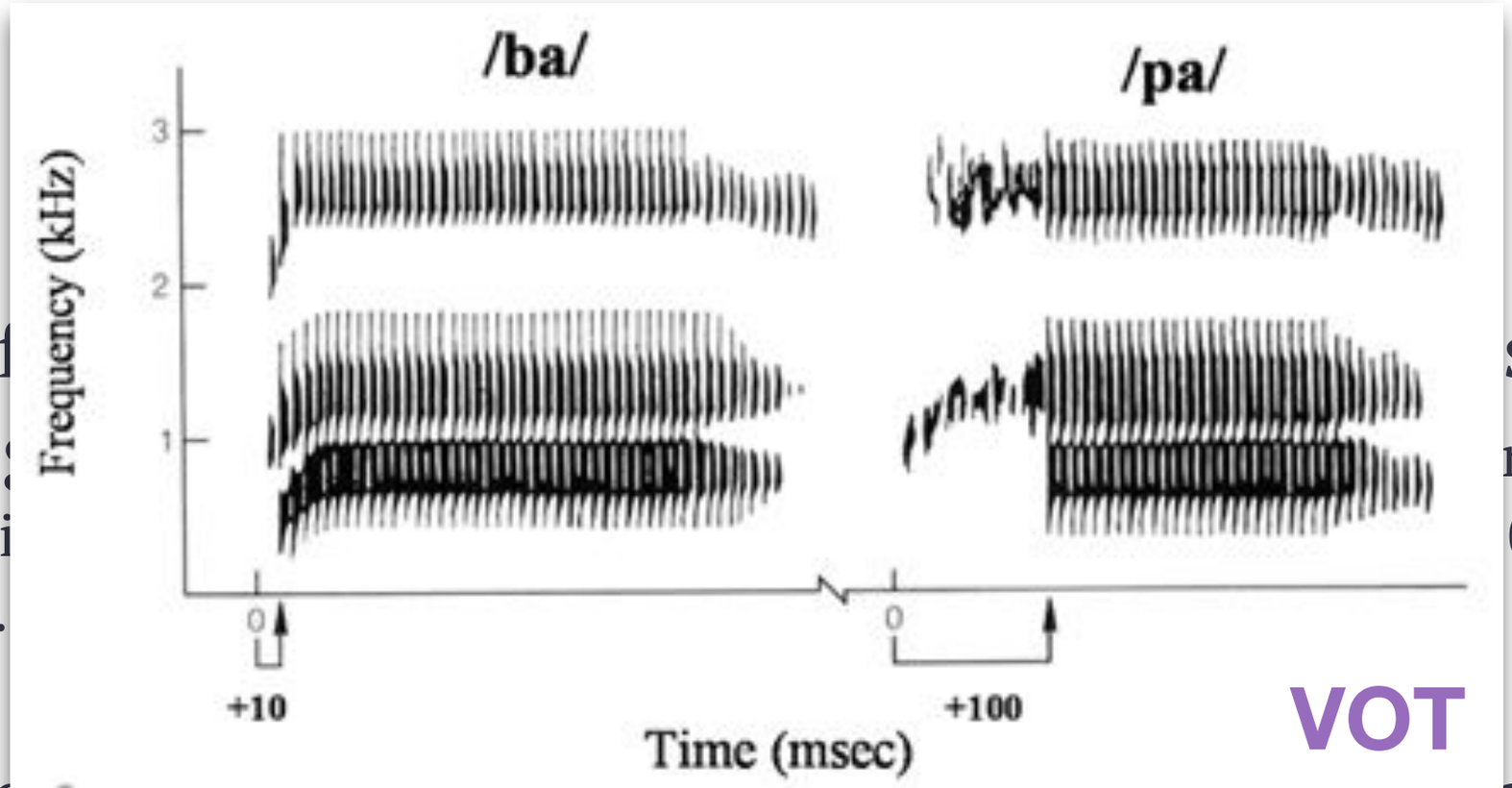
þ a r a k **ee** t

Roadmap

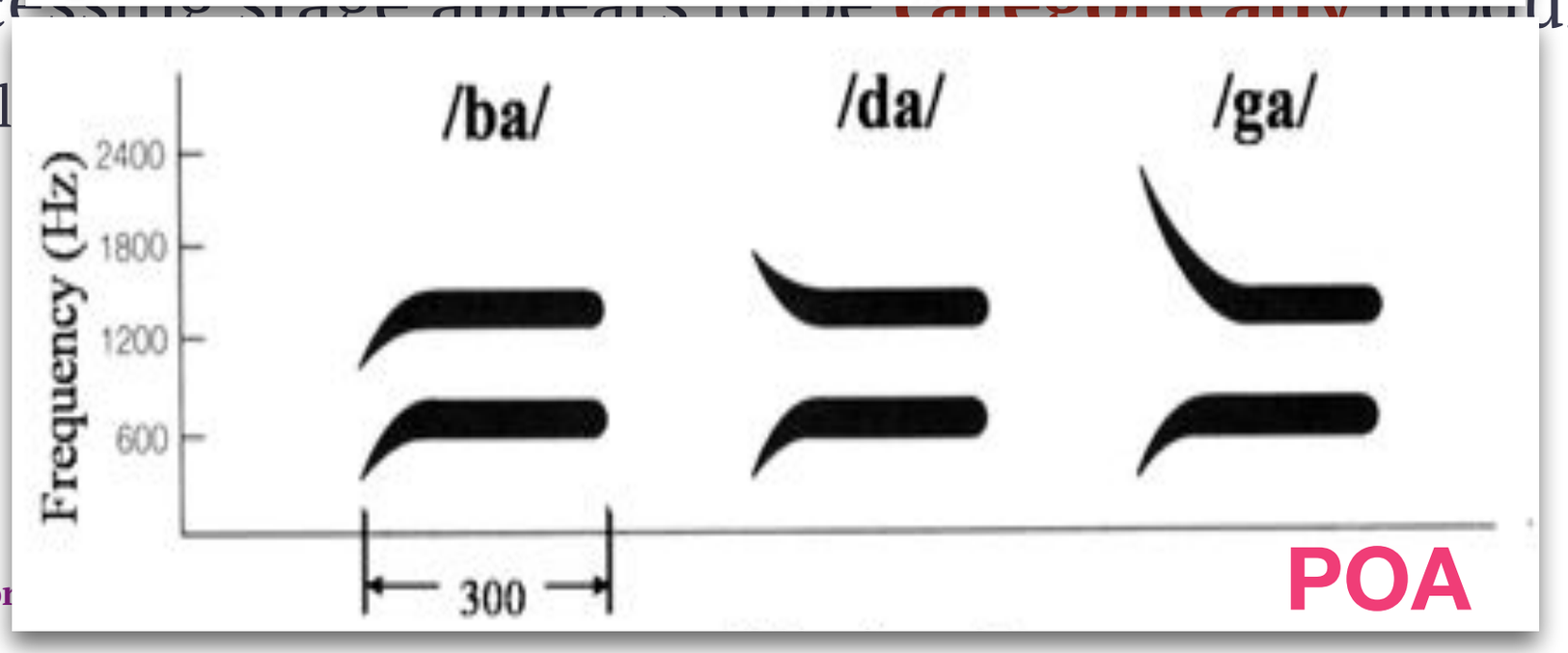
- Completed projects
  - **Bottom-up processes**: Transforming acoustic signal into discrete phonological categories
  - **Top-down processes**: Revising that categorisation based on subsequent context
- Future directions
  - Related processes in **continuous speech**
  - **Accent** adaptation
  - **Pitch perception** in music cognition

Bottom-up processing of phonemes

- Phonetic f… superior temporal g… r et al., 2003; Papani… 014; Di Liberto et al.…

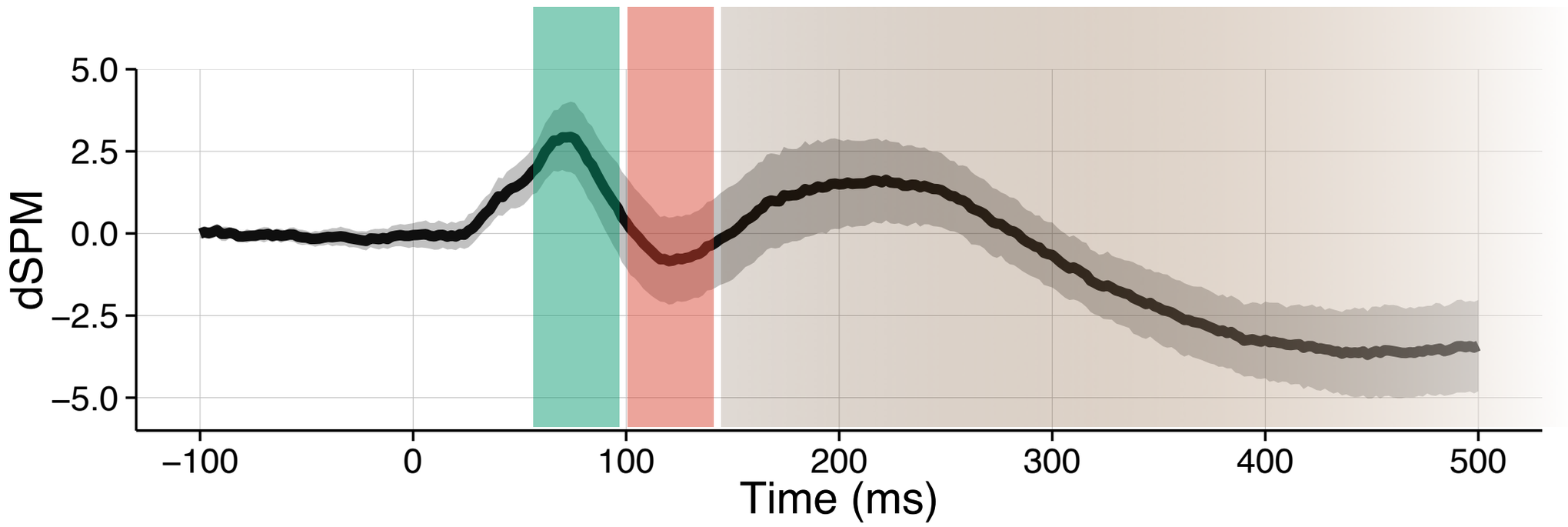


- This proce… ulated (Chang et al…

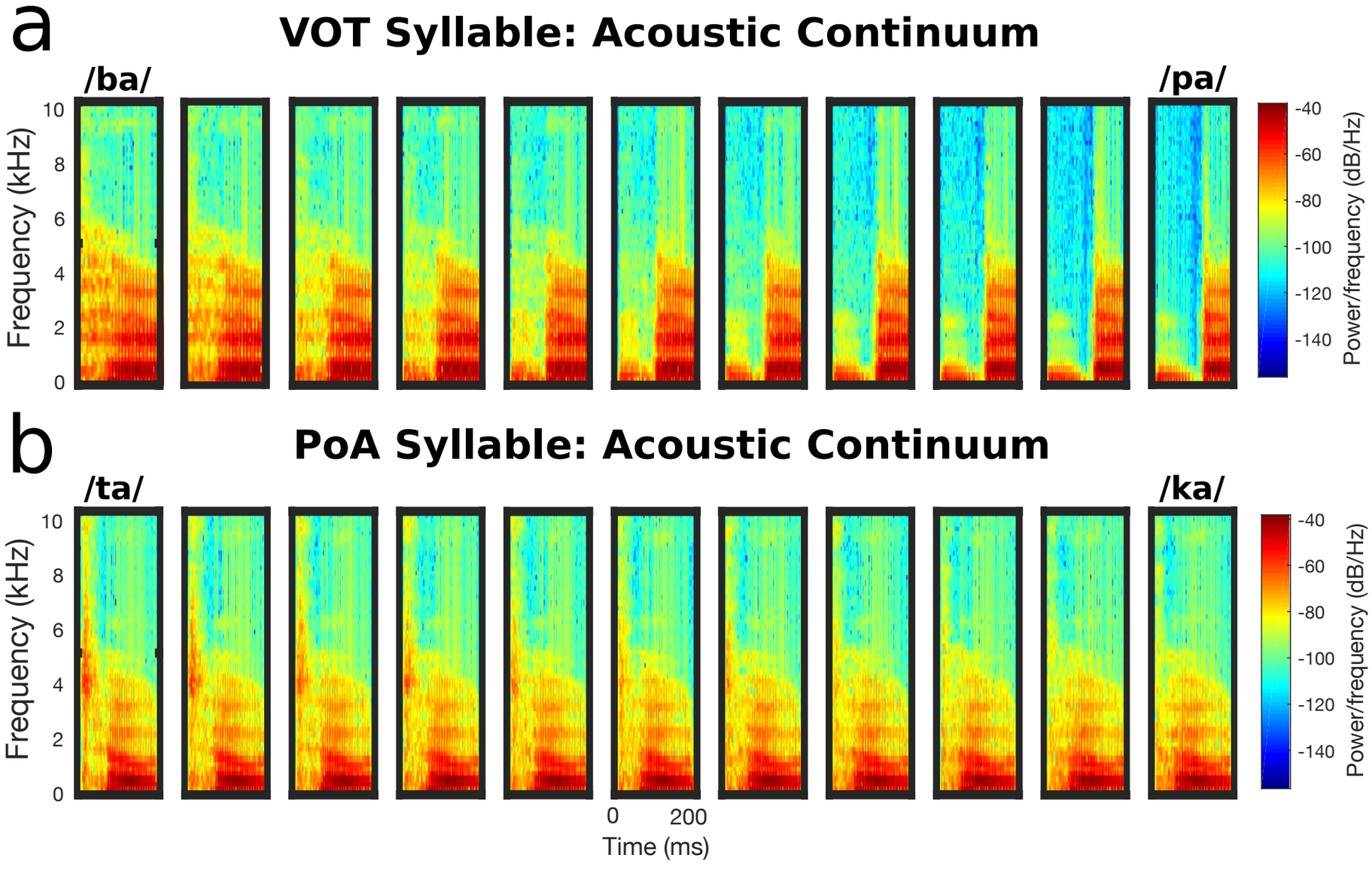


Neutralising ambiguity

At what stage of processing is phonological **ambiguity** alleviated?

