
Phonetic vs. contextual cues in communication between merged and unmerged speakers

Lacey Arnold Wade
University of Pennsylvania
NWAV45

Background

Herzog's Principle:

Mergers expand at the expense of distinctions

Background

Herzog's Principle:

Mergers expand at the expense of distinctions

But why?

Background

One proposal:

1. One-phoneme speakers come into contact with two-phoneme speakers
2. “The two-phoneme speakers often misunderstand what the one-phoneme speakers have said”
3. “The one-phoneme speakers, on the other hand, do not misunderstand the two-phoneme speakers — not because they can use the phonemic difference, but because they do not attend to it in deciding what is being said. Given what is for them a complete homonymy between *Don* and *Dawn*, they rely on the same types of syntactic, semantic and pragmatic information that they use in their own speech.”
4. “After a certain period of time, the two-phoneme speakers cease to attend to this phonemic distinction themselves, since it is not reliable”

Labov (1994) paraphrasing Herold (1990)

Background

Can you get me
a coffee?



merged speaker

[kafi]
/ \
[kapi] [kɔfi]



unmerged speaker

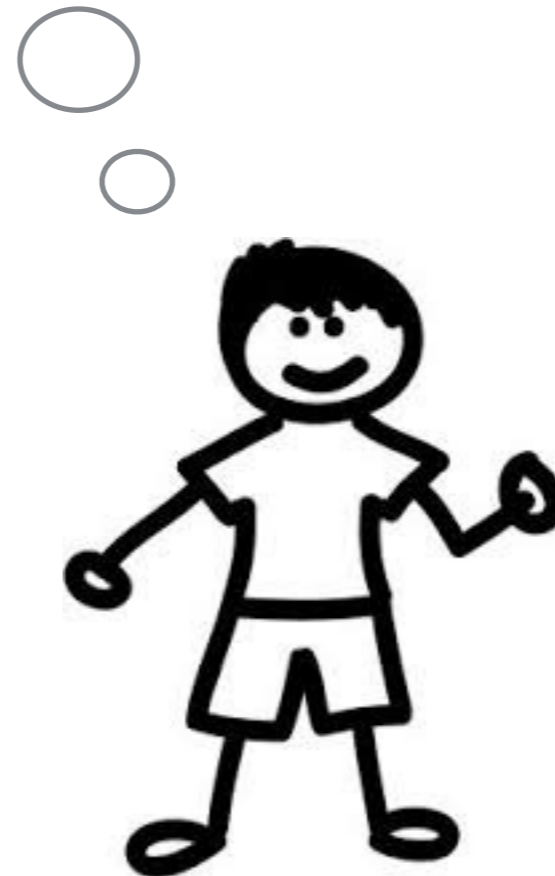
Background

Can you get me
a coffee?



merged speaker

[kafi]
/ \
[kapi] [kɔfi]