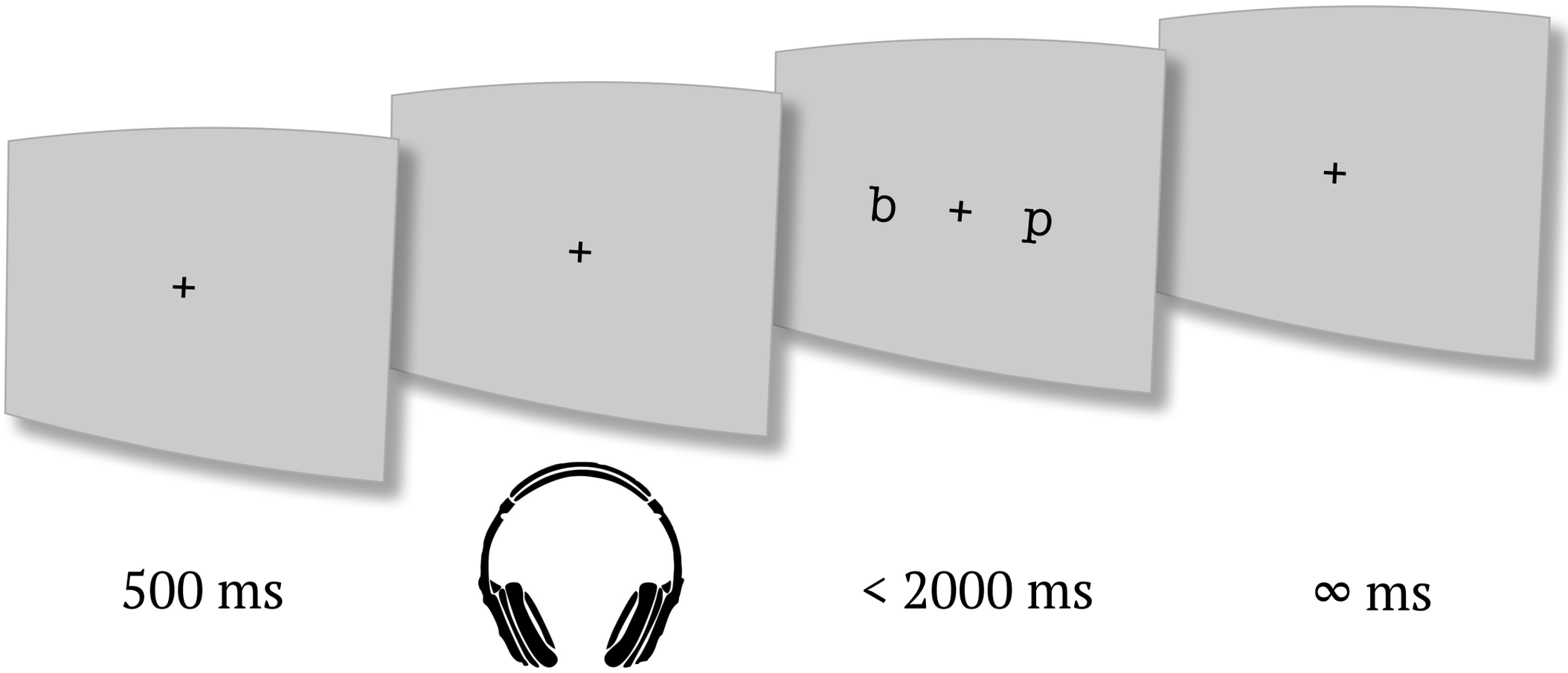


Materials

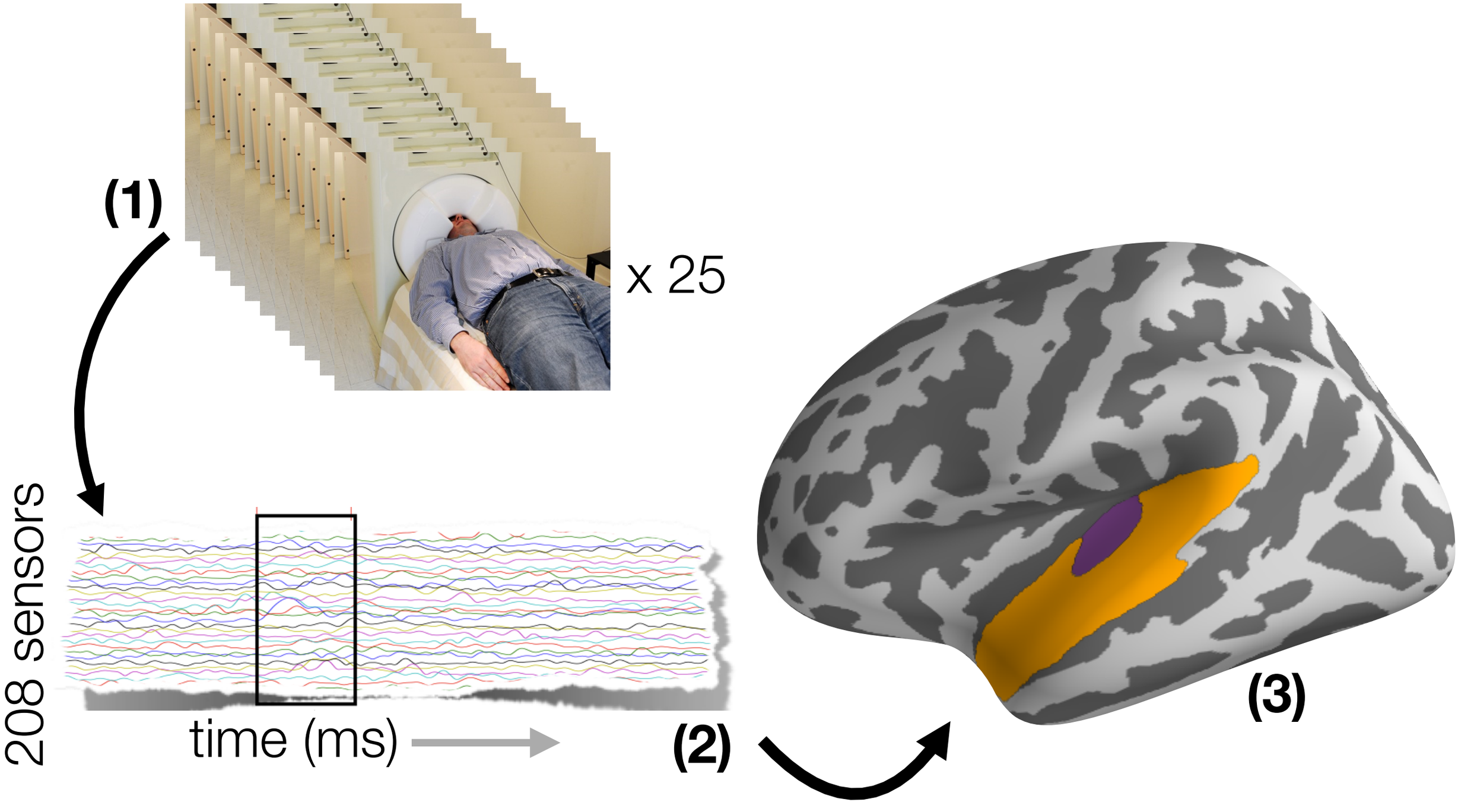


- VOT (31 pairs) {p-b, t-d, k-g} and POA (22 pairs) {t-k, p-t}

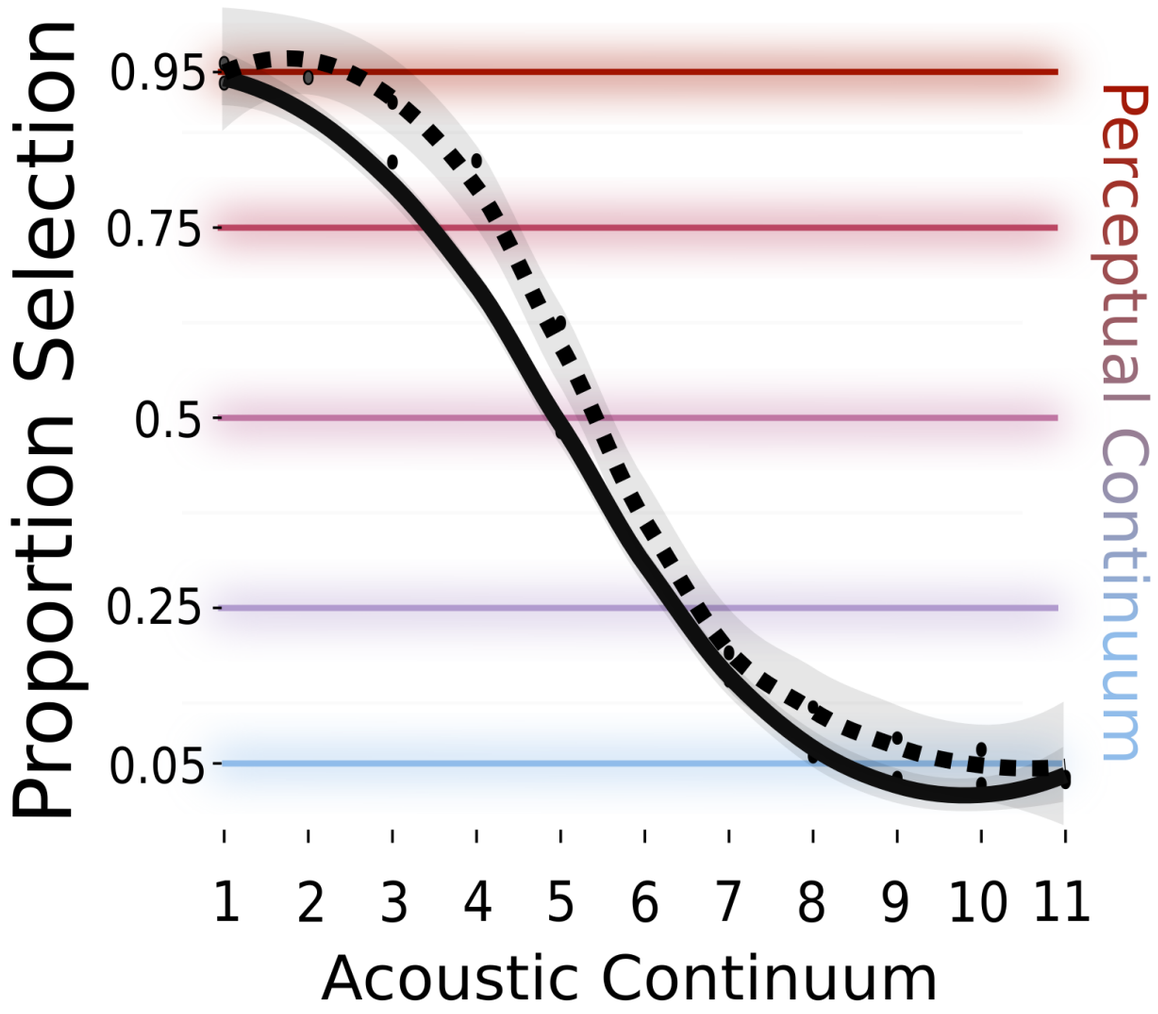
Design

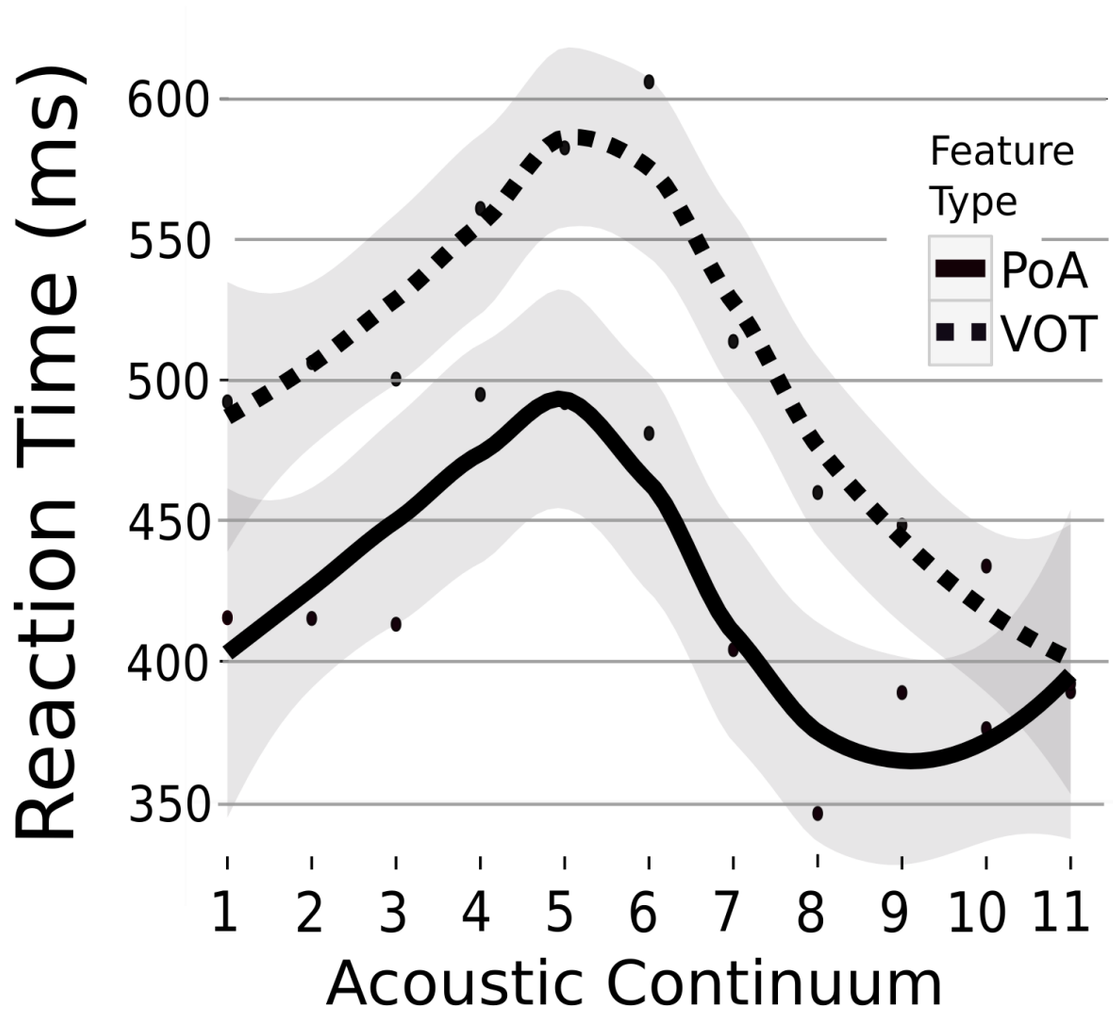


Procedure & Analysis



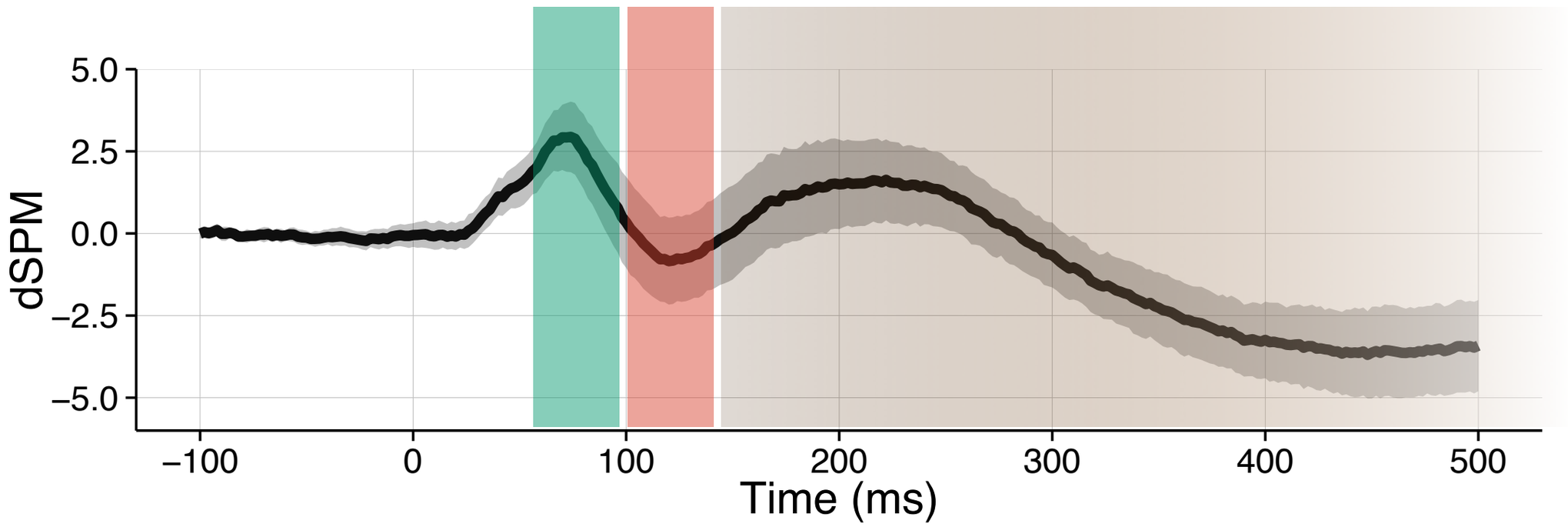
Behaviour


Proportion Selection
0.95
0.75
0.5
0.25
0.05
Perceptual Continuum
1 2 3 4 5 6 7 8 9 10 11
Acoustic Continuum

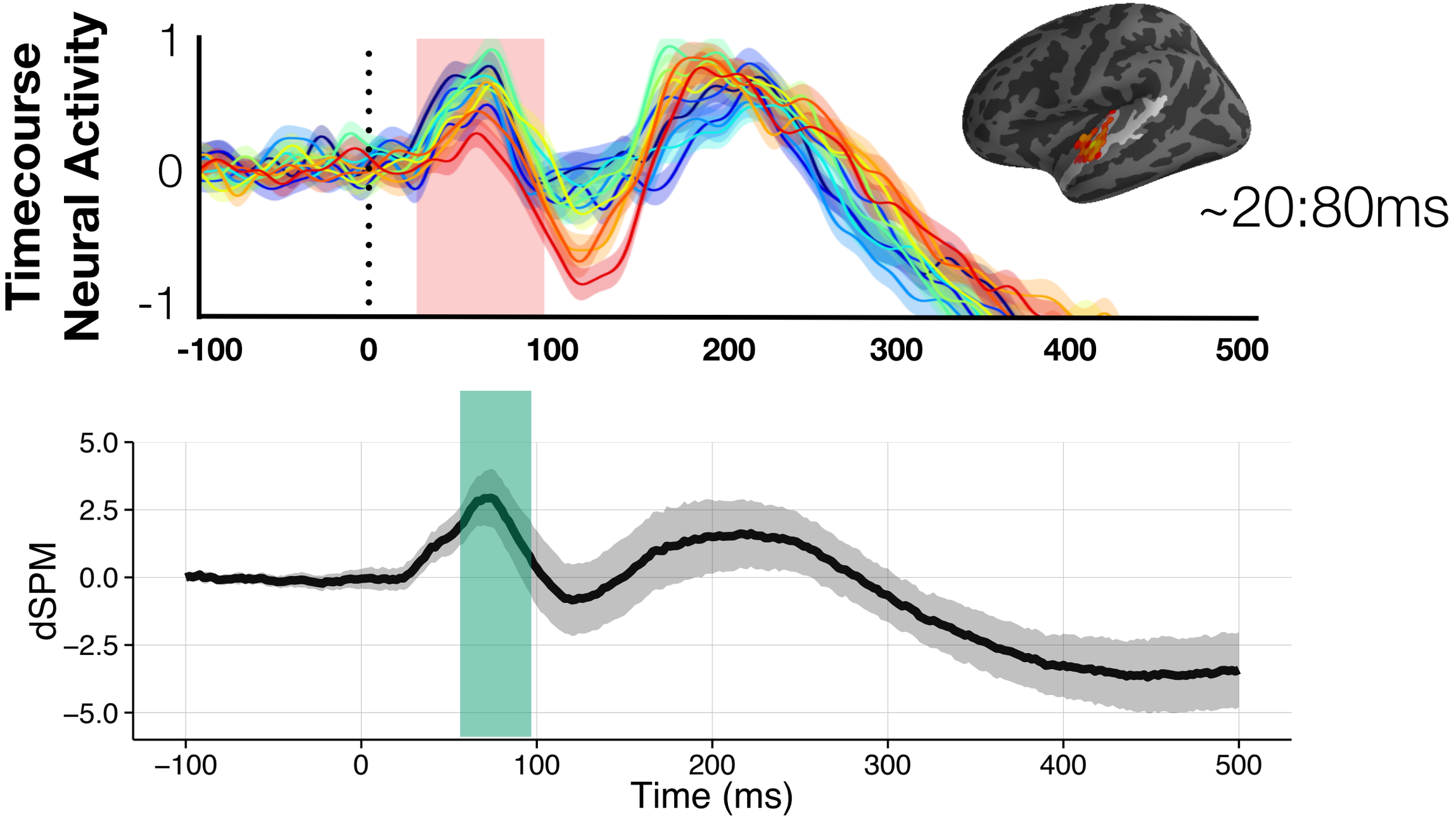

Reaction Time (ms)
600
550
500
450
400
350
Feature Type
PoA
VOT
1 2 3 4 5 6 7 8 9 10 11
Acoustic Continuum

Neutralising ambiguity

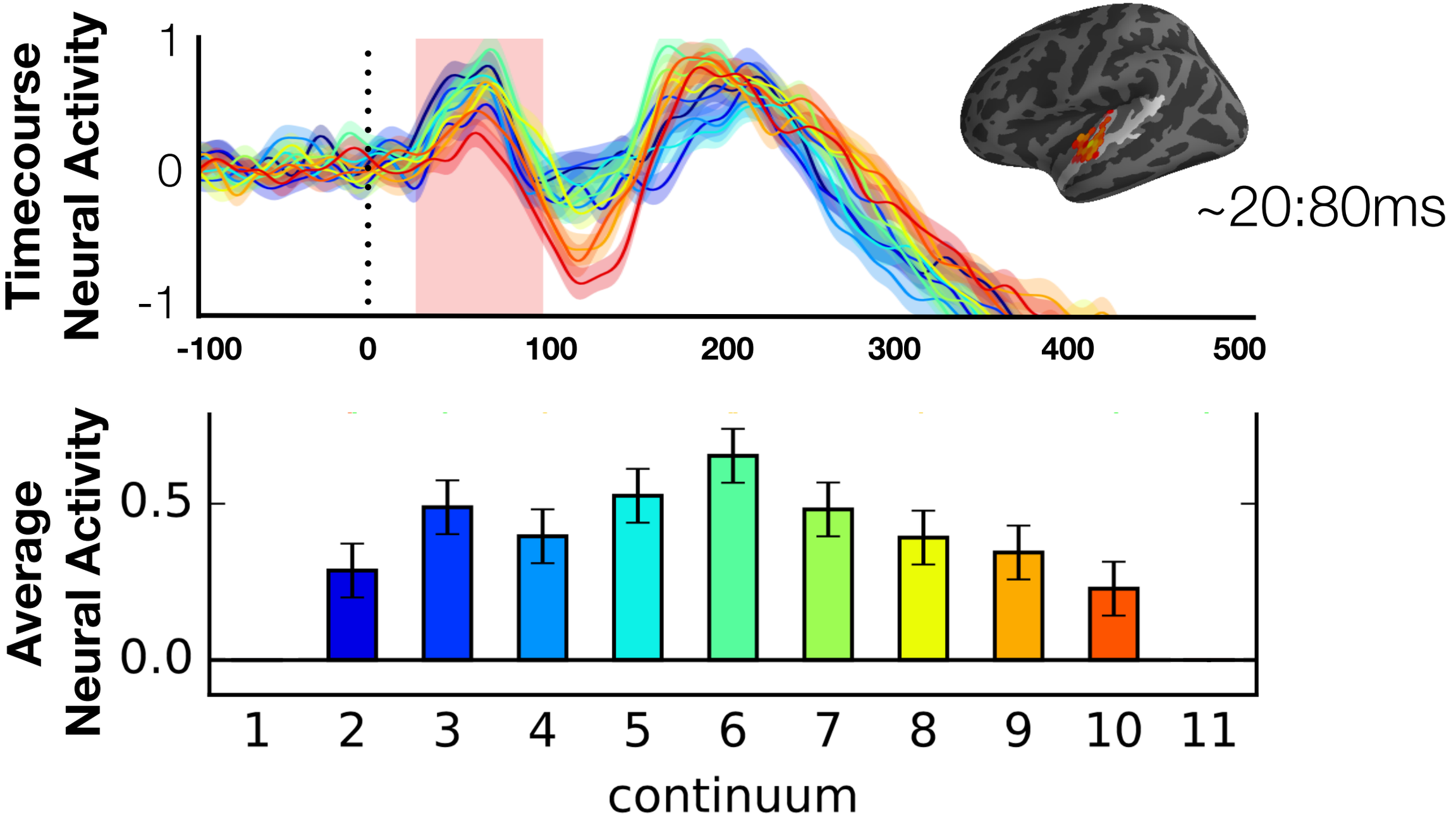
At what stage of processing is phonological **ambiguity** alleviated?



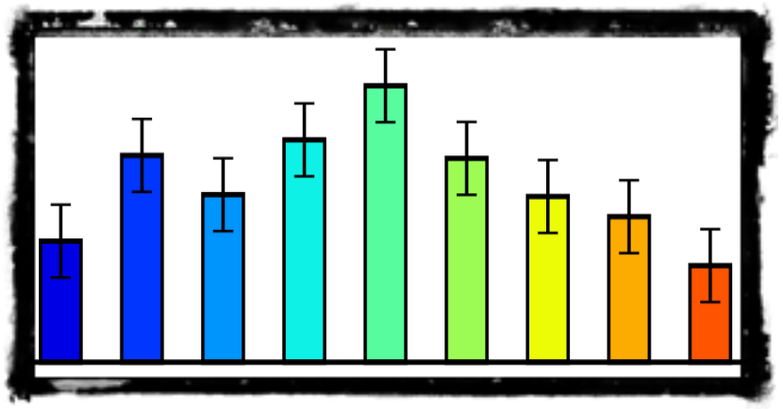
Early ambiguity responses in Heschl's gyrus



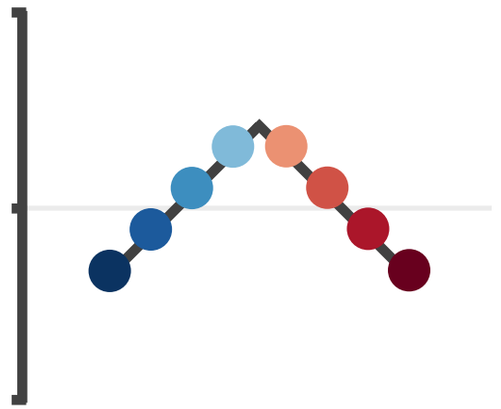
Early ambiguity responses in Heschl's gyrus

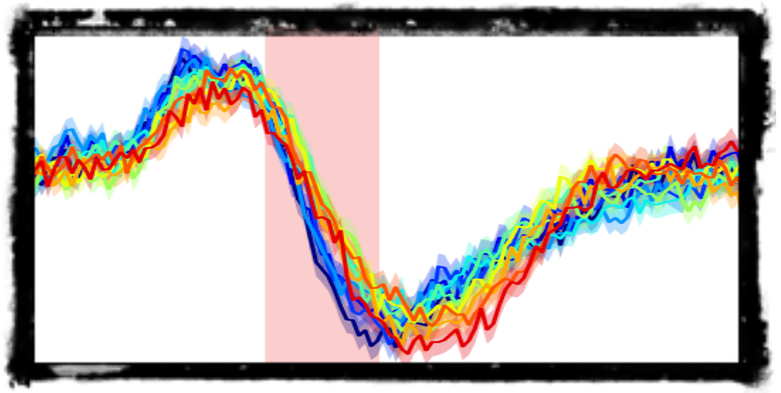


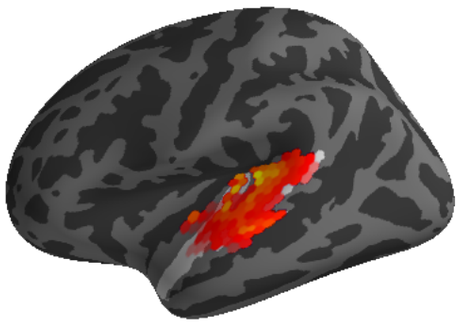
Replicating the categorical trajectory using MEG



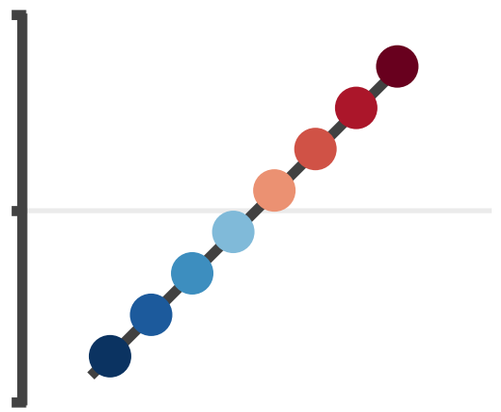
~20:80ms
ambiguity

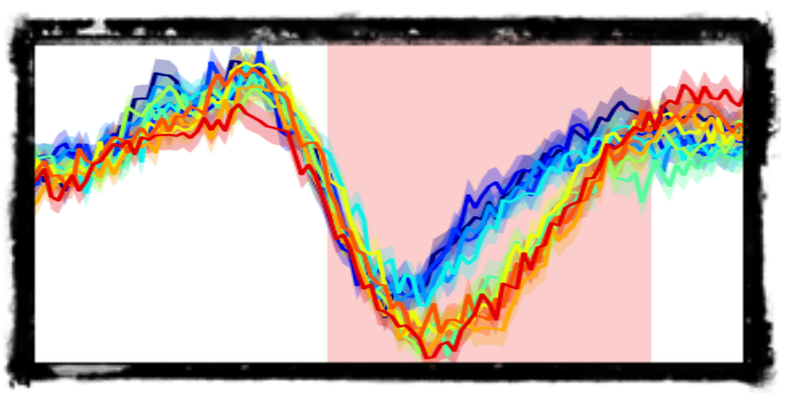




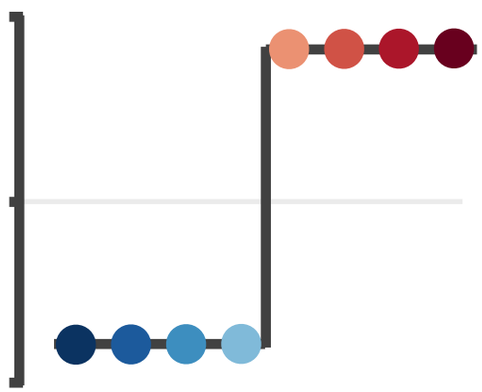


~80:150ms
linear



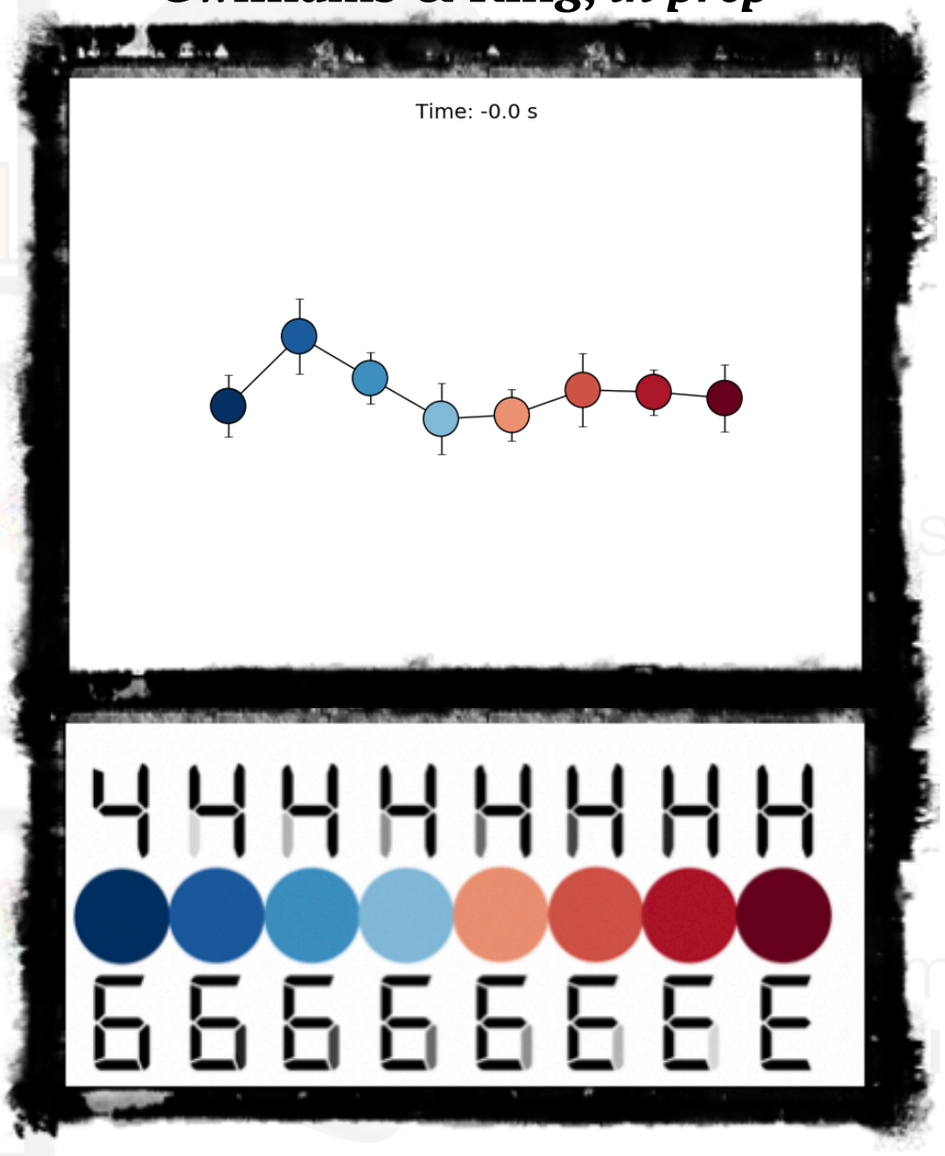


~100:200ms
categorical



Interesting links in a different domain

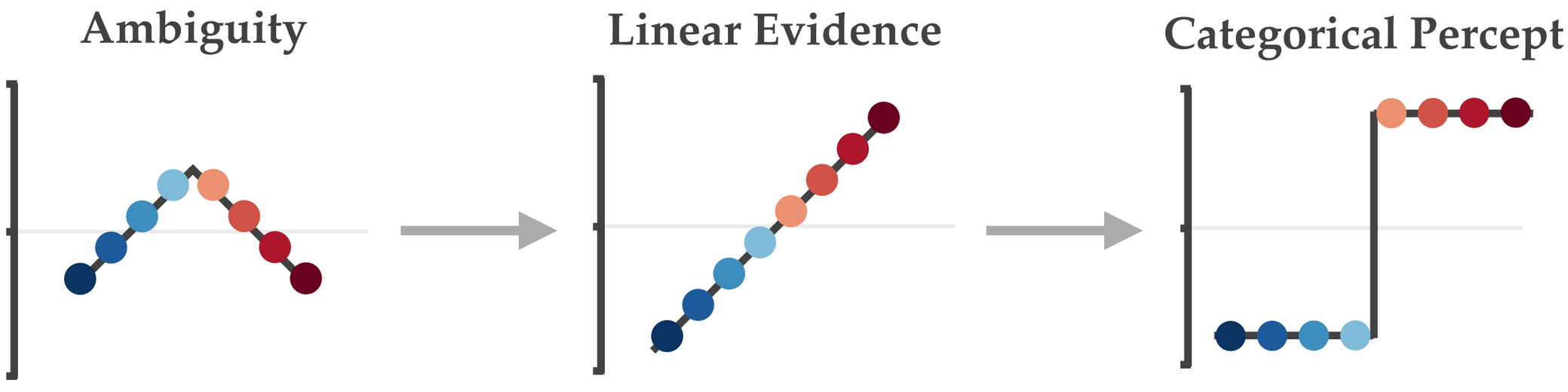
Gwilliams & King, *in prep*



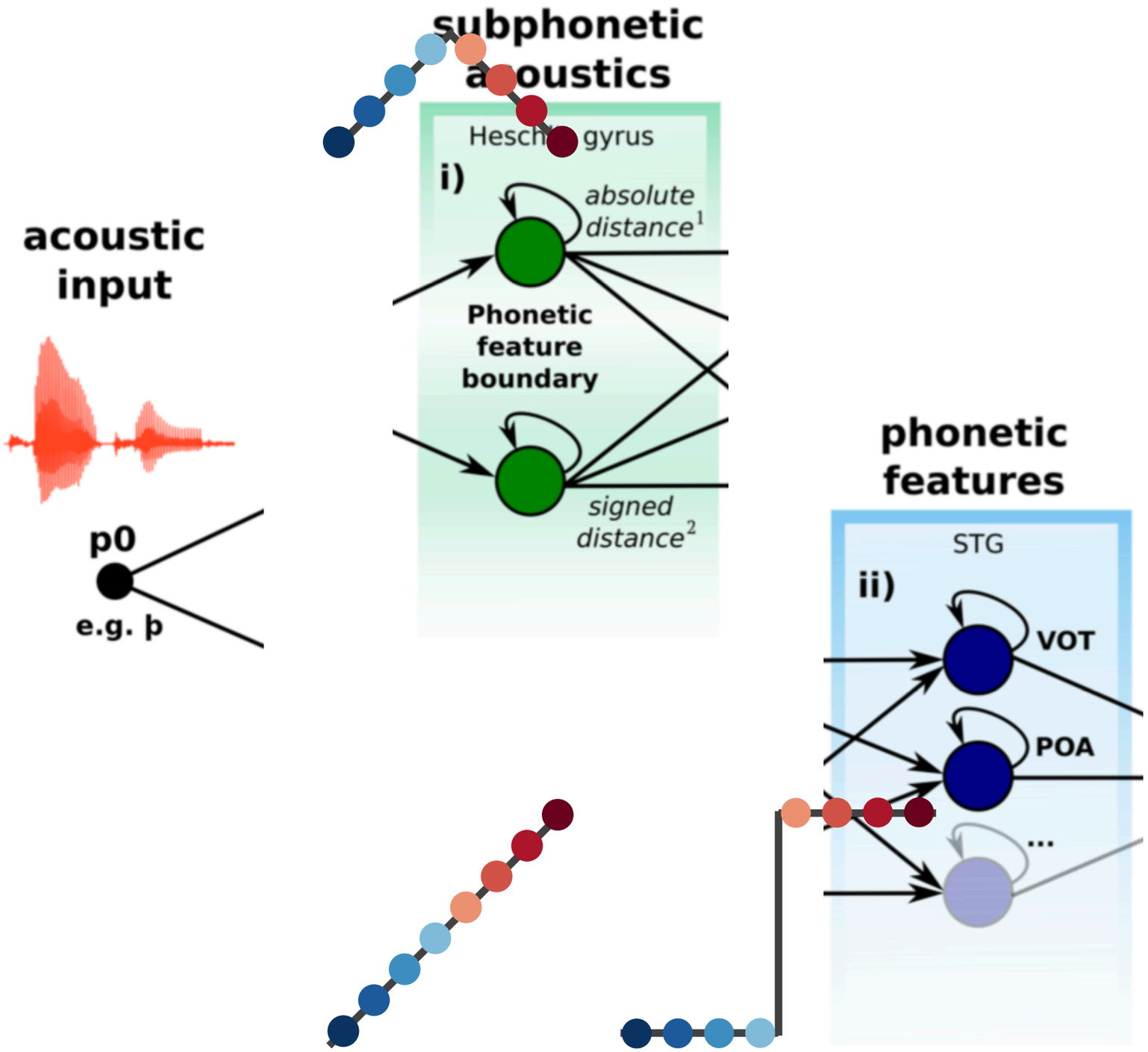
Experiment 1 Conclusions

- 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?)



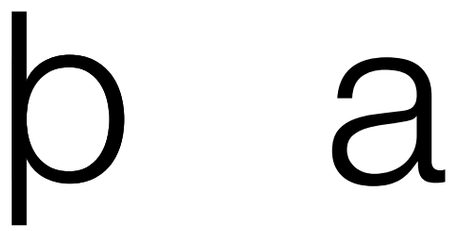
Putting together the processing pieces



Laura Gwilliams | New York University | @GwilliamsL

Roadmap

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



Roadmap

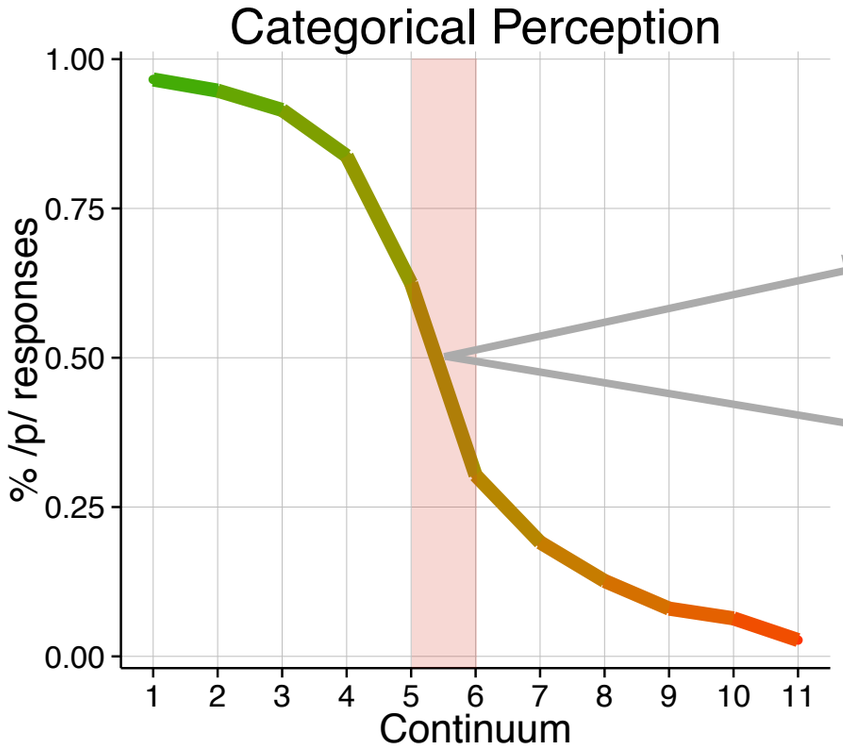
- Completed projects
  - **Bottom-up processes**: Transforming acoustic signal into discrete phonological categories
  - **Top-down processes**: Revising that categorisation based on subsequent context

þ 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



p

b

ee t

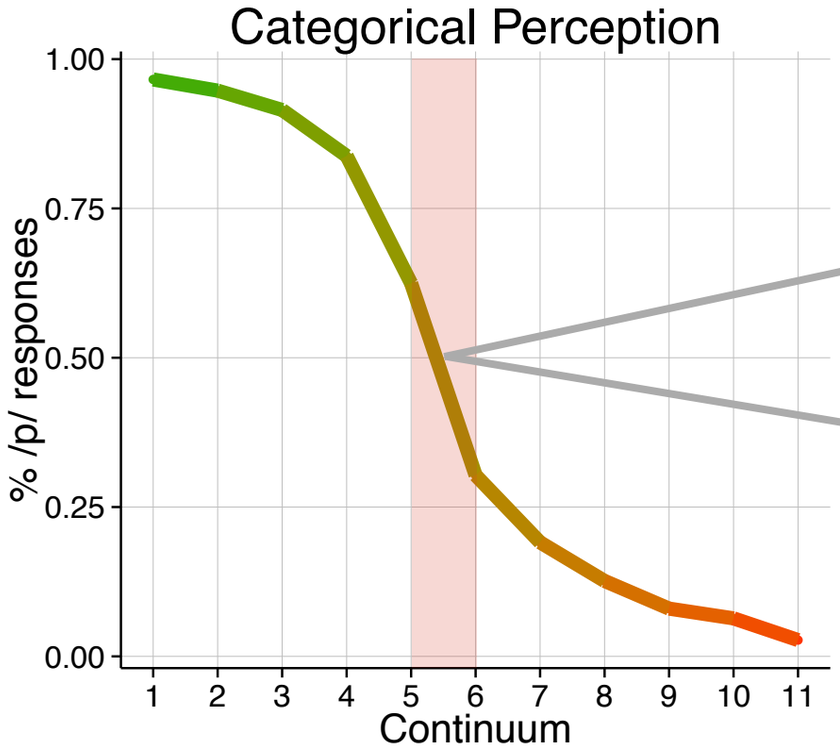
ai d

(this is a parakeet)

(this is a barricade)

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

Future Influences on Perception



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)

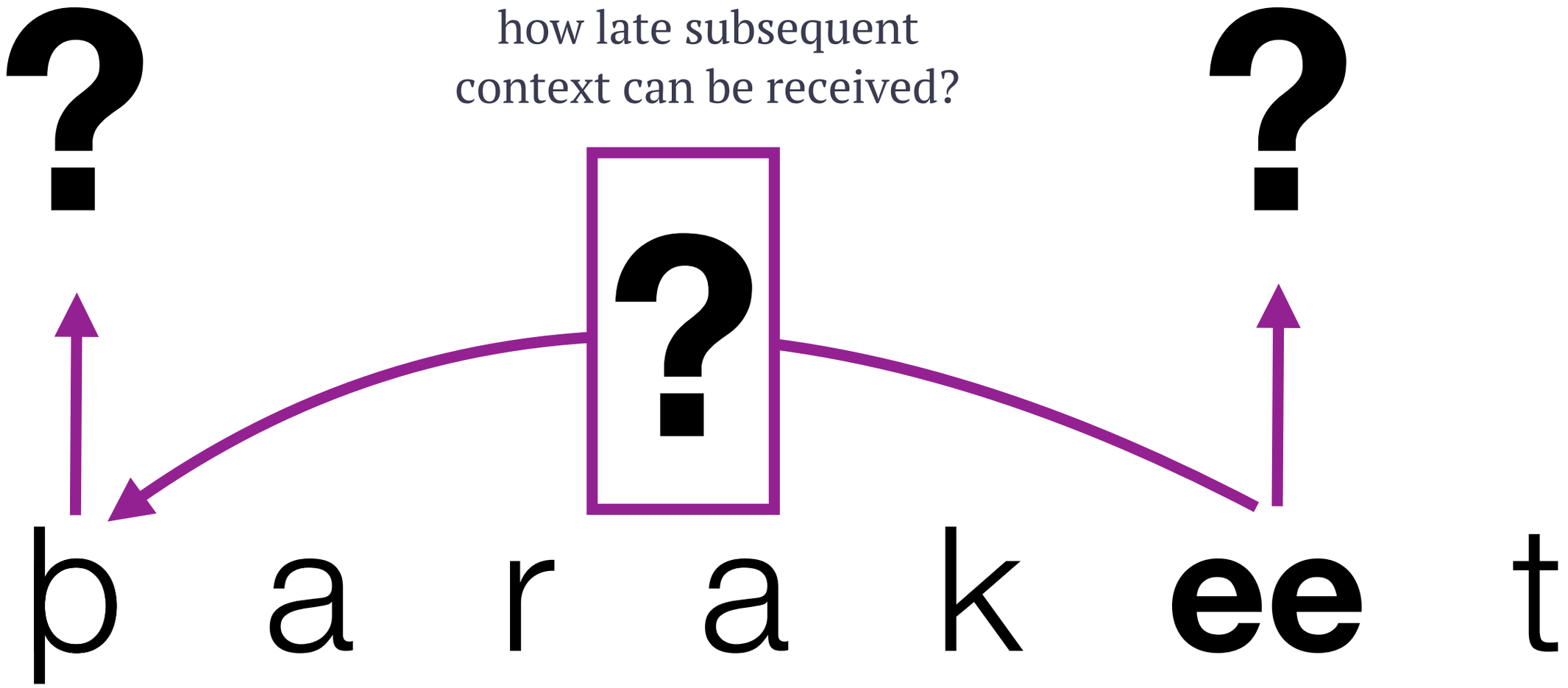
Today's Questions

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

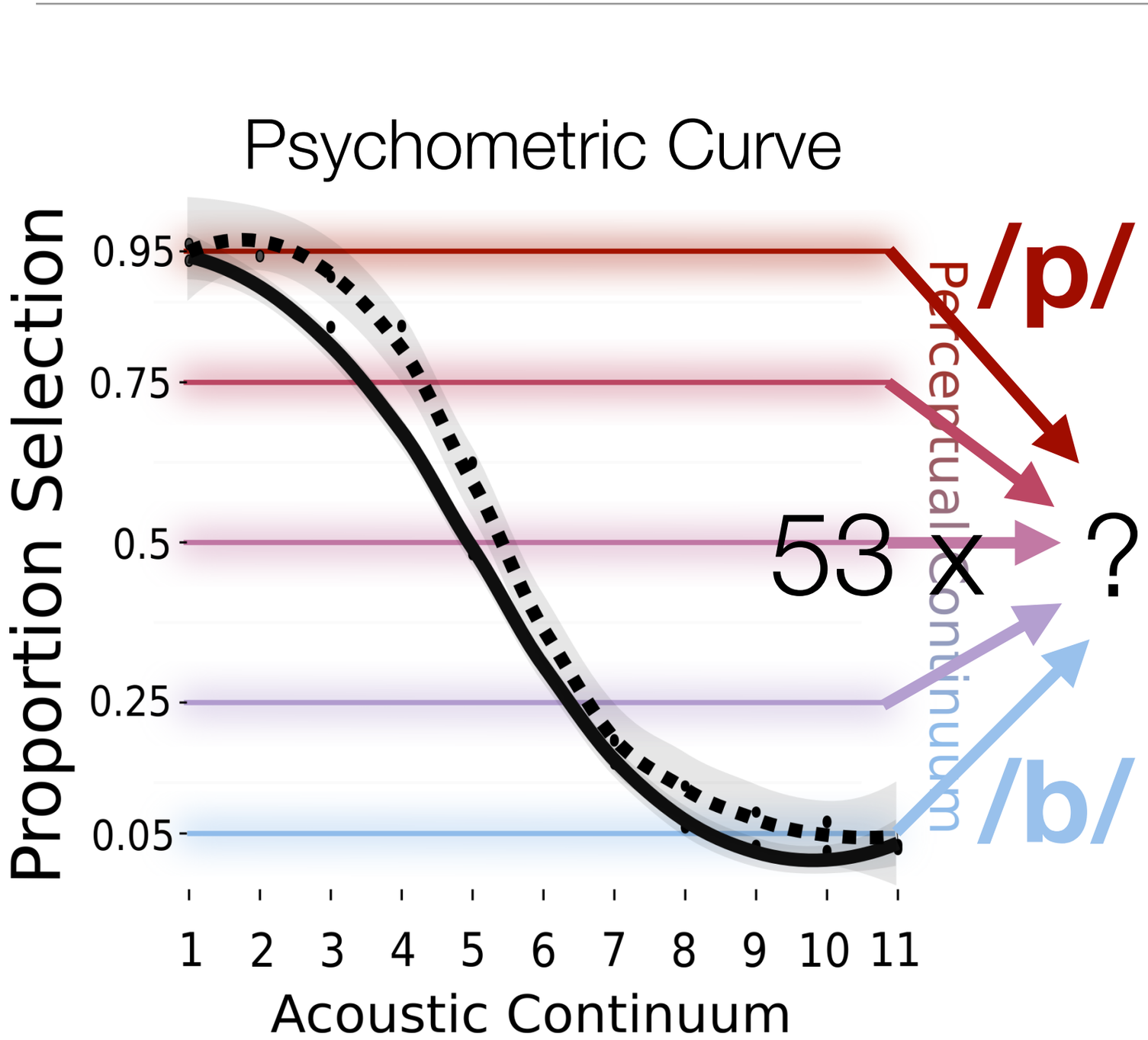
How does the auditory cortex **respond** to phonological ambiguity?