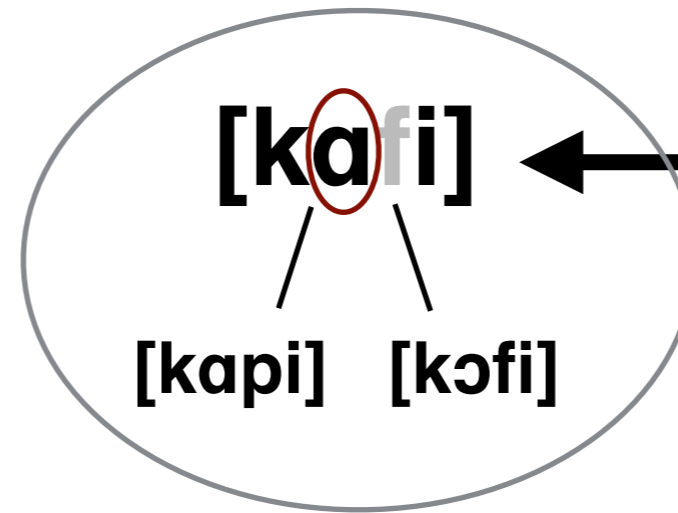
unmerged speaker

Background

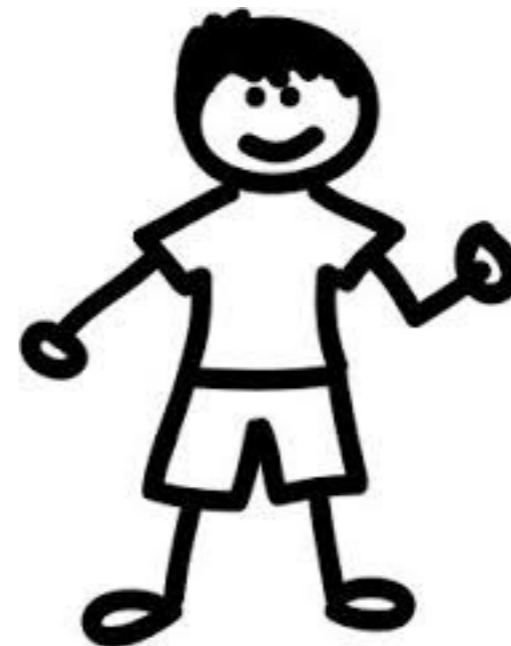
Can you get me
a coffee?



merged speaker



← **phonetic**



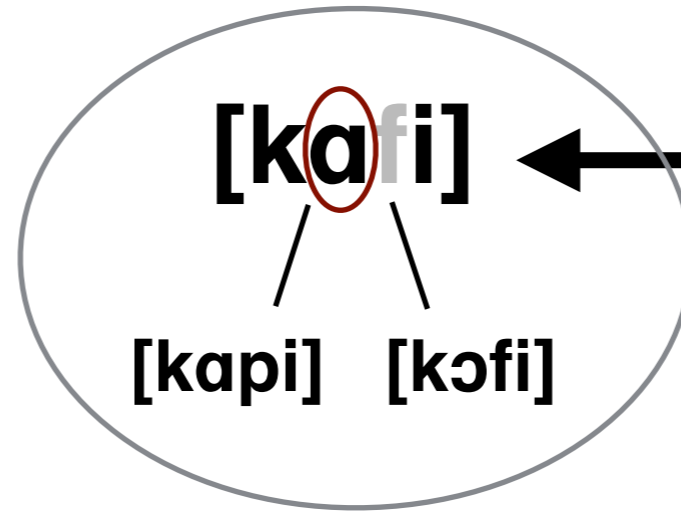
unmerged speaker

Background

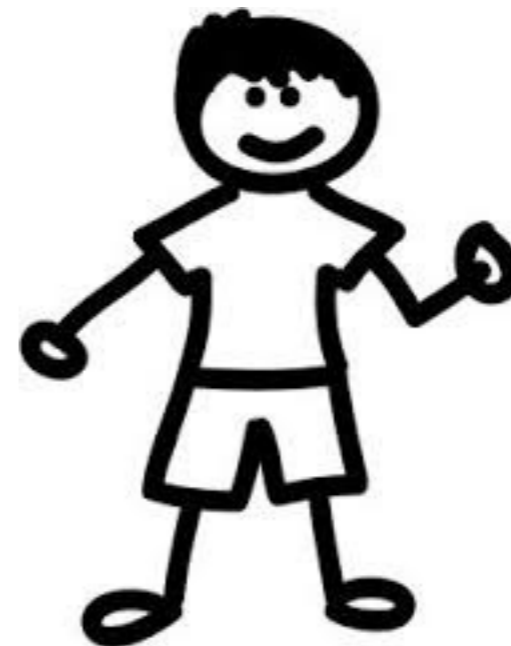
Can you get me
a coffee?