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

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

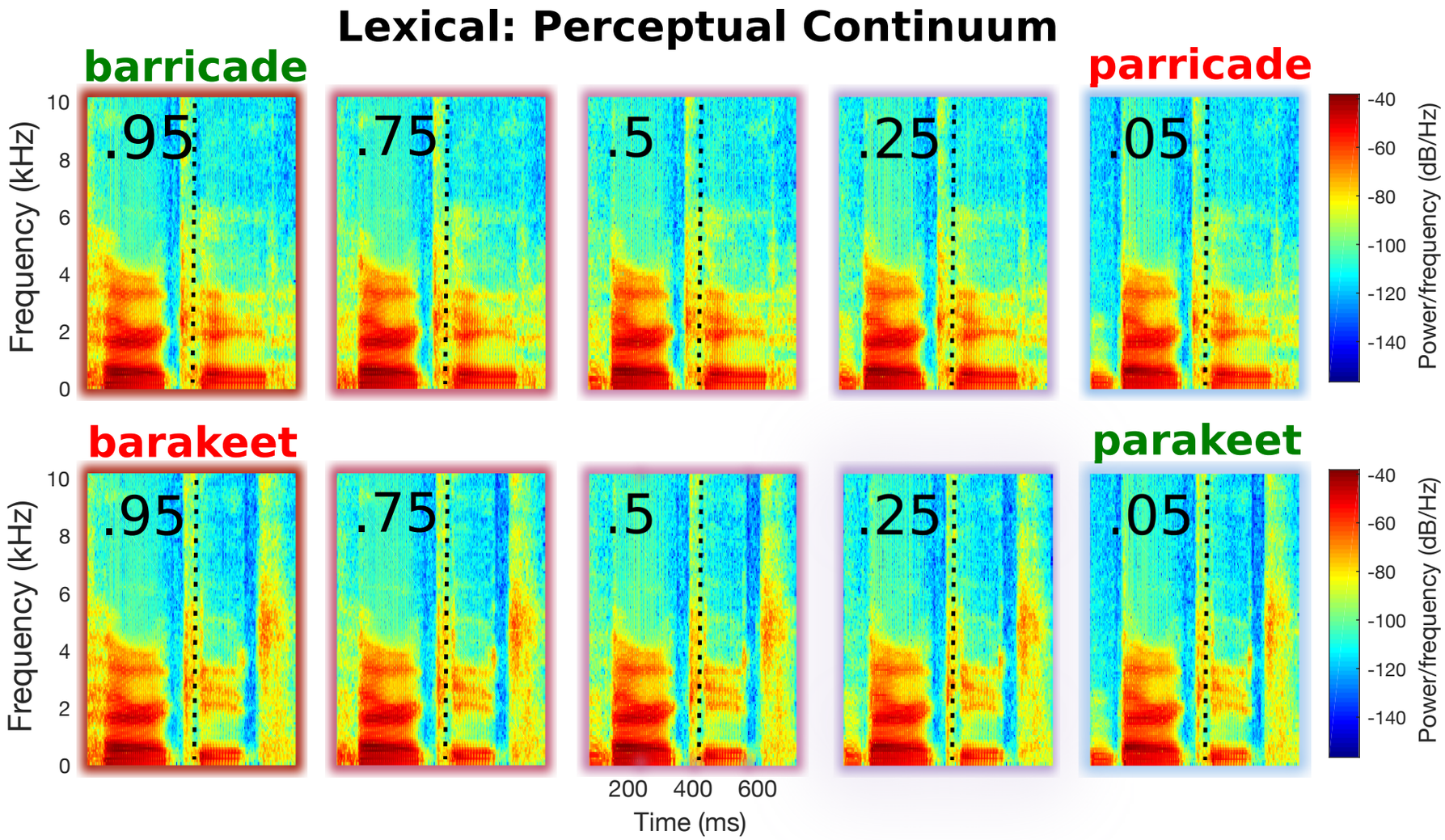


Design & Materials



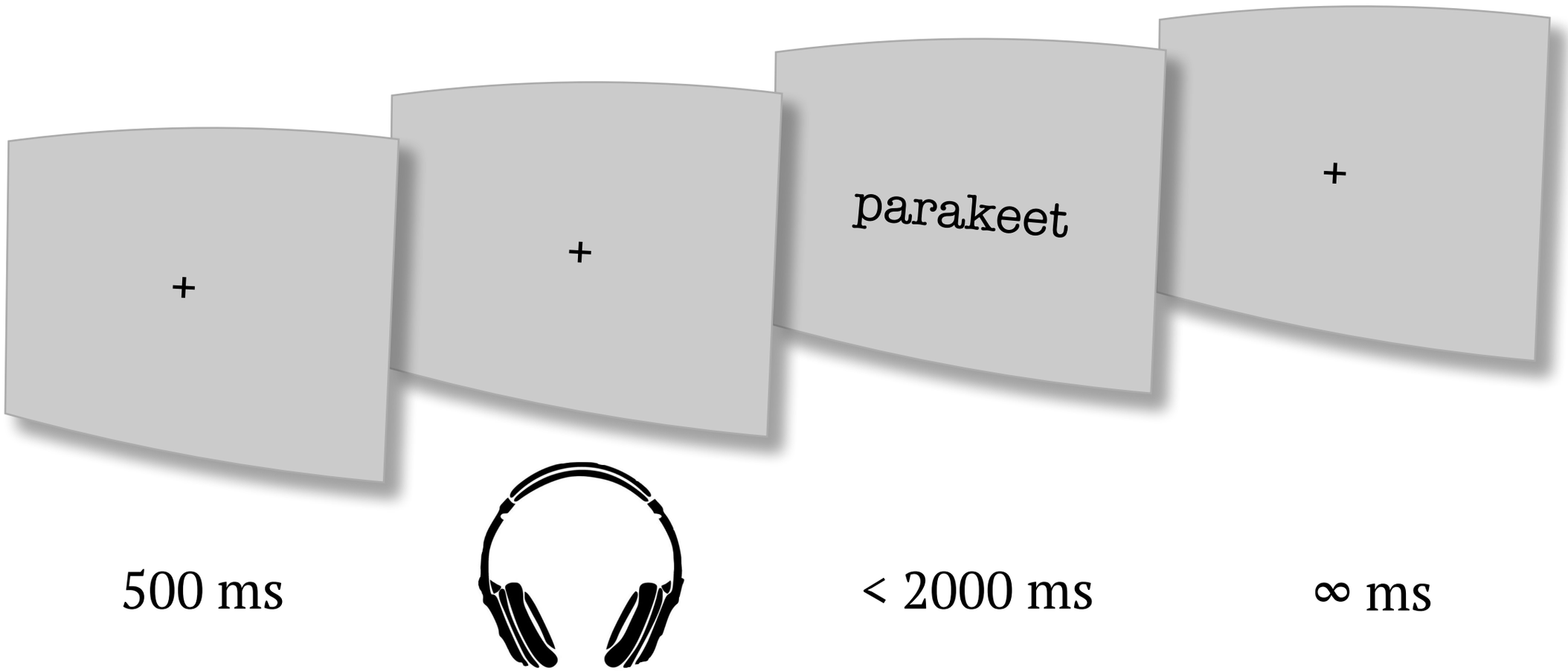
p p p b b

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}

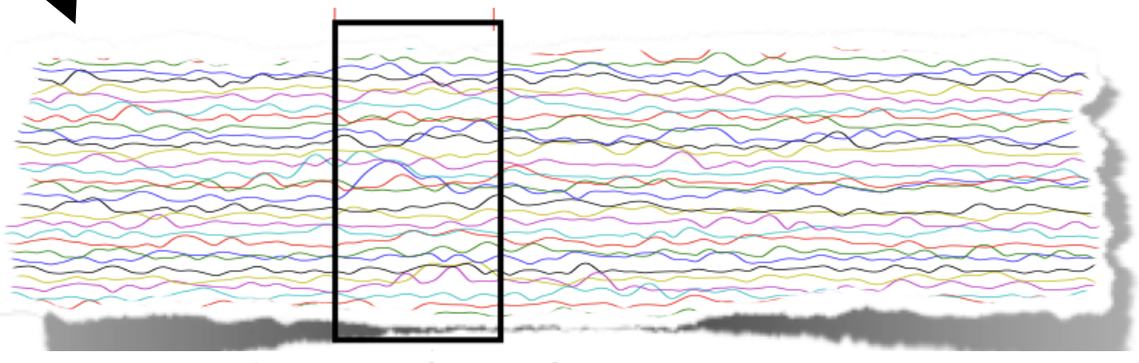
Design & Materials



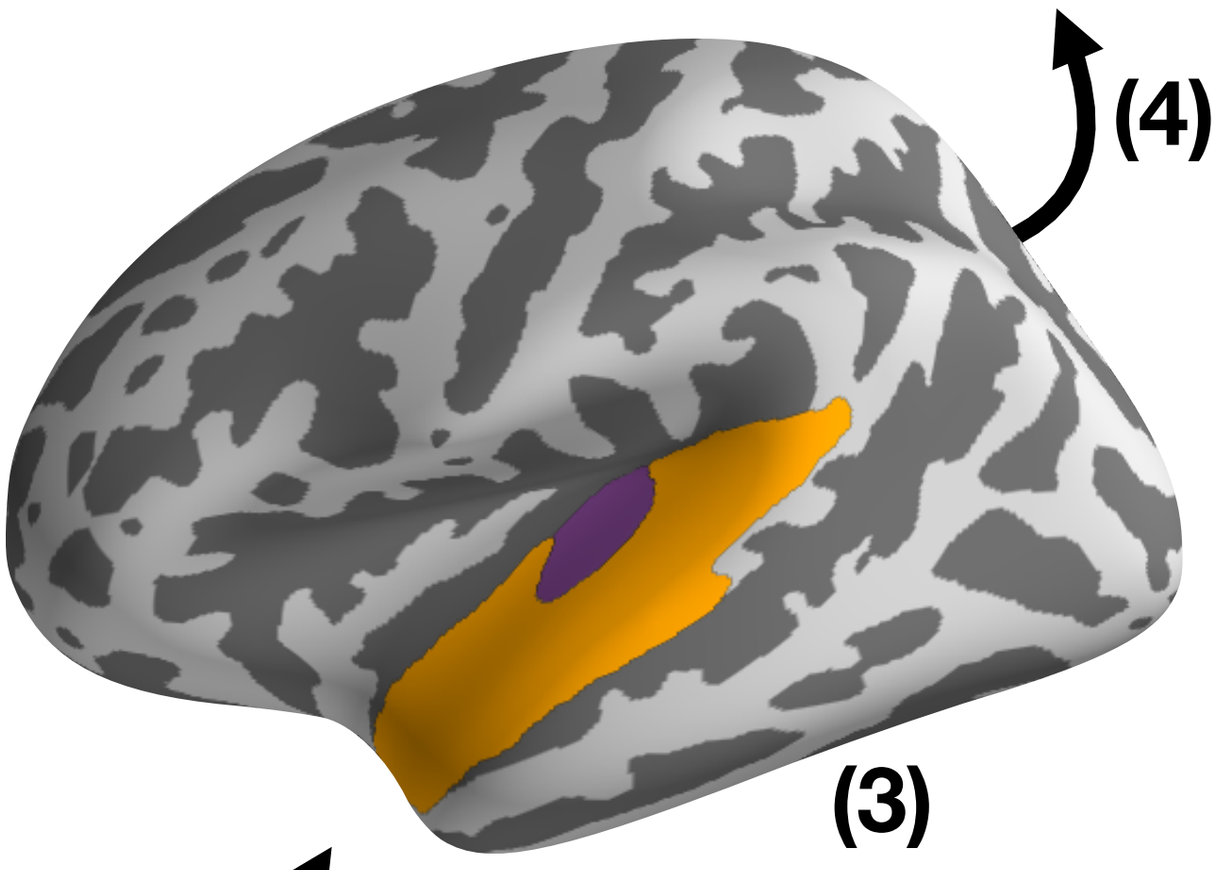
Procedure & Analysis

(1)

x 25



(2)



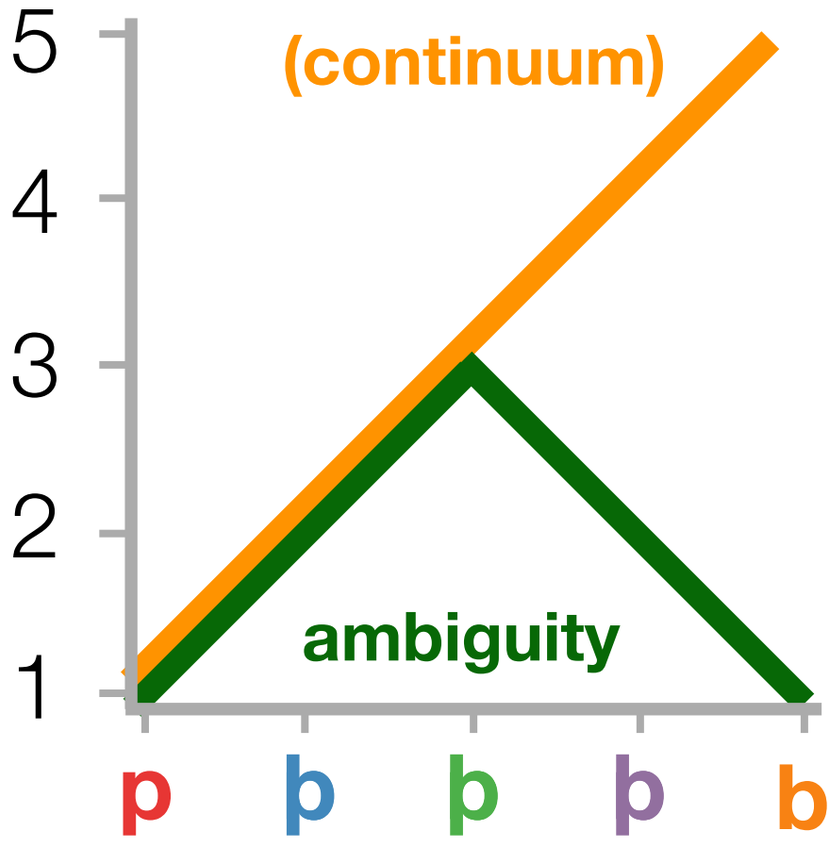
(3)

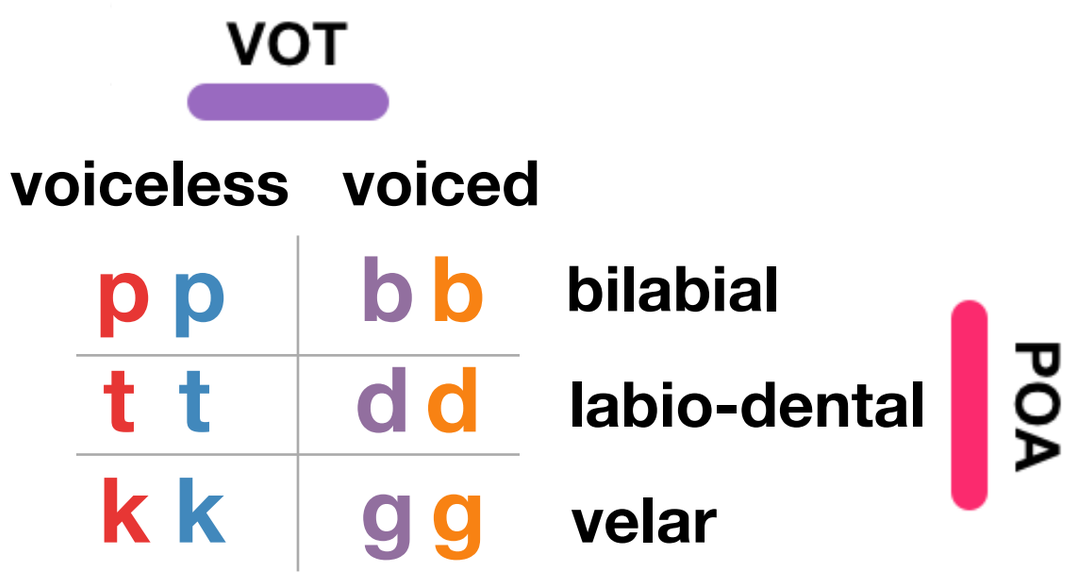
(4)

- Multiple regression at each ms and vertex
- Spatio-temporal cluster test on beta coefficients

(for more info see Gwilliams, Lewis & Marantz, 2016)

Four Experimental Variables



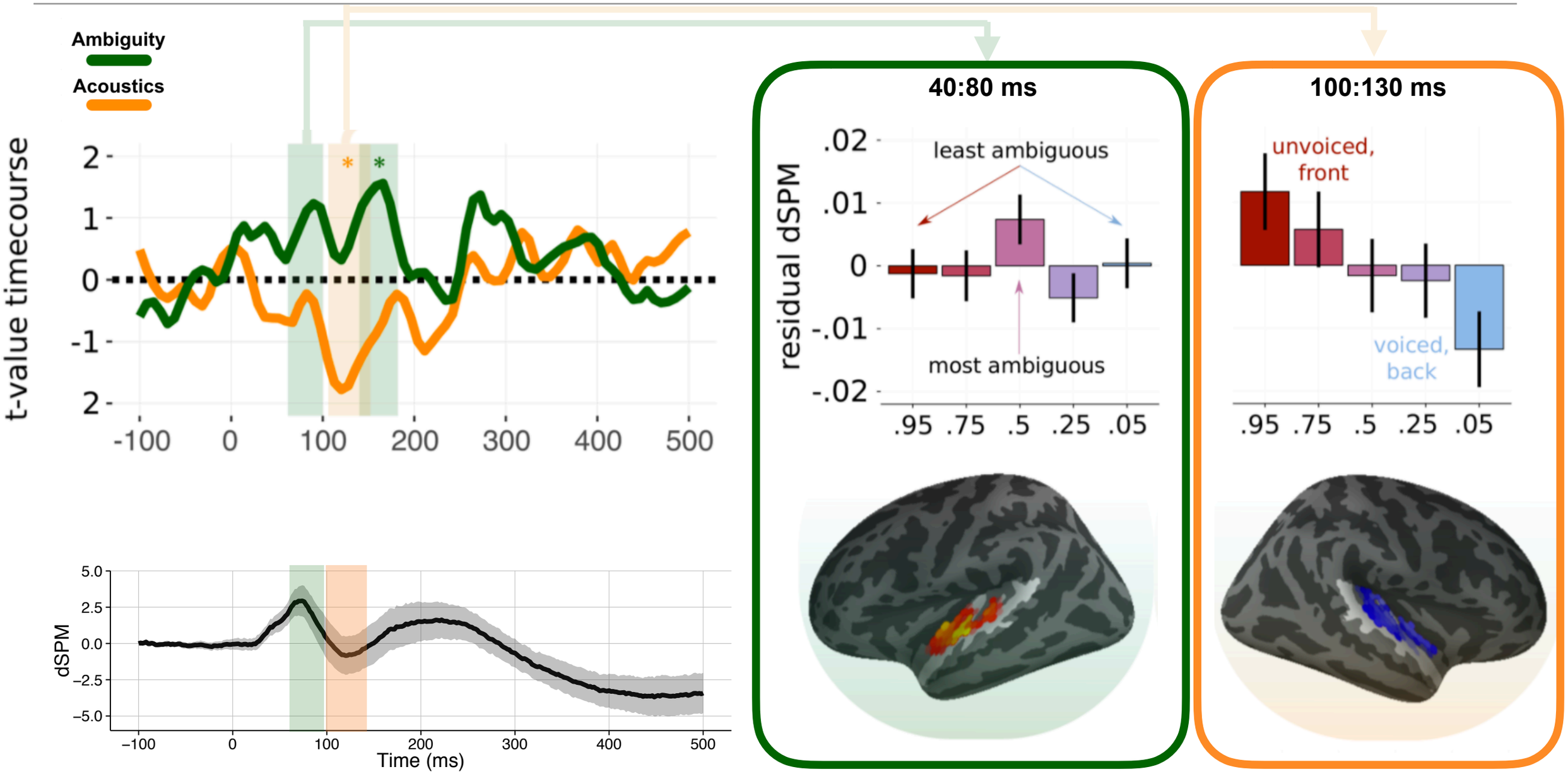


Today's Questions

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

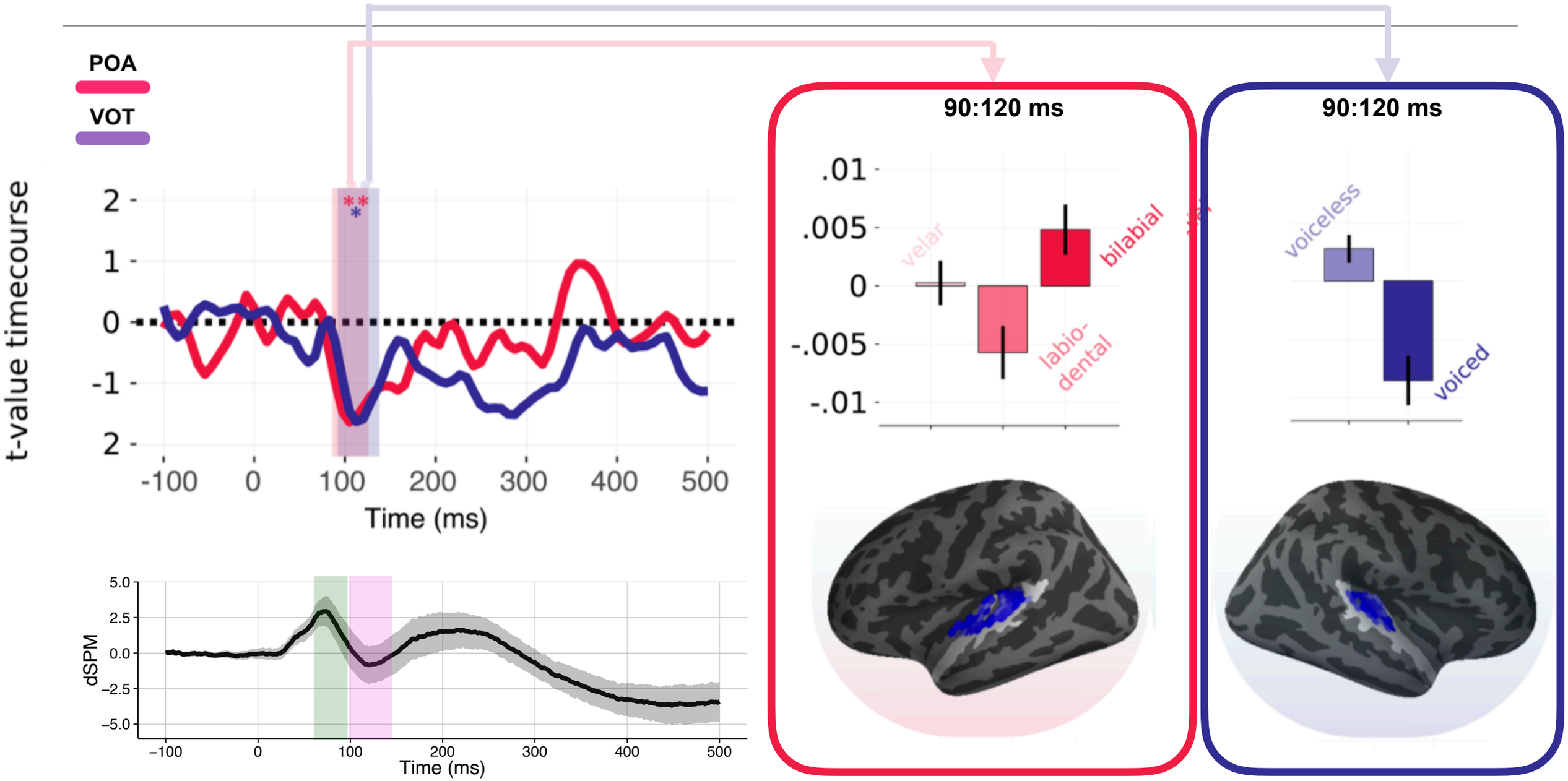
p p b b b

Subphonetics at Onset

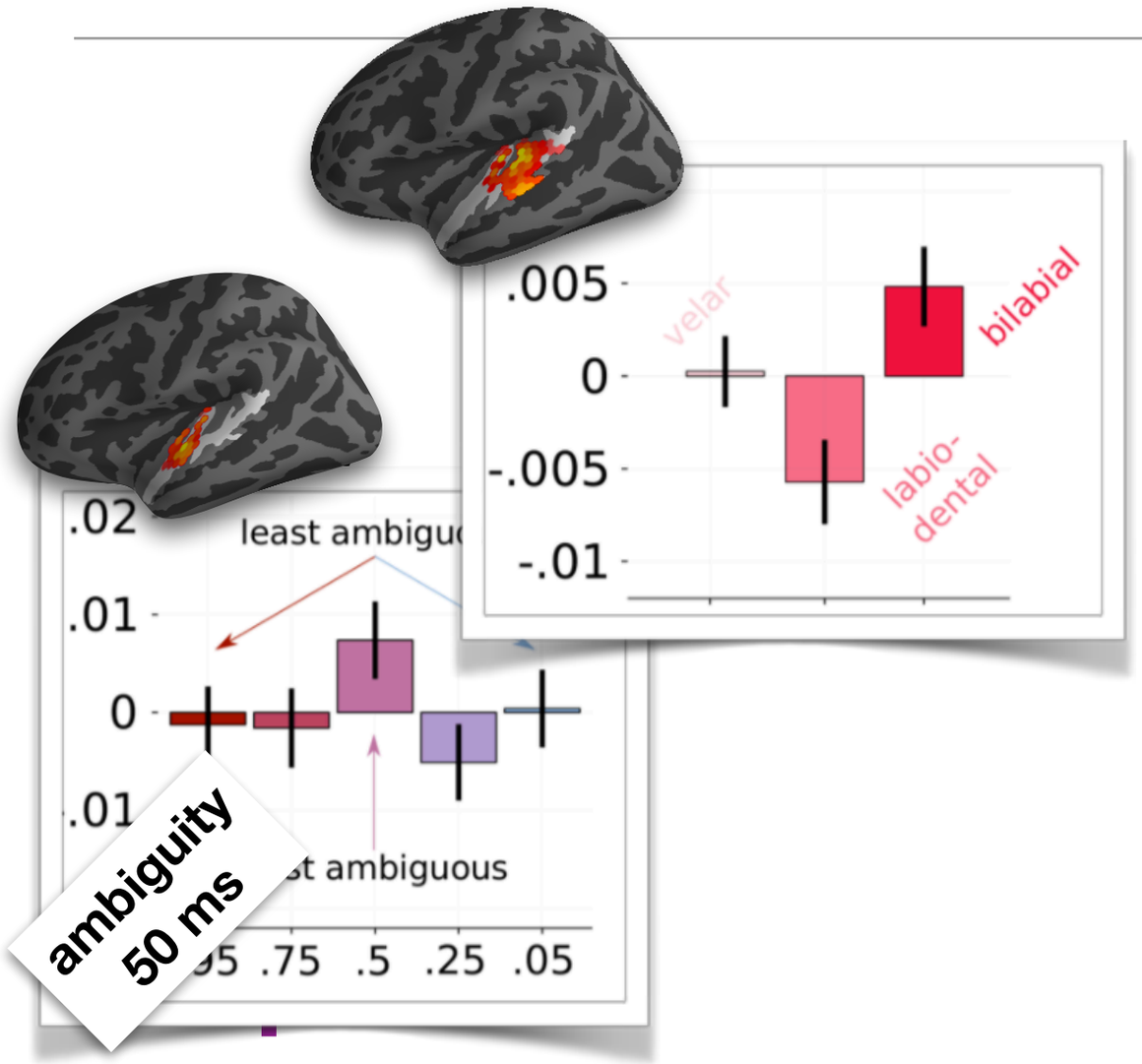


p p p b b

Phonetic Features at Onset

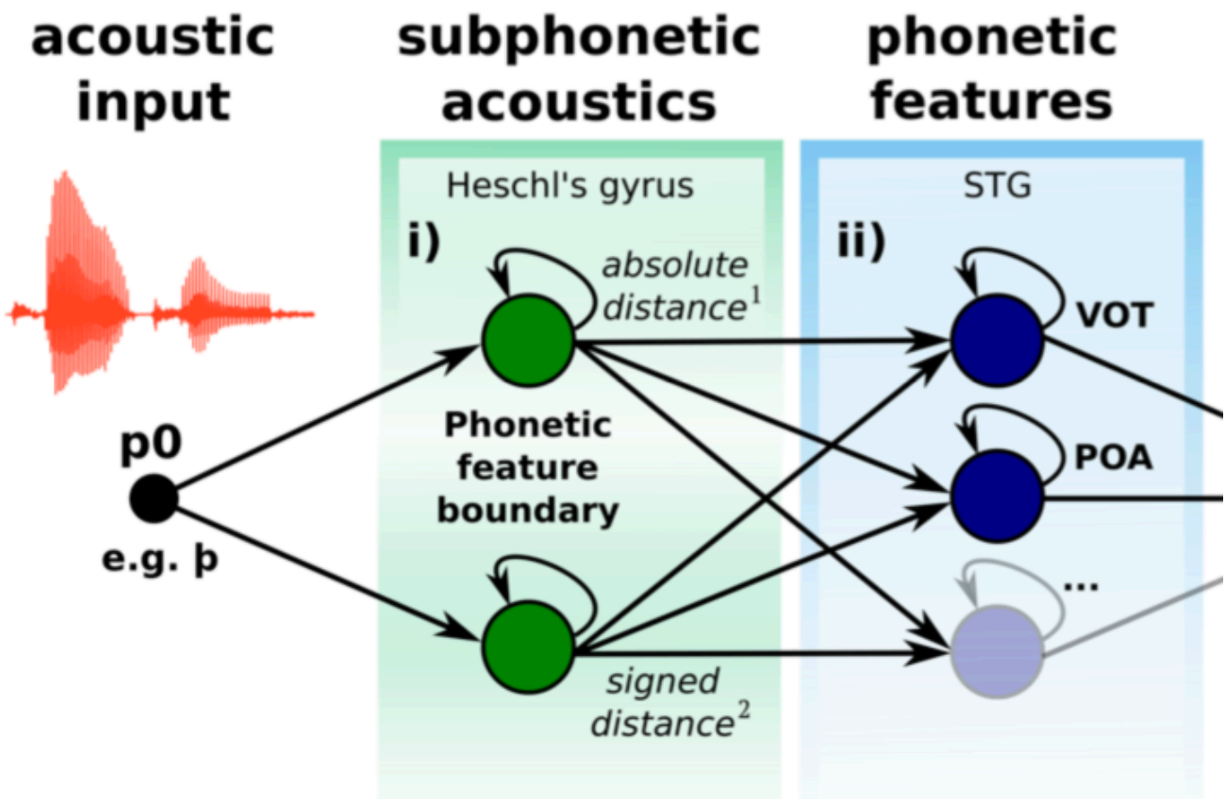


Interim Conclusion



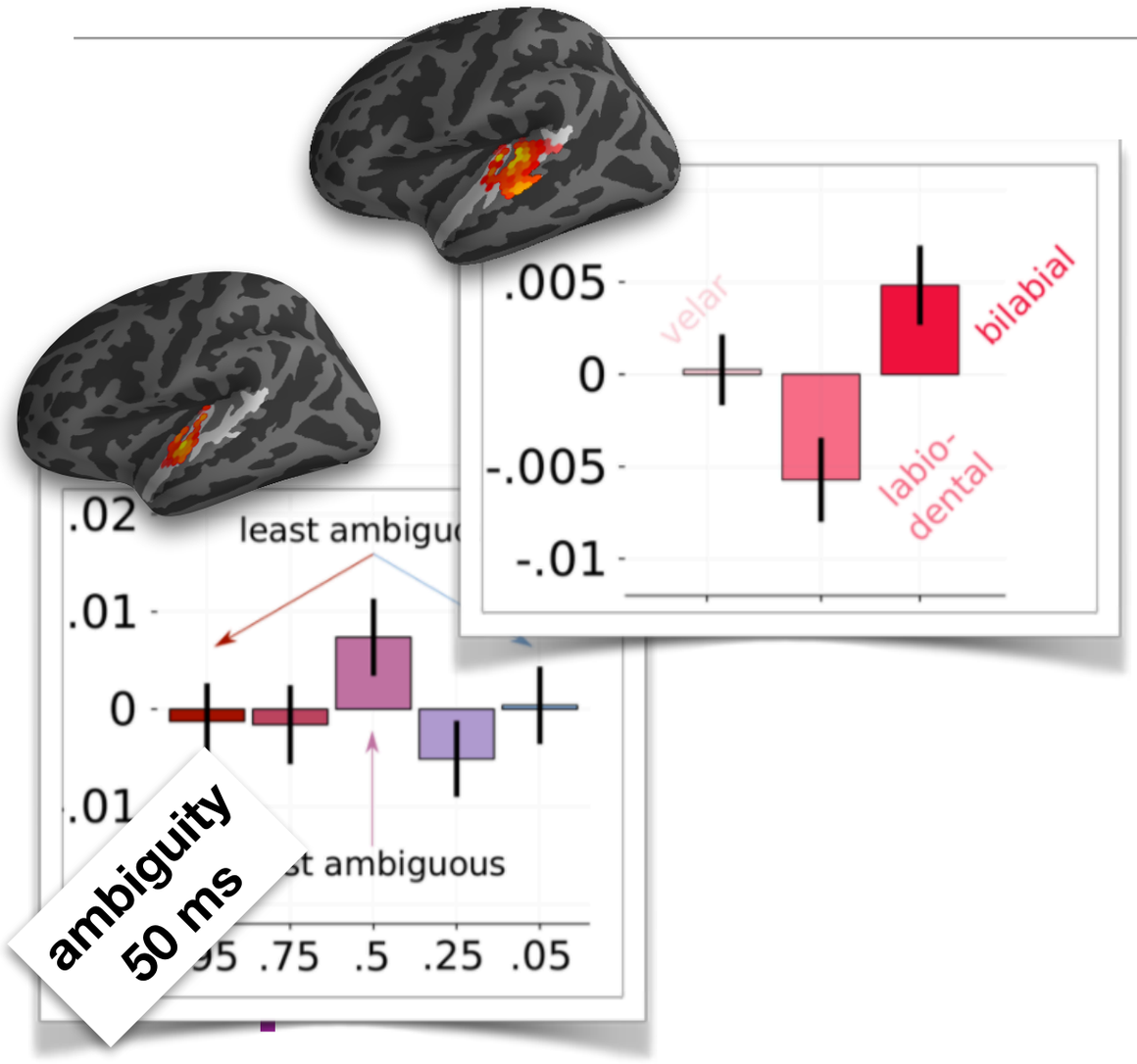
þ a r a k **ee** t

Putting together the processing pieces



Gwilliams et al. (2018) *JNeuro*

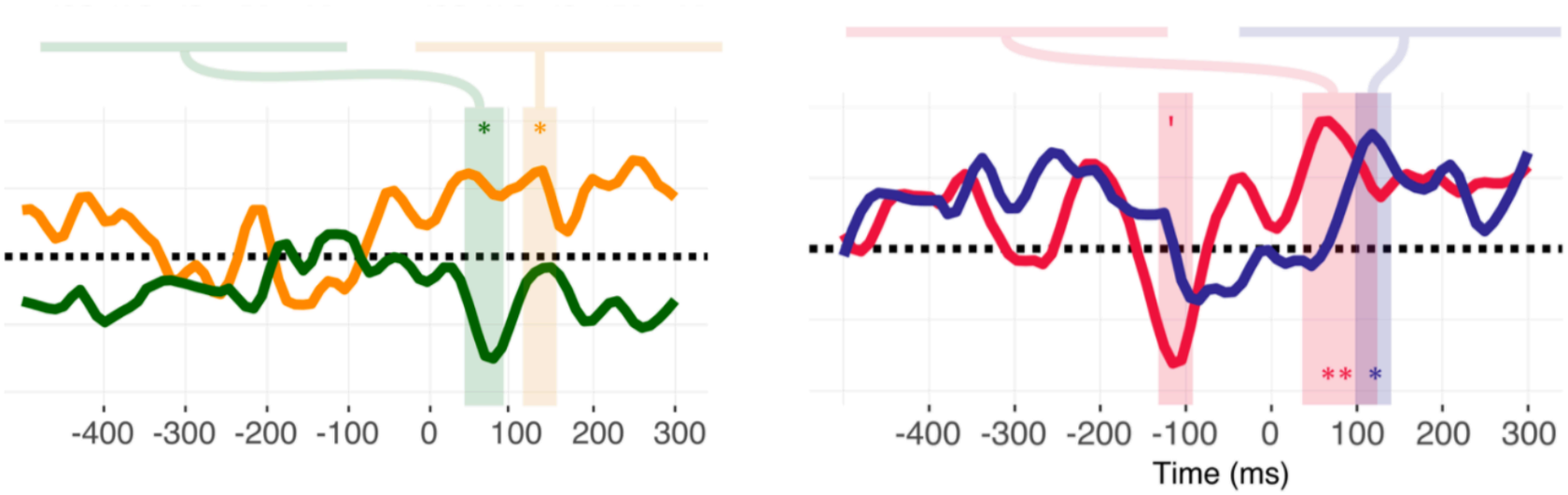
Interim Conclusion



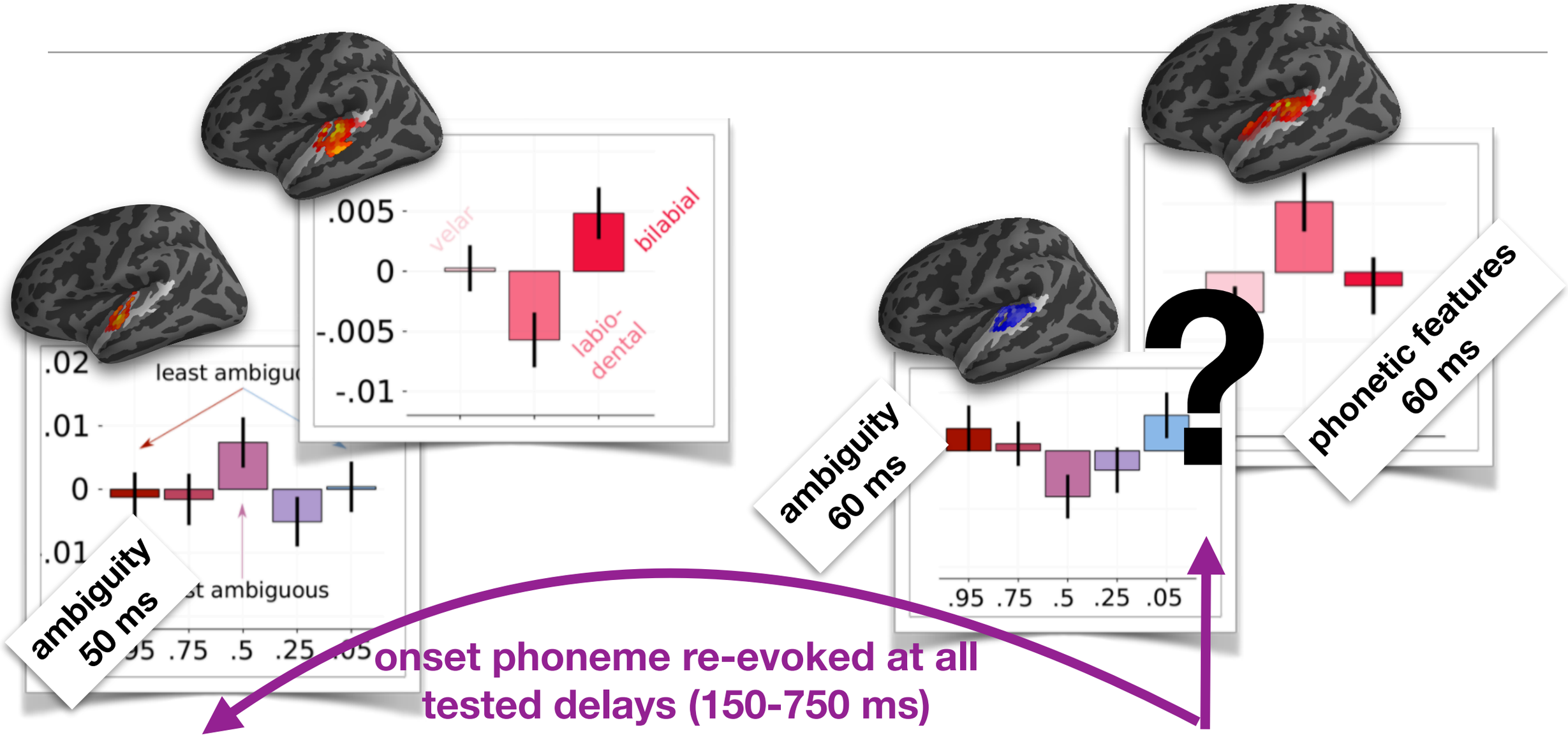
?