merged speaker



← **phonetic**



unmerged speaker

contextual

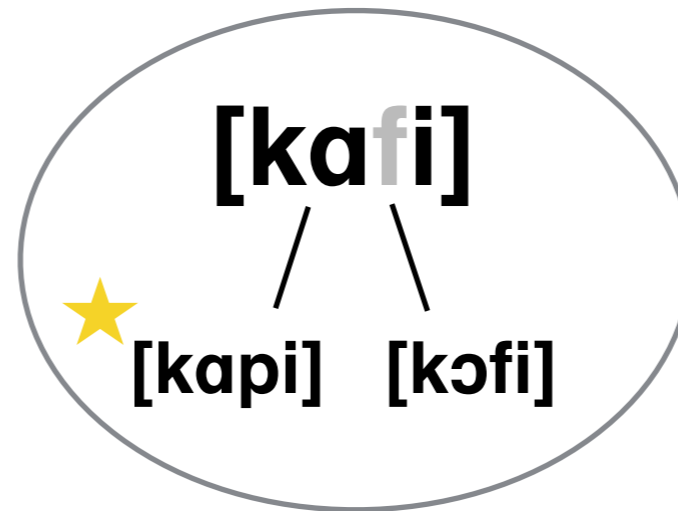


Background

Can you get me
a coffee?



merged speaker



unmerged speaker

A copy? of
what?

Background

Some evidence from the CDC project (Labov 1989):

- Collection of natural misunderstandings

Table 11.1 Distribution of naturally occurring misunderstandings of /o/ and /oh/ by phonemic system of speaker and listener

	<i>Listener</i>			
	<i>Two-phoneme</i>		<i>One-phoneme</i>	
<i>Speaker</i>				
Two-phoneme	a	2	b	1
One-phoneme	c	23	d	1

- Participants more accurately identified problematic words of a different dialect when they were left blank than when a misleading (e.g., Northern-cities-shifted) word was heard in sentence context.

Background

“misleading phonetic information can be a worse obstacle to comprehension than no phonetic information at all”

(Herold 1990)

Research Question

Are merged speakers actually more accurate than unmerged speakers when listening to merged speech?

Methods

Task

-Participants ...

- heard full sentences with a potentially ambiguous item (e.g., dawn/don)
- typed out the entire sentence that they heard
- completed a same/different judgment task for ALL test items
- completed a questionnaire asking about all places they've lived

-The task was administered in PsychoPy

-All sentences and same/different questions were randomized

Methods

Stimuli

- 18 word pairs
- 3 sentences contexts:

LOT-biased: Her thick hair is always *knotty*

THOUGHT-biased: The child was acting *naughty*

neutral: They described it as being *knotty/naughty*

Semantically biased rather than syntactically biased

Methods

Stimuli

The same sound clip was spliced into all three sentences

Participants received 1 of 3 lists so that they only heard one sentence for each word pair

Methods

Stimuli

Frequency (Log10cd measure from SUBTLEX) roughly balanced across word classes

LOT class Freq = 1.95

THOUGHT class Freq = 2.08

Methods

Stimuli

Pink noise (+10db SNR) added to each sound file using Audacity

1. to avoid ceiling effects
2. for ecological validity
3. to facilitate confusion between near-minimal pairs (coffee~copy) without masking vowels

Methods

Stimuli

Minimal pairs

- stock / stalk
- donned / dawned
- pond / pawned

Methods

Stimuli

Minimal pairs

- stock / stalk
- donned / dawned
- pond / pawned

Pseudo-minimal pairs

- coffee / copy
- Anne applauded / Anna plotted
- Mike's water / Mike's swatter

Methods

Stimuli

Minimal pairs

- stock / stalk
- donned / dawned
- pond / pawned

Pseudo-minimal pairs

- coffee / copy
- Anne applauded / Anna plotted
- Mike's water / Mike's swatter

Methods

Stimuli

Minimal pairs

- stock / stalk
- donned / dawned
- pond / pawned

Pseudo-minimal pairs

- coffee / copy
- Anne applauded / Anna plotted
- Mike's water / Mike's swatter

Methods

Stimuli

A merged female speaker produced all tokens

This speaker merges toward LOT

Unmerged speakers would be expected to confuse THOUGHT words produced by this speakers as LOT words