þ a r a k **ee** t

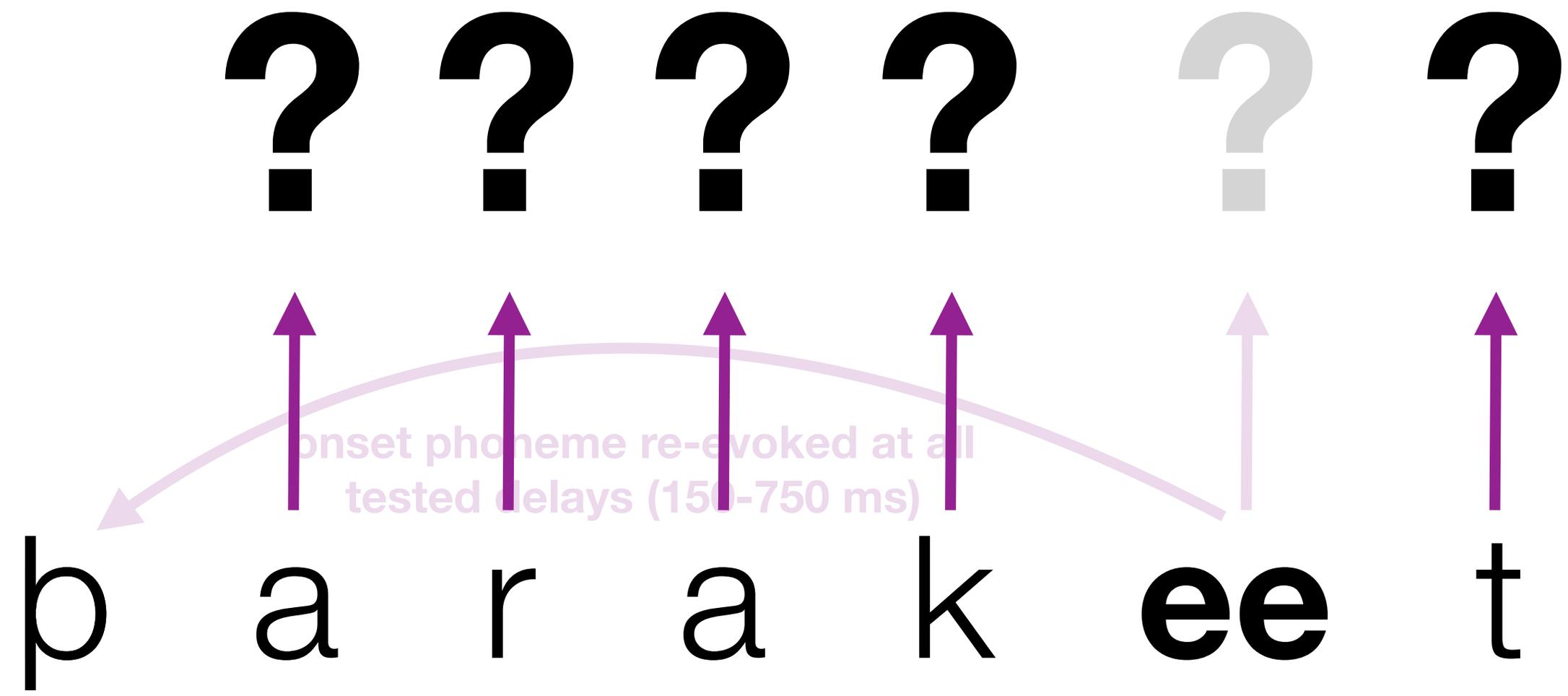
Ambiguity at POD



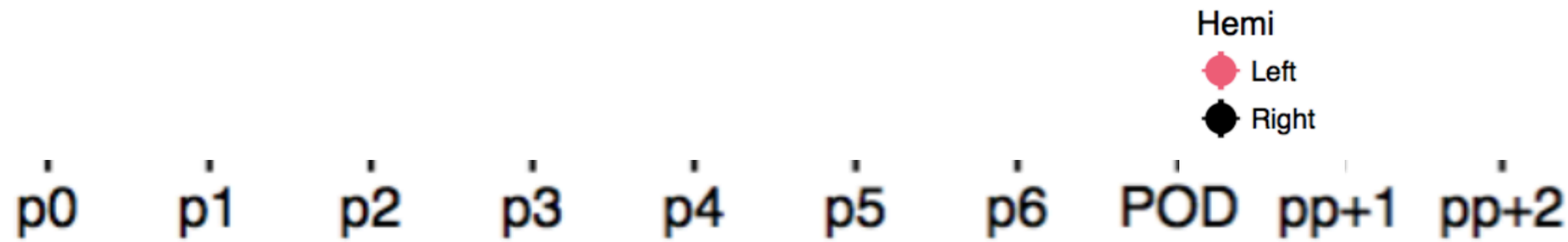
Interim Conclusion



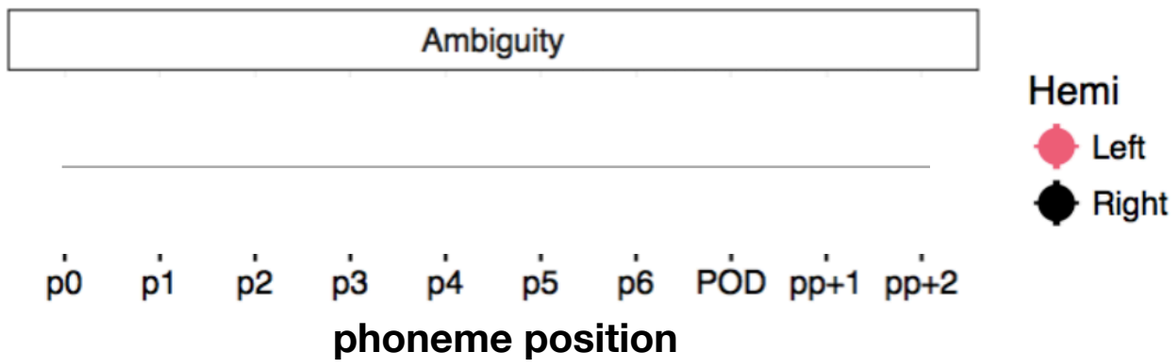
Interim Conclusion



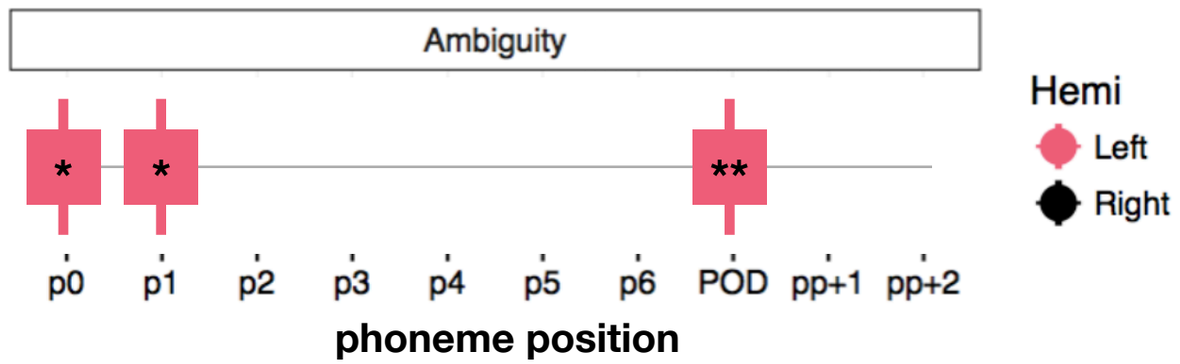
Reactivation in Intermediate Positions



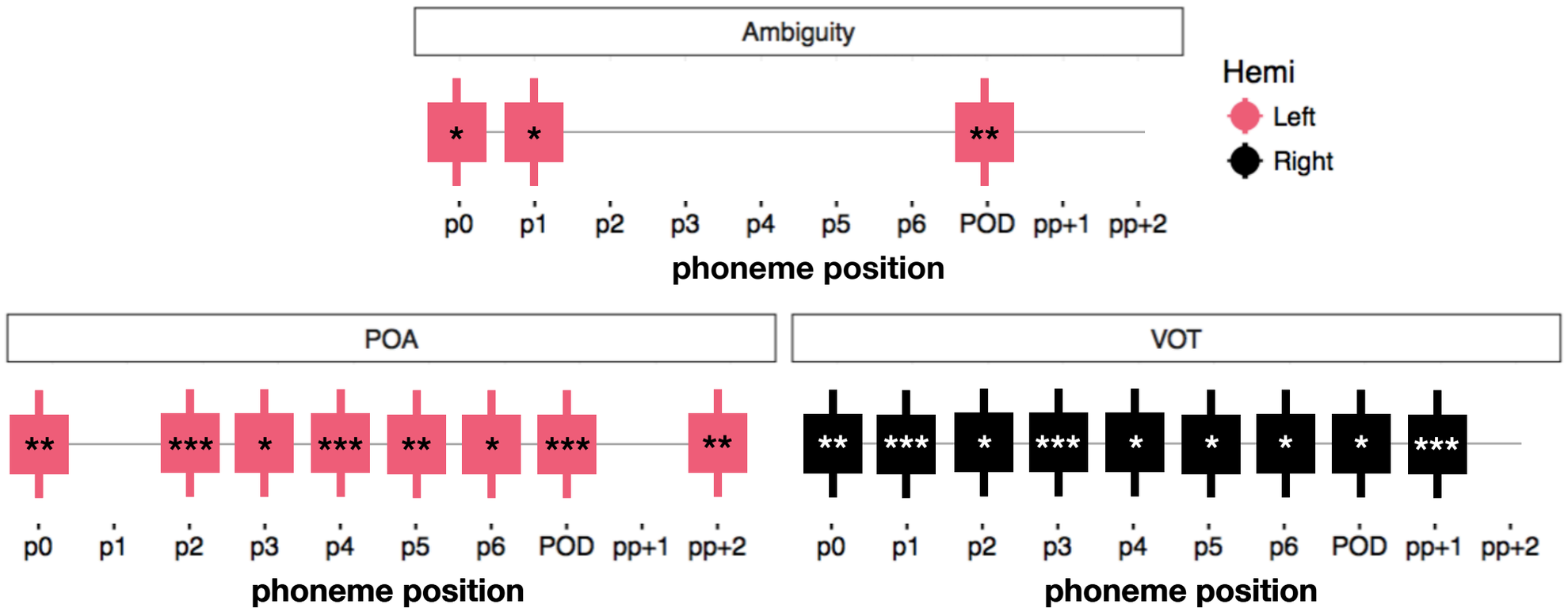
Reactivation in Intermediate Positions



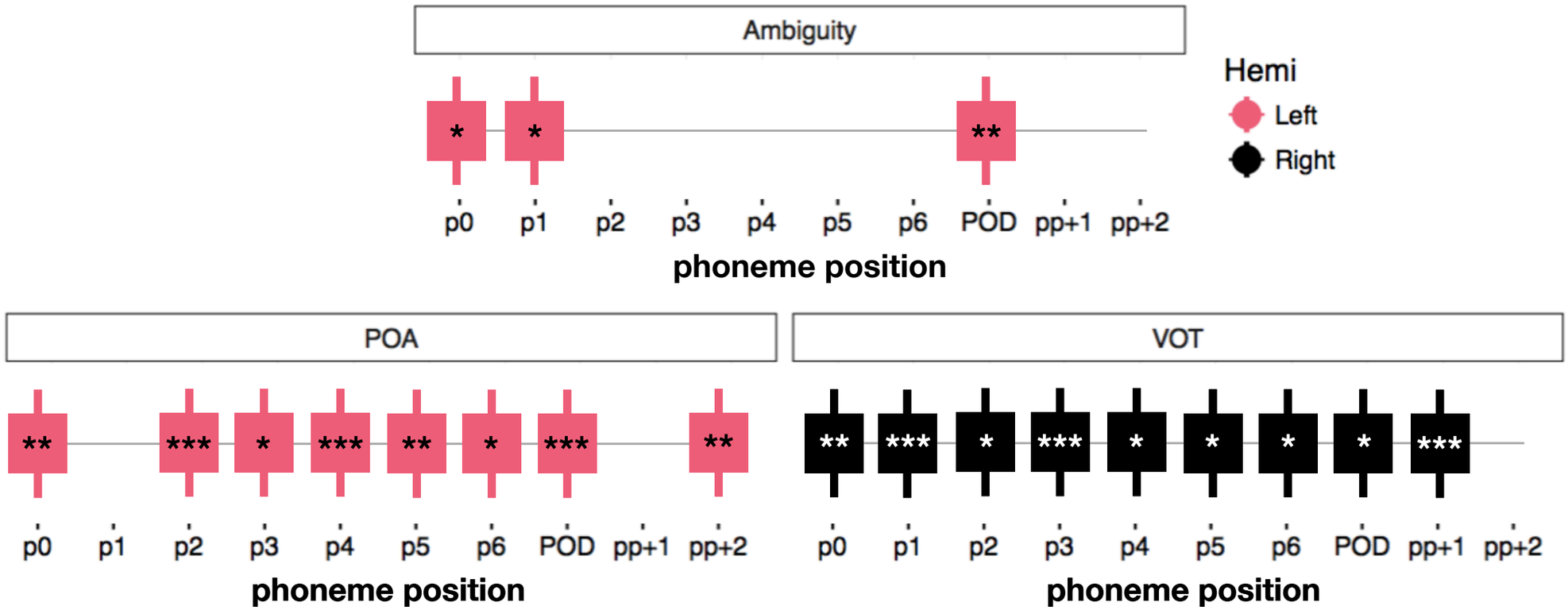
Reactivation in Intermediate Positions



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

/p/

phonological commitment

p a r a k ee t

acoustic-phonetic maintenance

ballet

prove

pin

bath

pacify

bond

palate

beef

book

p

pants

balance

bind

paddle

boast

poke

panda

ballet

prove

pin

bath

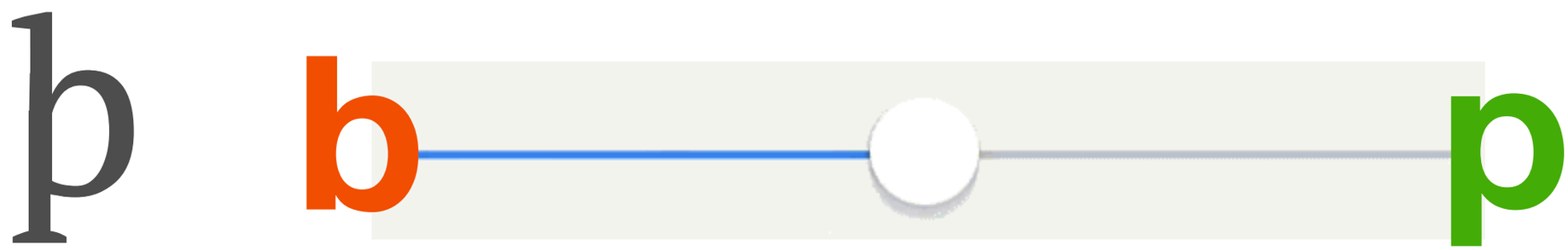
pacify

bond

palate

beef

book



pants

balance

bind

paddle

boast

poke

panda

ballet

prove

pin

bath

pacify

bond

palate

beef

book

b

pants

balance

bind

paddle

boast

poke

panda

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

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

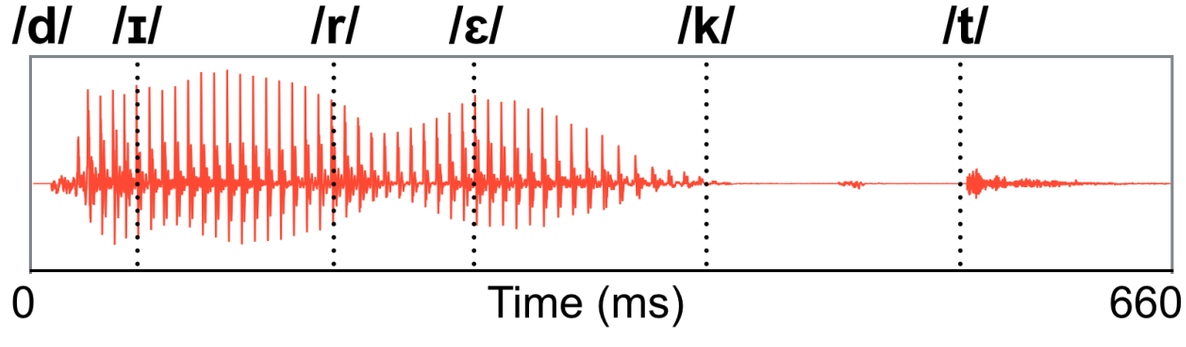
Model Setup

- **Critical variables**:

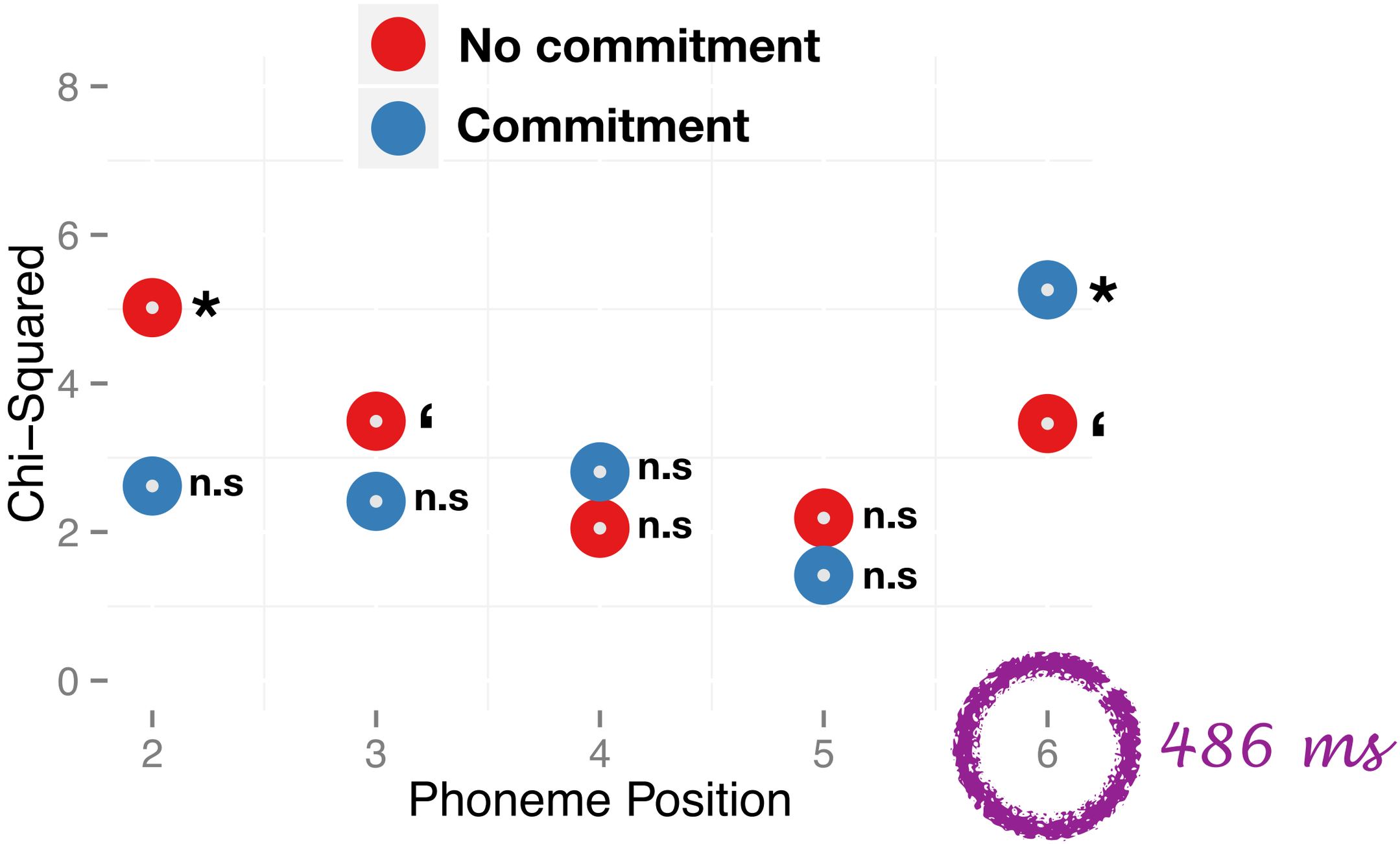
no commitment entropy
no commitment surprisal
commitment entropy
commitment surprisal

- **Control variables**:

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

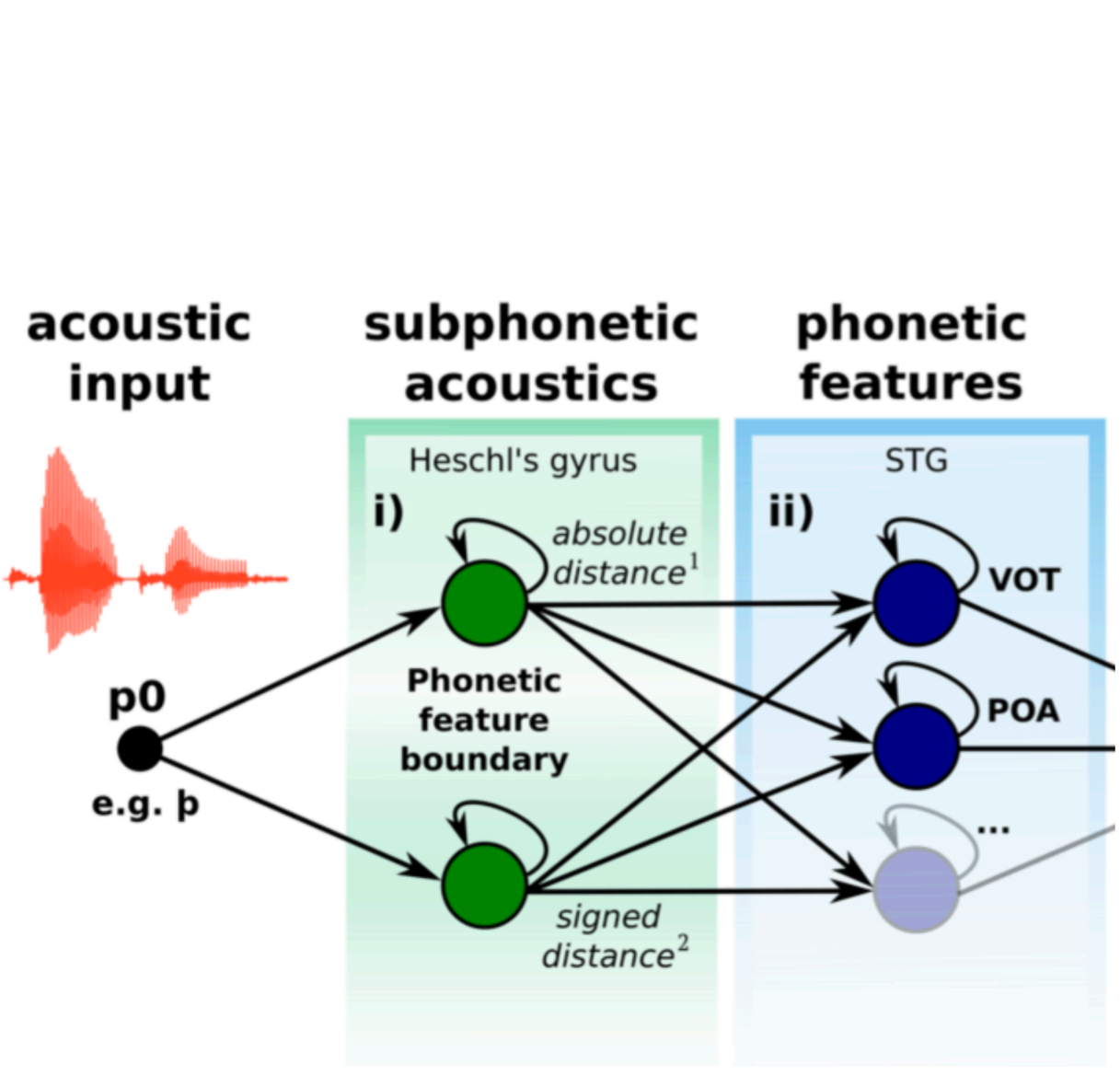


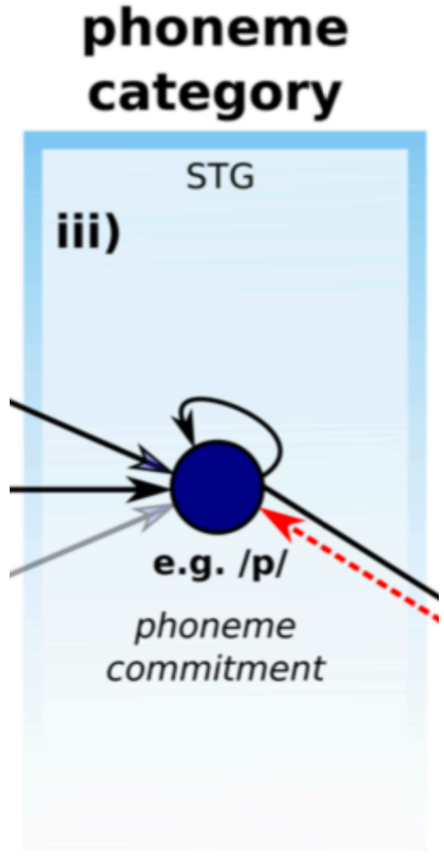
Results



Gwilliams et al., *CMCL*, 2018

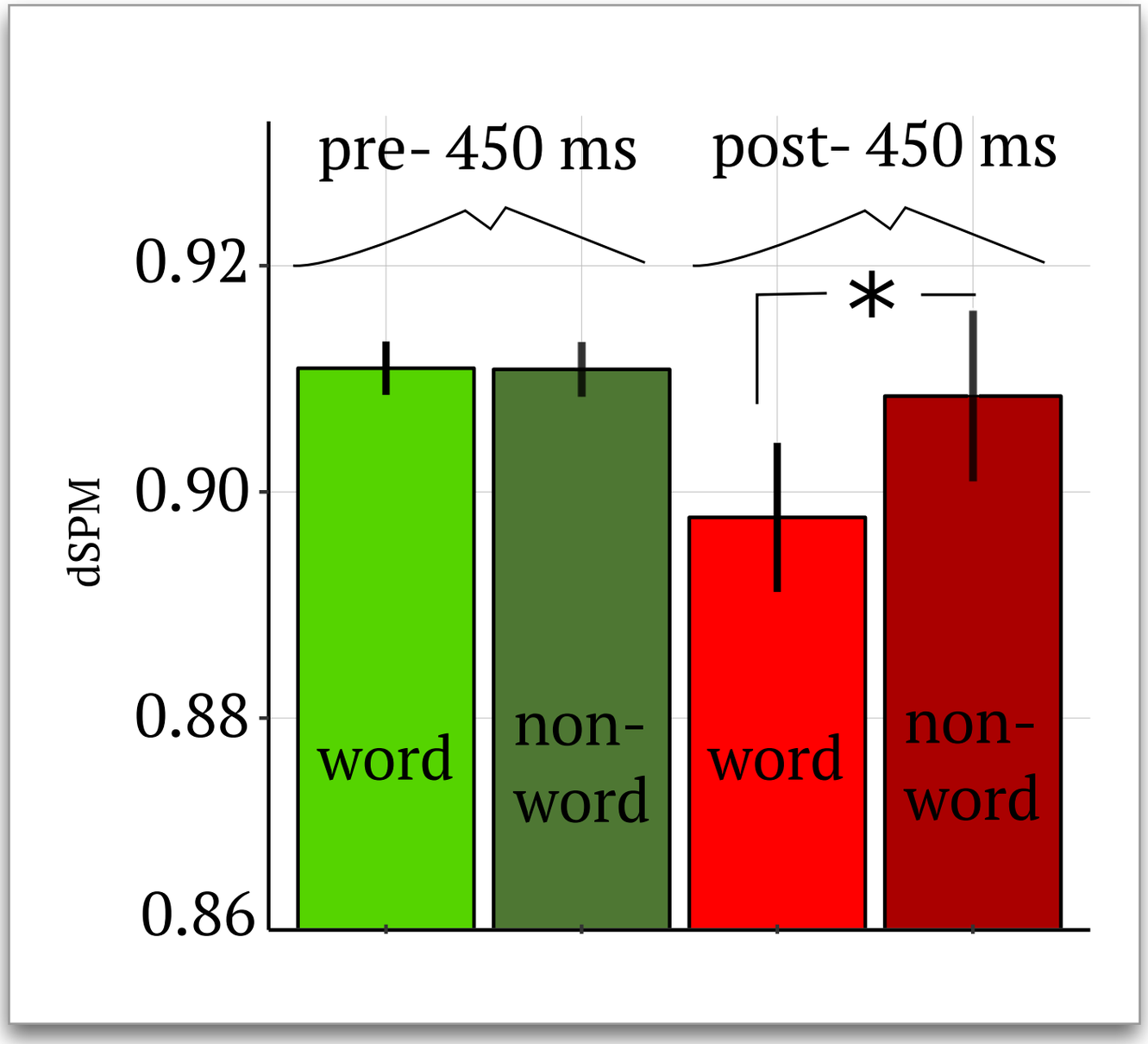
Putting together the processing pieces





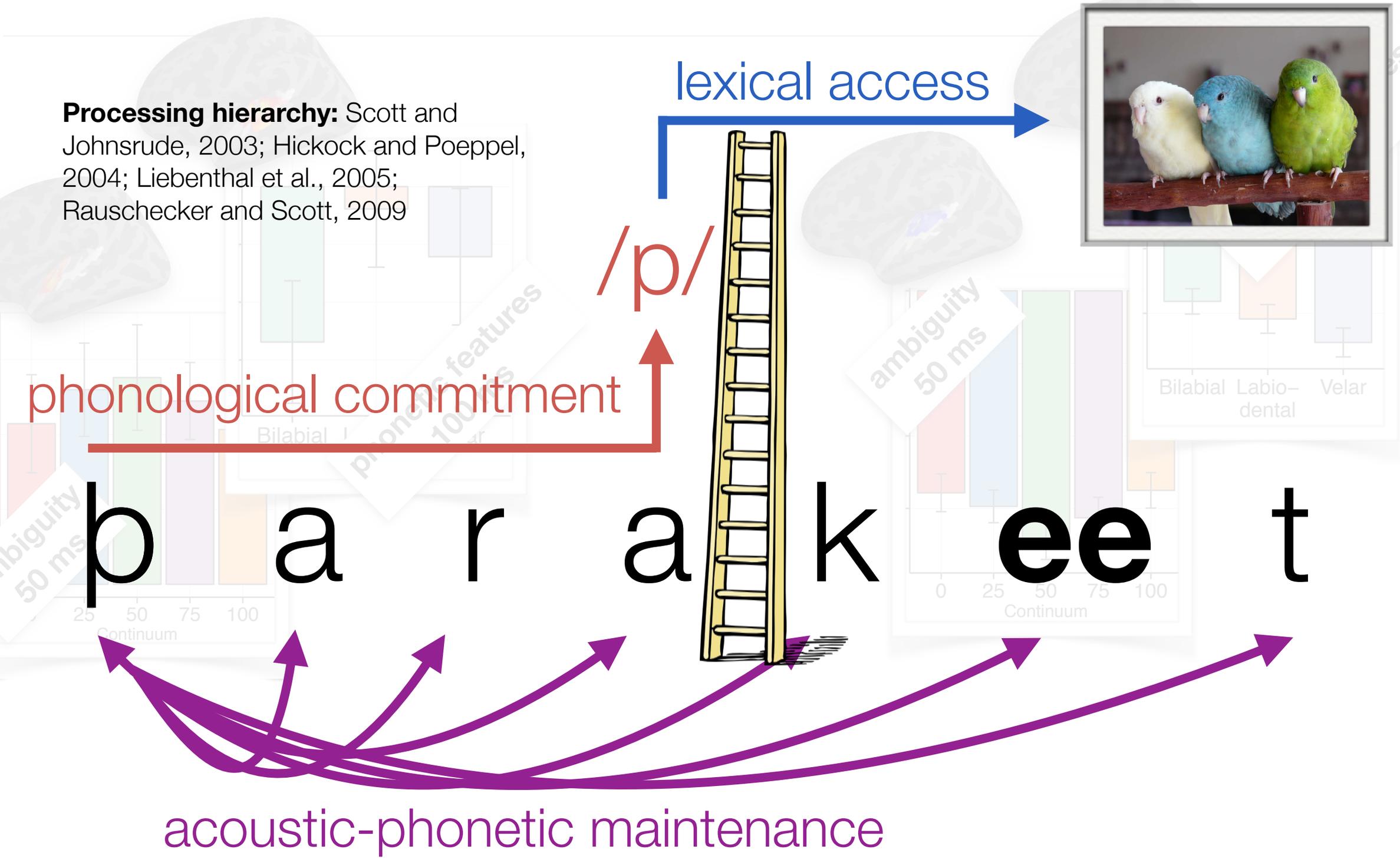
Gwilliams et al. (2018) *JNeuro*

Further test of commitment



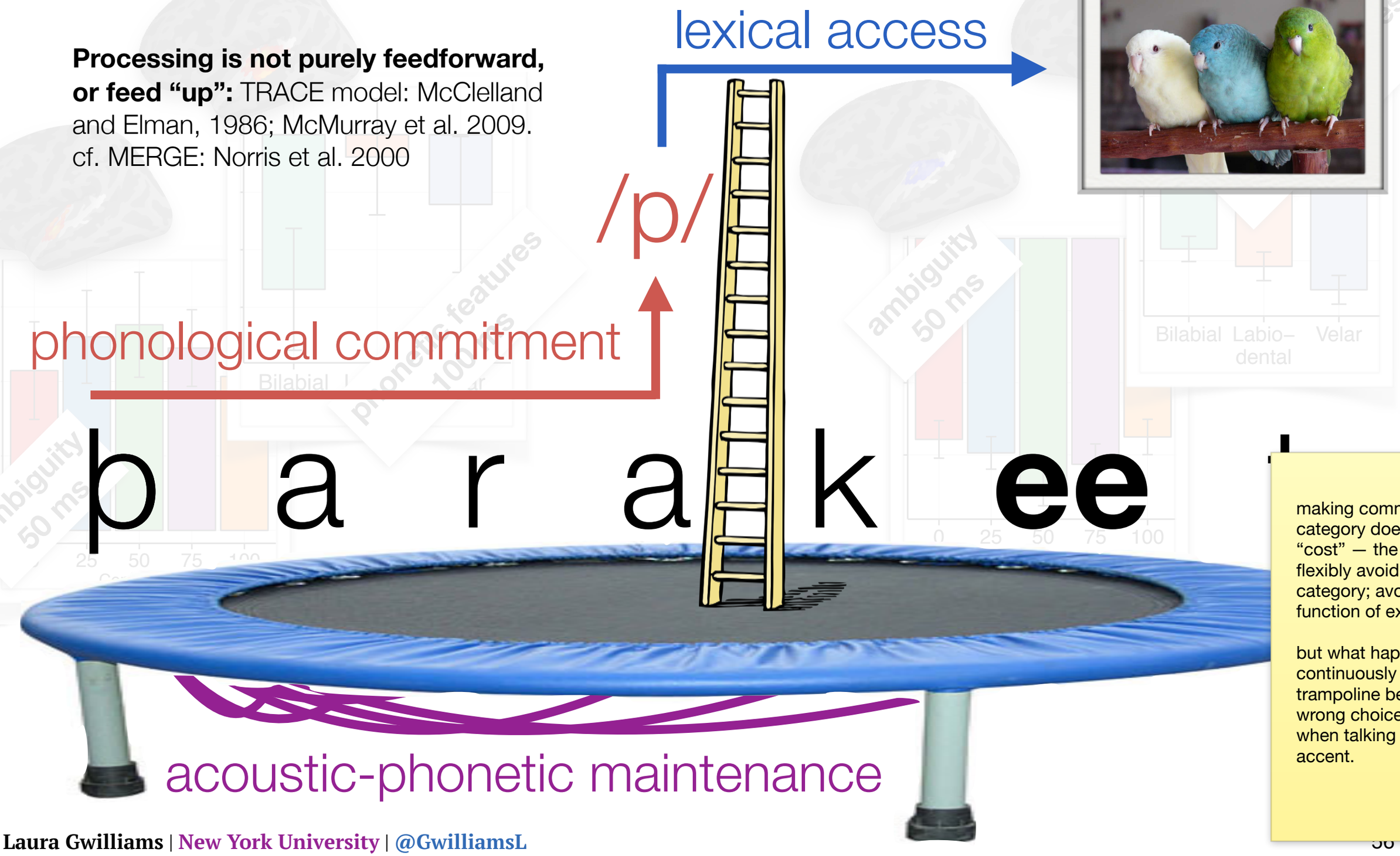
Interpretation

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



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 commitmen
category does not c
"cost" — the syster
flexibly avoid comm
category; avoiding t
function of exposur

but what happens i
continuously jumpir
trampoline because
wrong choice? well
when talking to som
accent.

acoustic input

subphonetic acoustics

phonetic features

phoneme category

phoneme sequence

lexical candidates

Heschl's gyrus

i)

absolute distance[1]

Phonetic feature boundary

signed distance[2]

p0

e.g. þ

STG

ii)

VOT

POA

...

STG

iii)

e.g. /p/

phoneme commitment

?

iv)

/ˈp/

/ˈpa/

/ˈparəkiːt/

Left MTG

v)

barricade

plant

bread

parakeet

ballot

possum

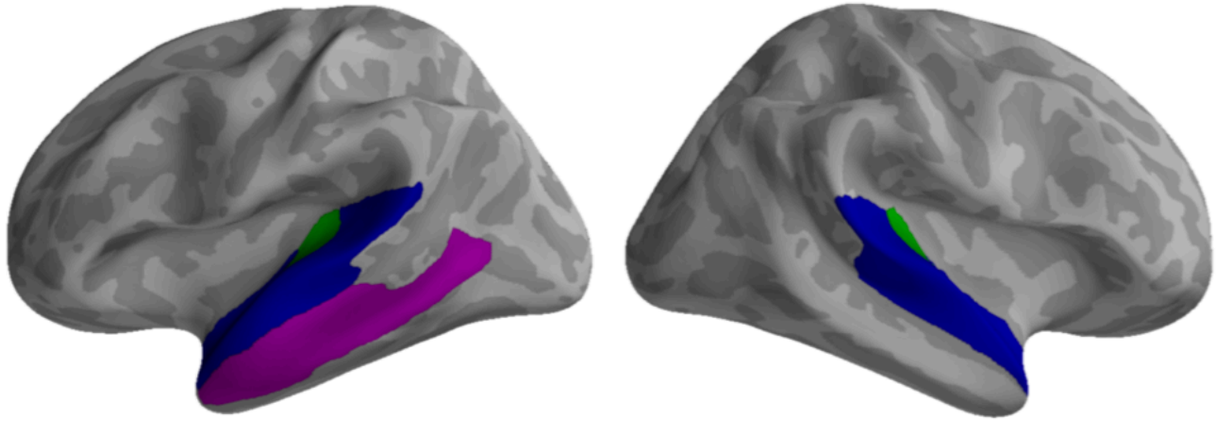
p1

e.g. a

e.g. /a/

1. Co-linear with phoneme ambiguity
2. Co-linear with perceptual continuum

phonemes



connection types

recurrent

feedforward

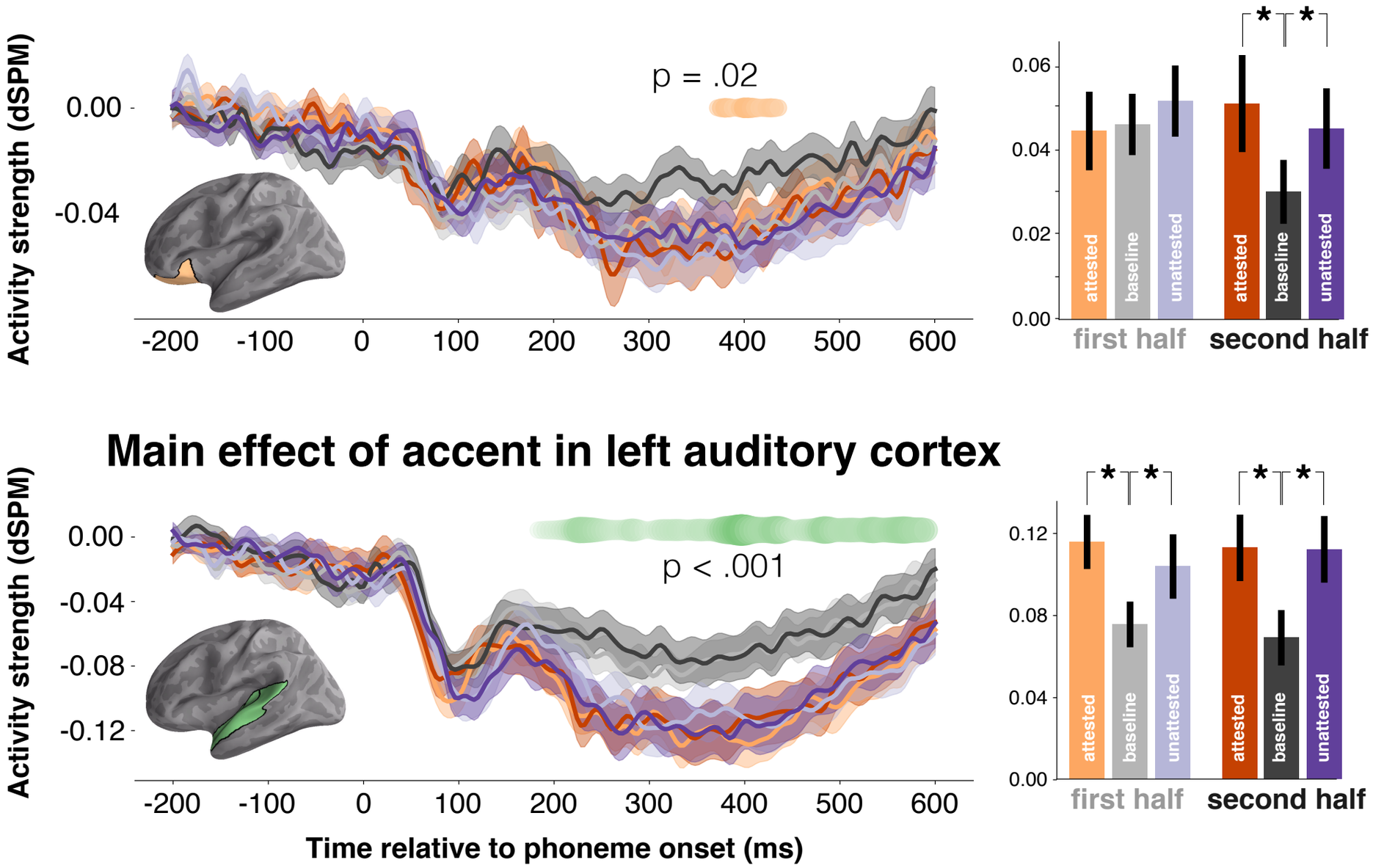
feedback

Laura Gwilliams | New York University | @Gwill

What's next?