Methods

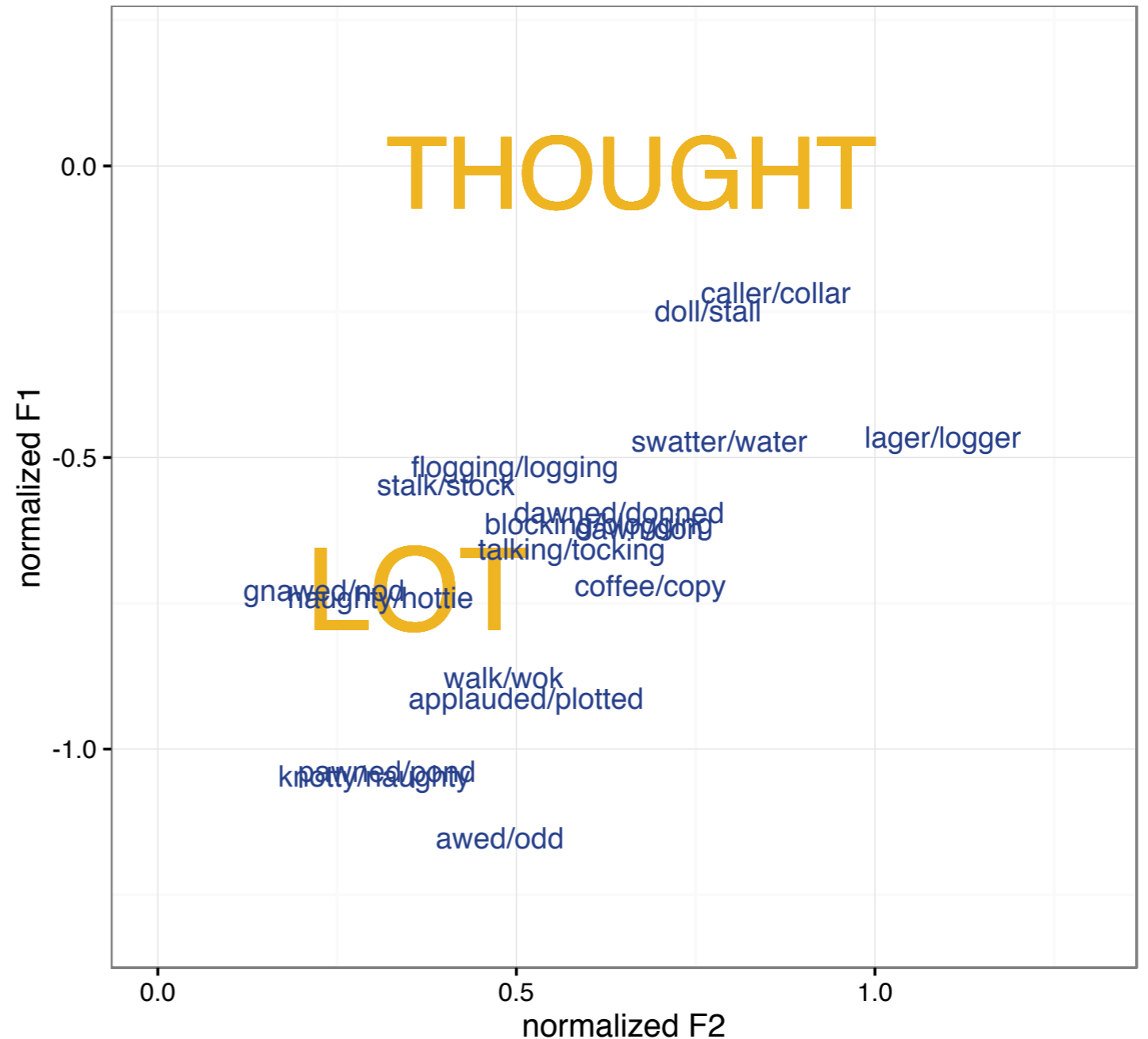
Stimuli

A merged female speaker produced all tokens

This speaker merges toward LOT

Unmerged speakers would be expected to confuse THOUGHT words produced by this speakers as LOT words

F1/F2 of (merged) stimuli plotted over mean F1 and F2 for an unmerged speaker



Methods

Fillers

16 sentences from a BKB Standardized Sentence List with the same level of pink noise as the target sentences

- Any LOT and THOUGHT words were changed
- A total score out of 50 was computed
- No difference in comprehension in noise between merged and unmerged speakers or across lists
 - **Merged = 49.5** **Unmerged = 49.5**
 - **A = 49.6** **B = 49.4** **C = 49.4**

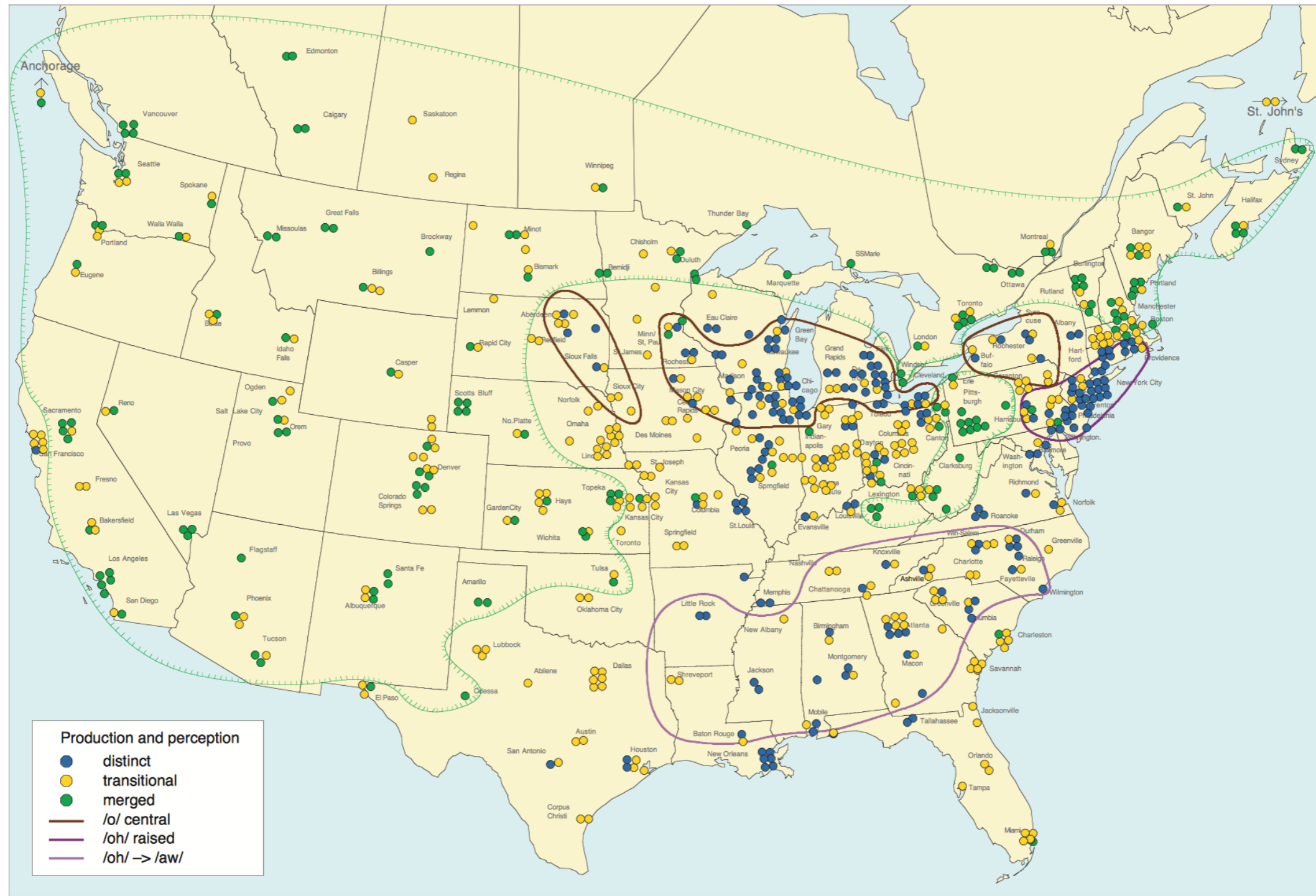
Methods

Participants

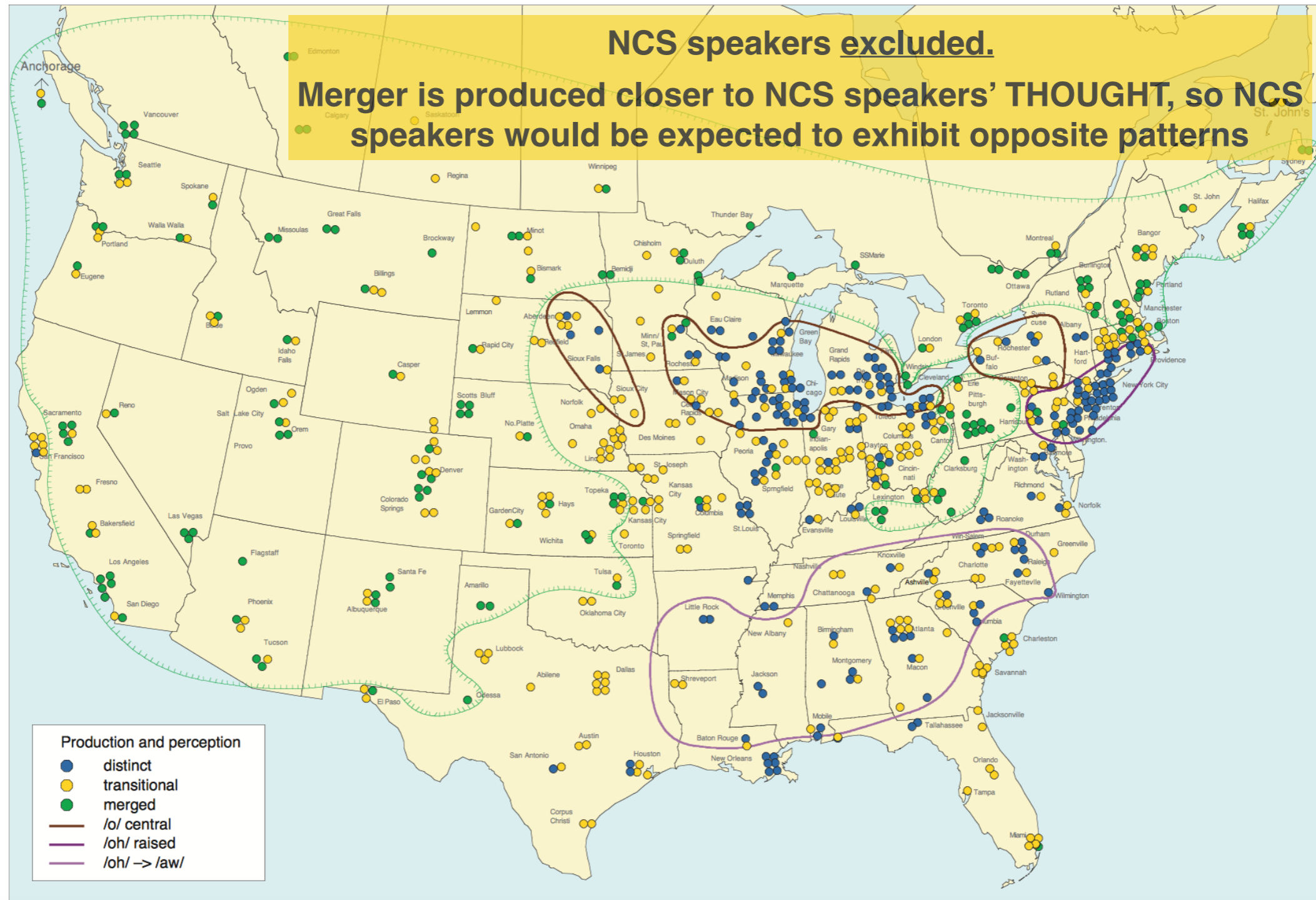
109 undergraduates from the University of Pennsylvania
all native speakers of American English
Roughly 1/3 were merged