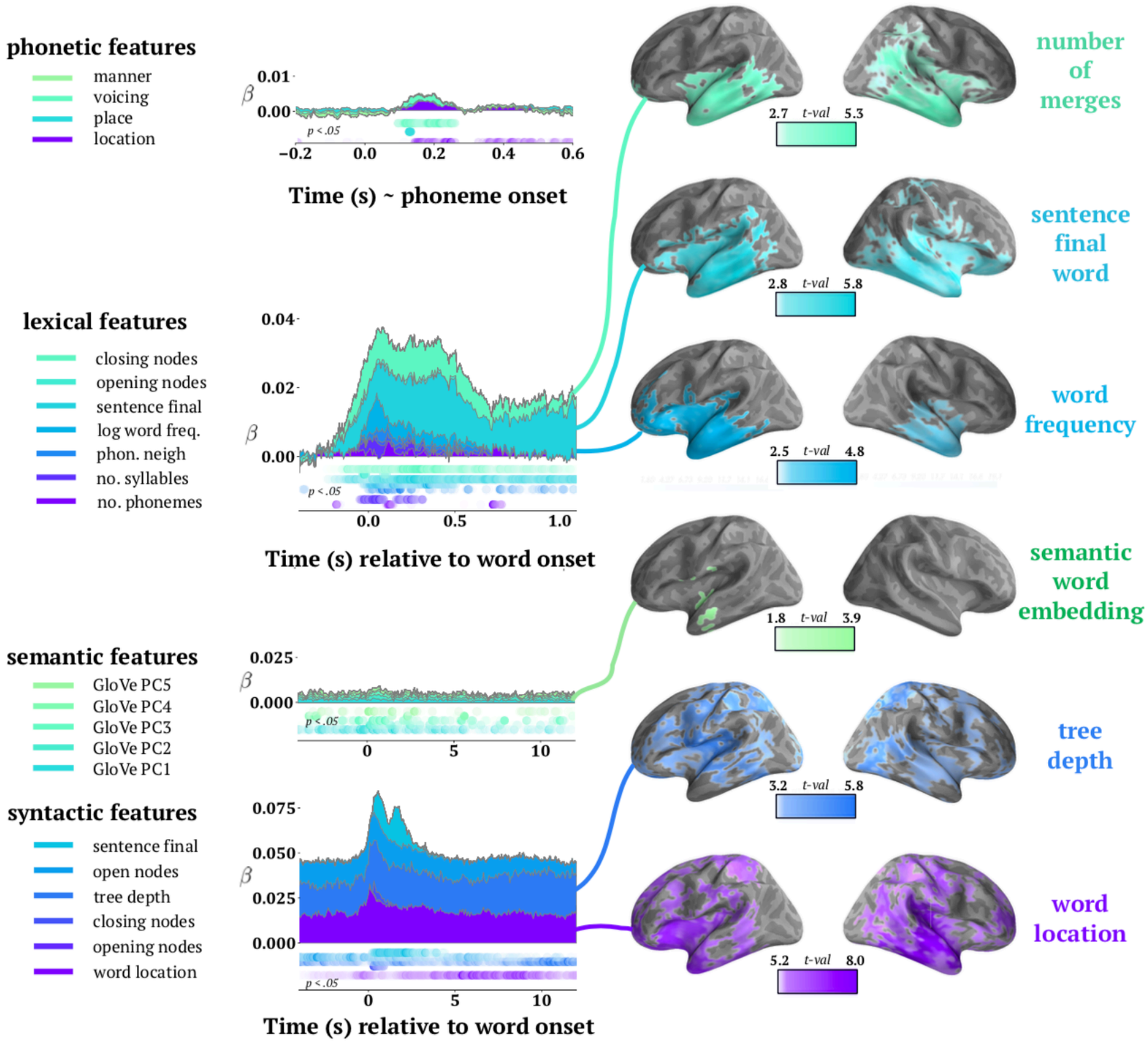
- **Commitment process**
  - **Question**: is the time-course of commitment temporally or informationally defined?
  - Data collection phase: Sandy Abu Adas; Alicia Chatten; Guy Tabachnick; Alec Marantz
- **Phoneme sequence representations**
  - **Question**: are there phoneme sequence representations in STG? Re-activation?
  - Data collection phase: Narayan Sankaran; Matthew Leonard; Edward Chang
- **Accent attunement**
  - **Question**: does lexical feedback serve to perceptually re-map acoustics to phonetics?
  - Analysis phase: Esti Blanco-Elorrieta; Liina Pylkkanen; Alec Marantz
- **Hierarchy in continuous speech**
  - **Question**: does a similar architecture support higher-level comprehension?
  - Analysis phase: Jean-Remi King; David Poeppel
- **Domain general auditory perception: pitch**
  - **Question**: how is a percept of pitch formed from an ambiguous input?
  - Analysis phase: Ellie Abrams; Alec Marantz

Accent attunement



Laura Gwilliams | New York University | @GwilliamsL

Hierarchy in continuous speech



Laura Gwilliams | New York University | @GwilliamsL

✉ 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

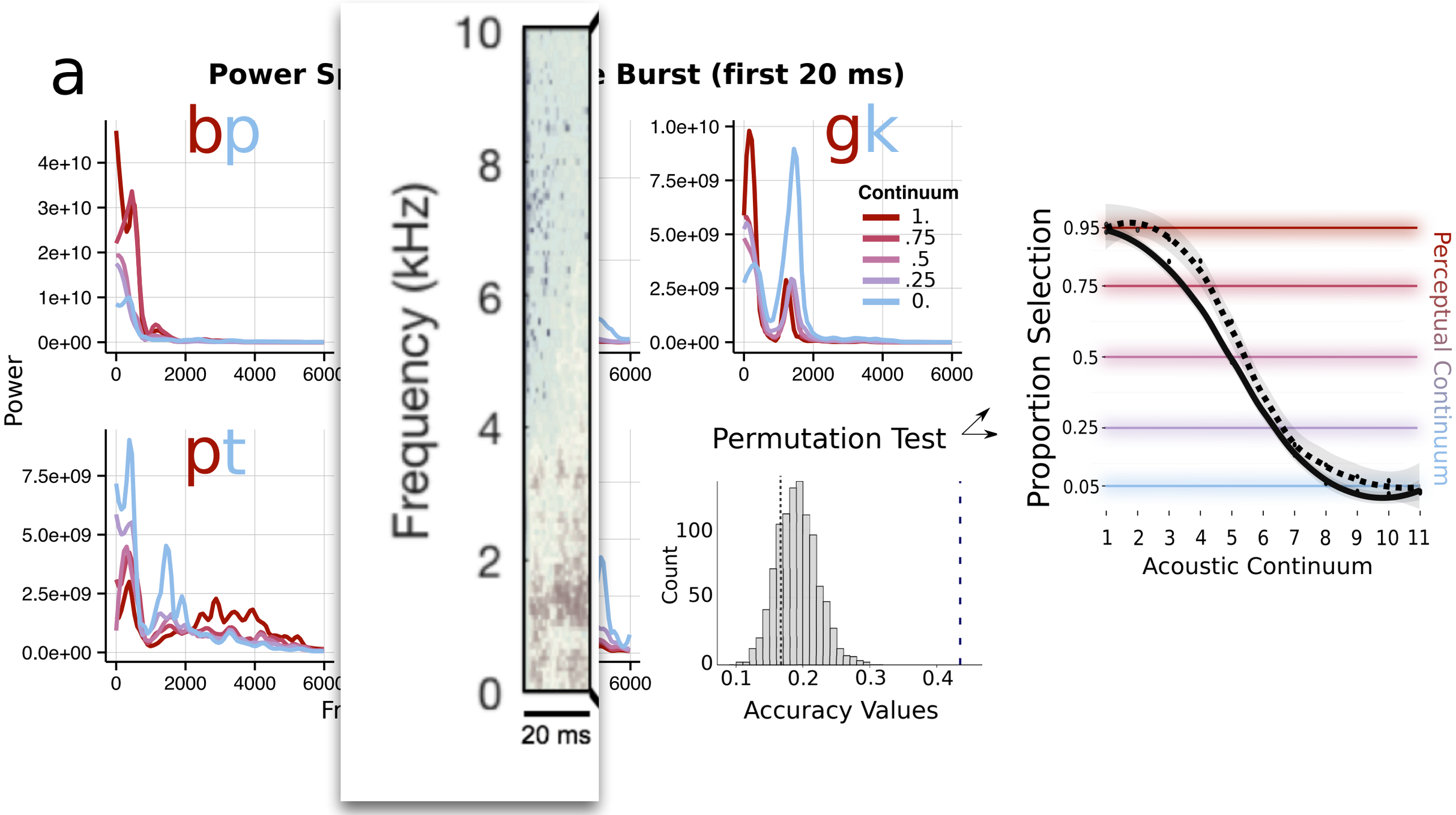
NEW YORK UNIVERSITY

laura.gwilliams@nyu.edu
@GwilliamsL

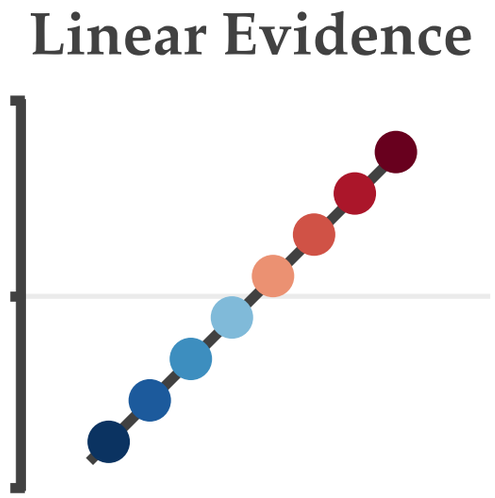
Thank you!

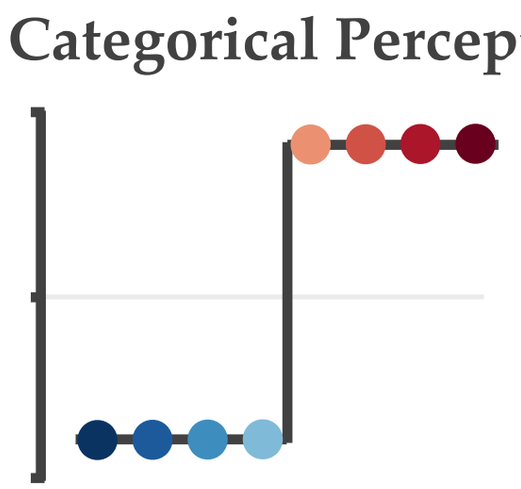
Laura Gwilliams | **New York University** | **@GwilliamsL**

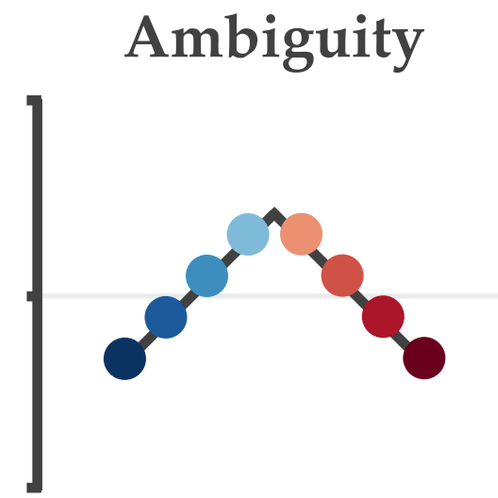
Is ambiguity correlated with acoustic properties?



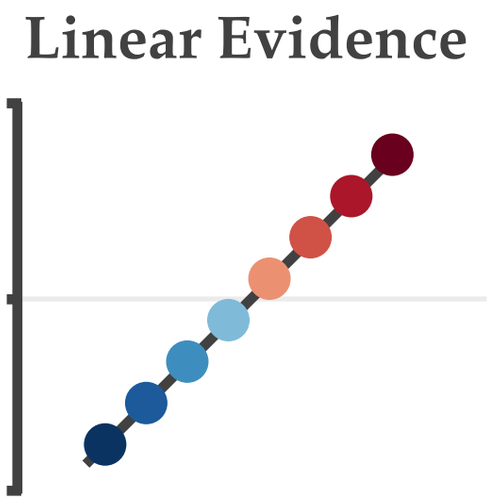
Predictive Coding

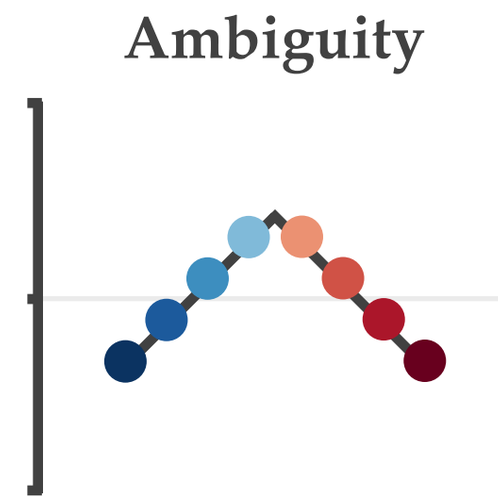


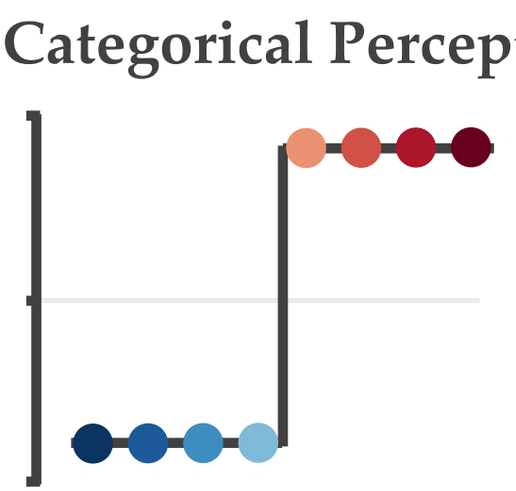




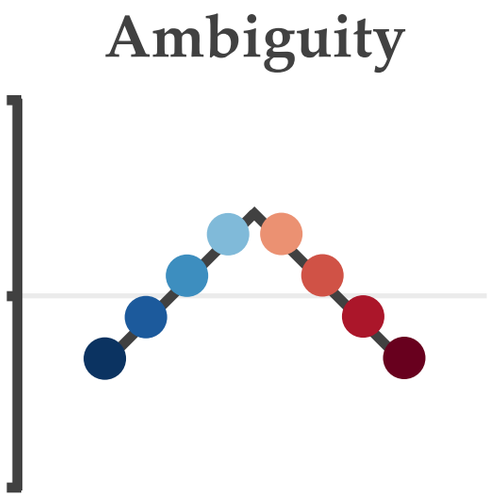
Neutralisation

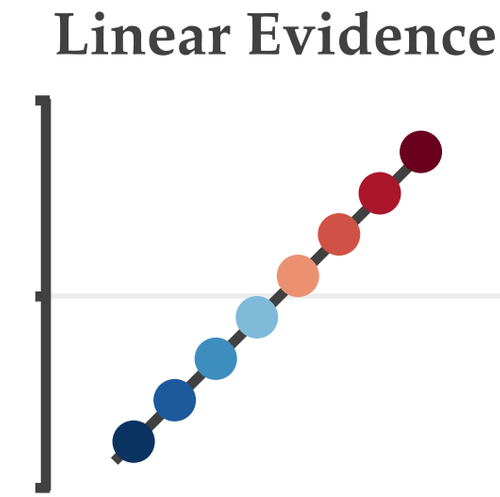


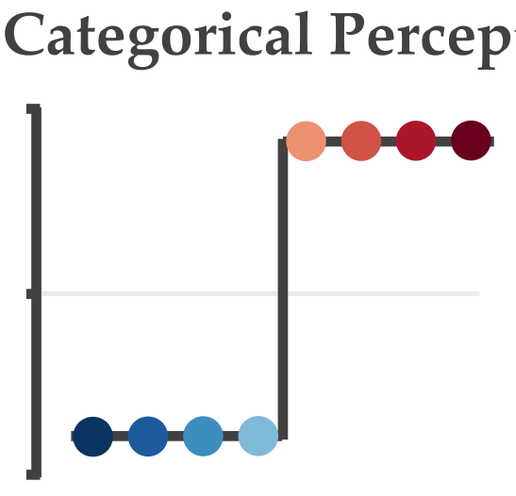




Cut-through connection

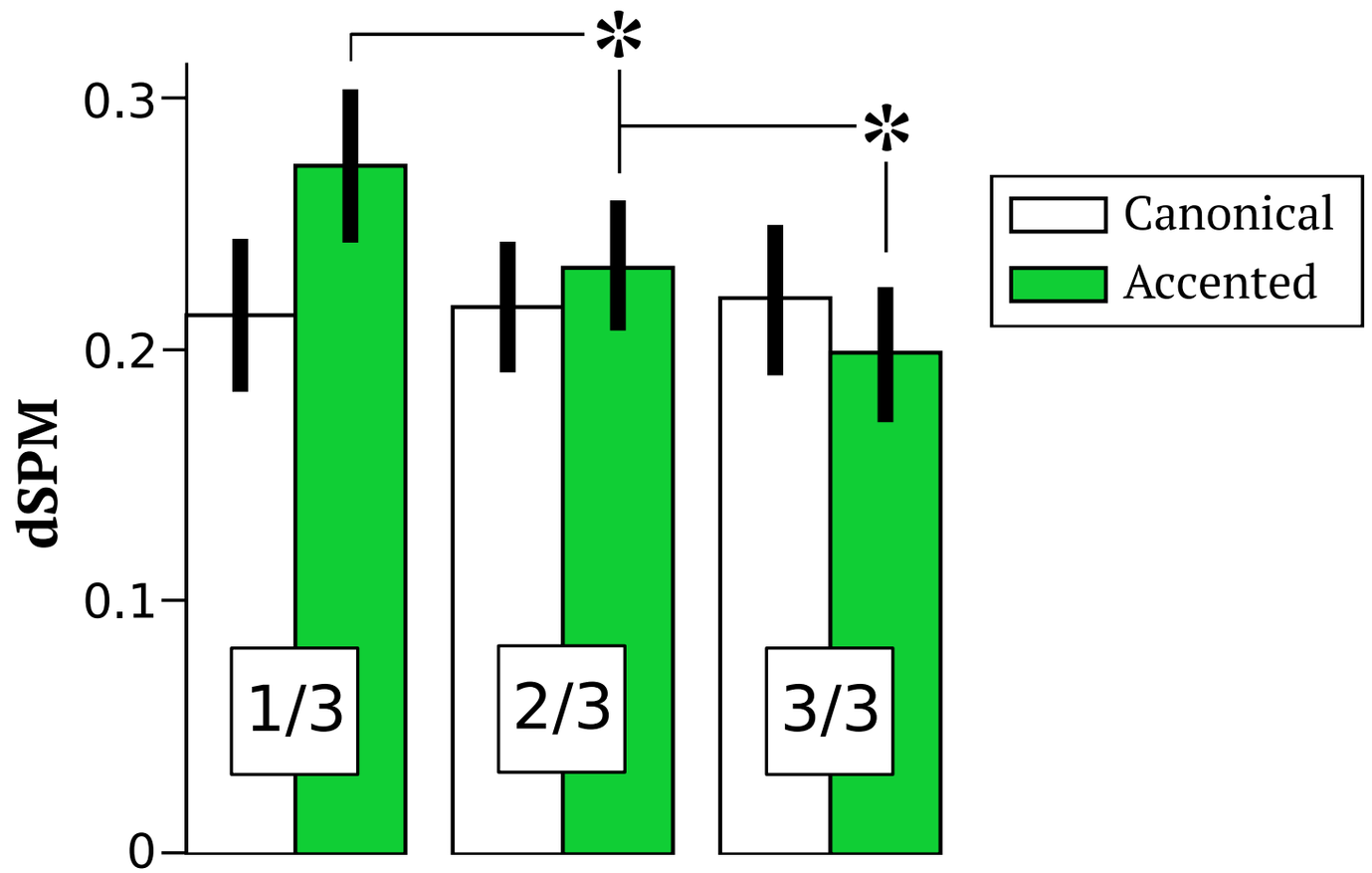




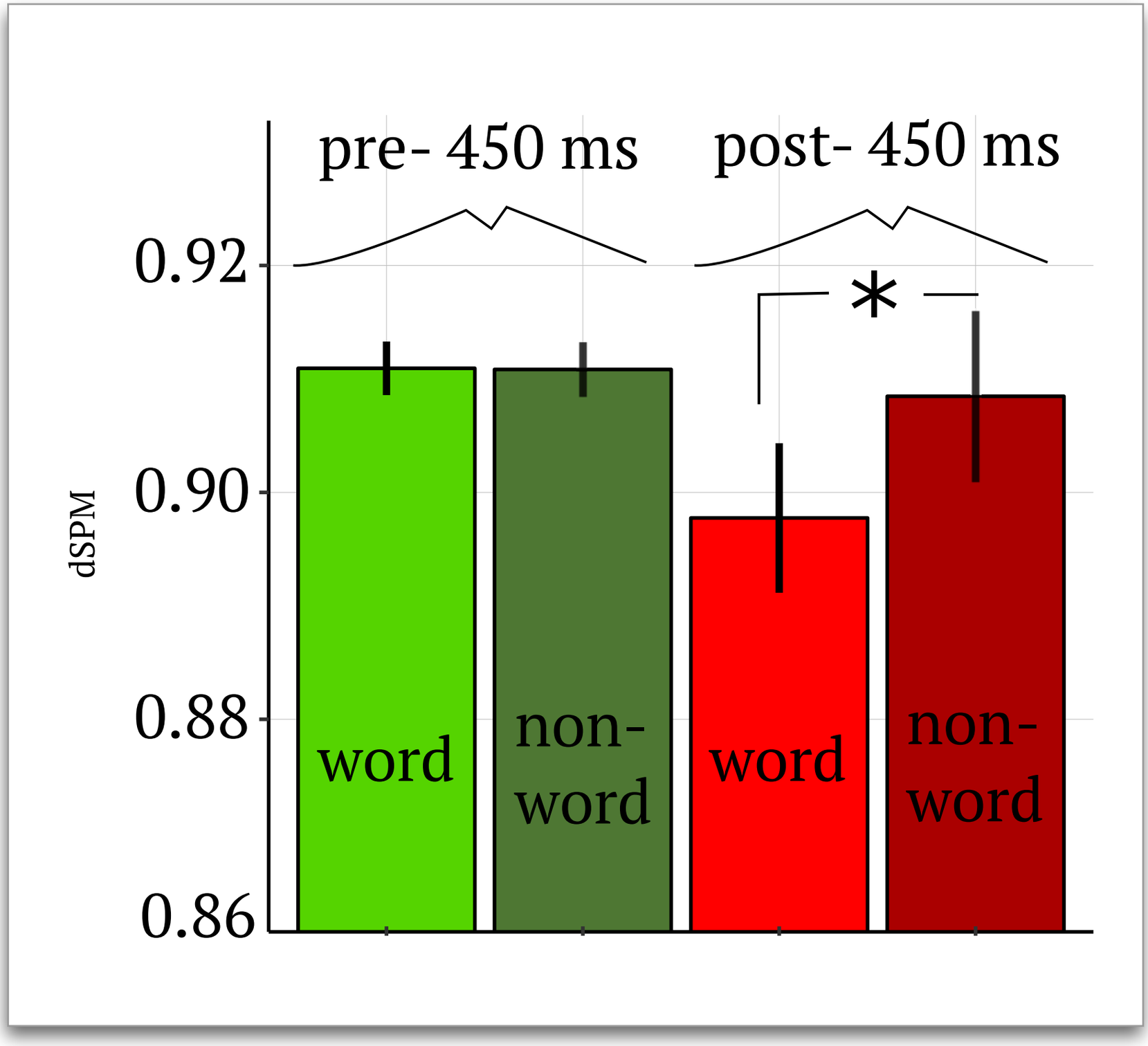


Laura Gwilliams | New York University | @GwilliamsL

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)



Laura Gwilliams | New York University | @GwilliamsL