Get number from East Coast area

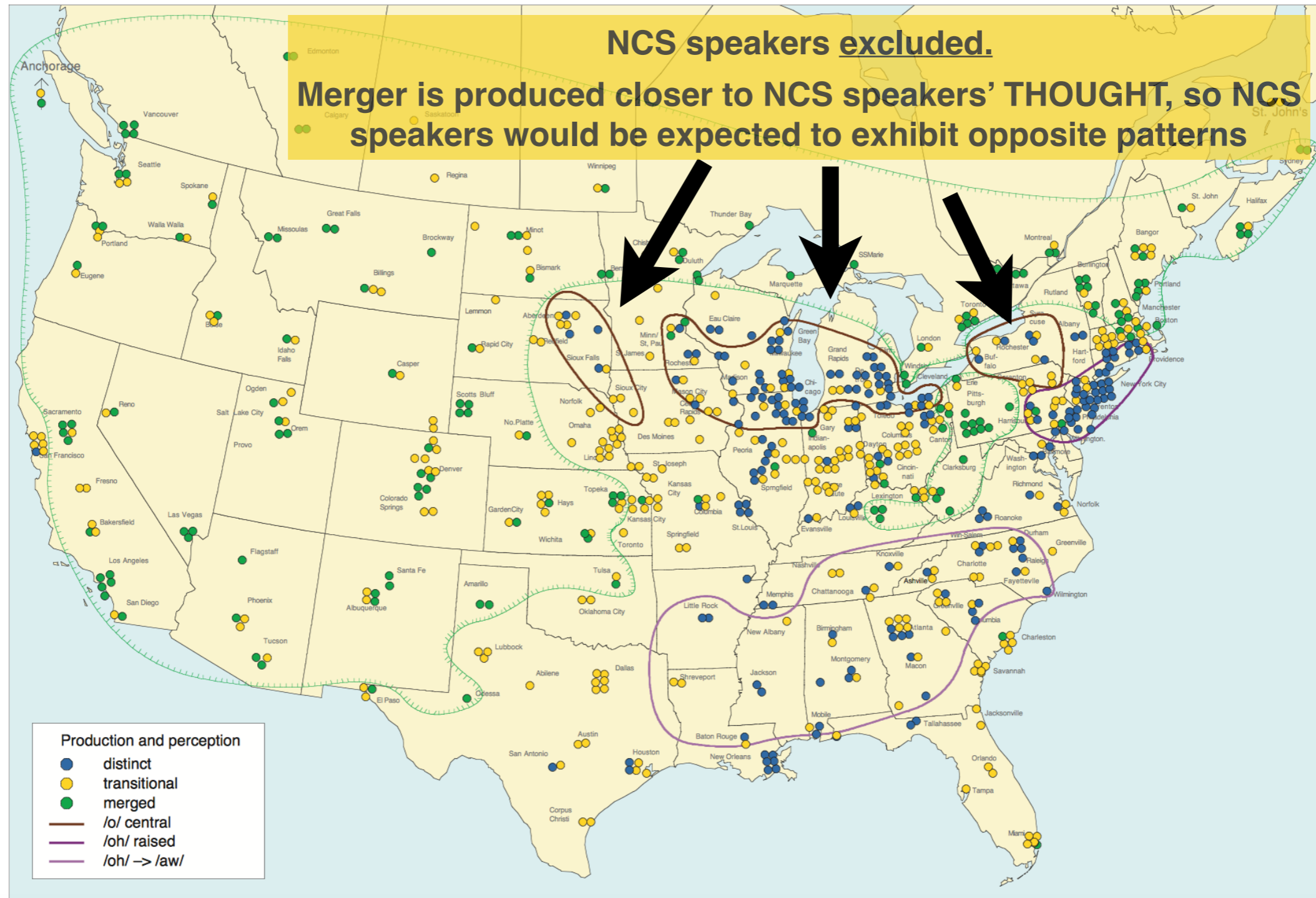
Methods



Methods



Methods



Methods

Analysis

Merged & Unmerged labels were given to each participant **for each word pair**

Labels determined by their responses to all 18 word-pairs in the same/different judgment task administered after the listening task

Excluded responses where a word not in the LOT or THOUGHT class was guessed

Mixed Effects Logistic Regression models fit in R

Results

Are merged speakers more accurate?

Prediction

Unmerged speakers should have overall lower accuracy.

Unmerged speakers may be misled by phonetic information that conflicts with contextual information.

Merged speakers should rely primarily on context and should be more accurate.

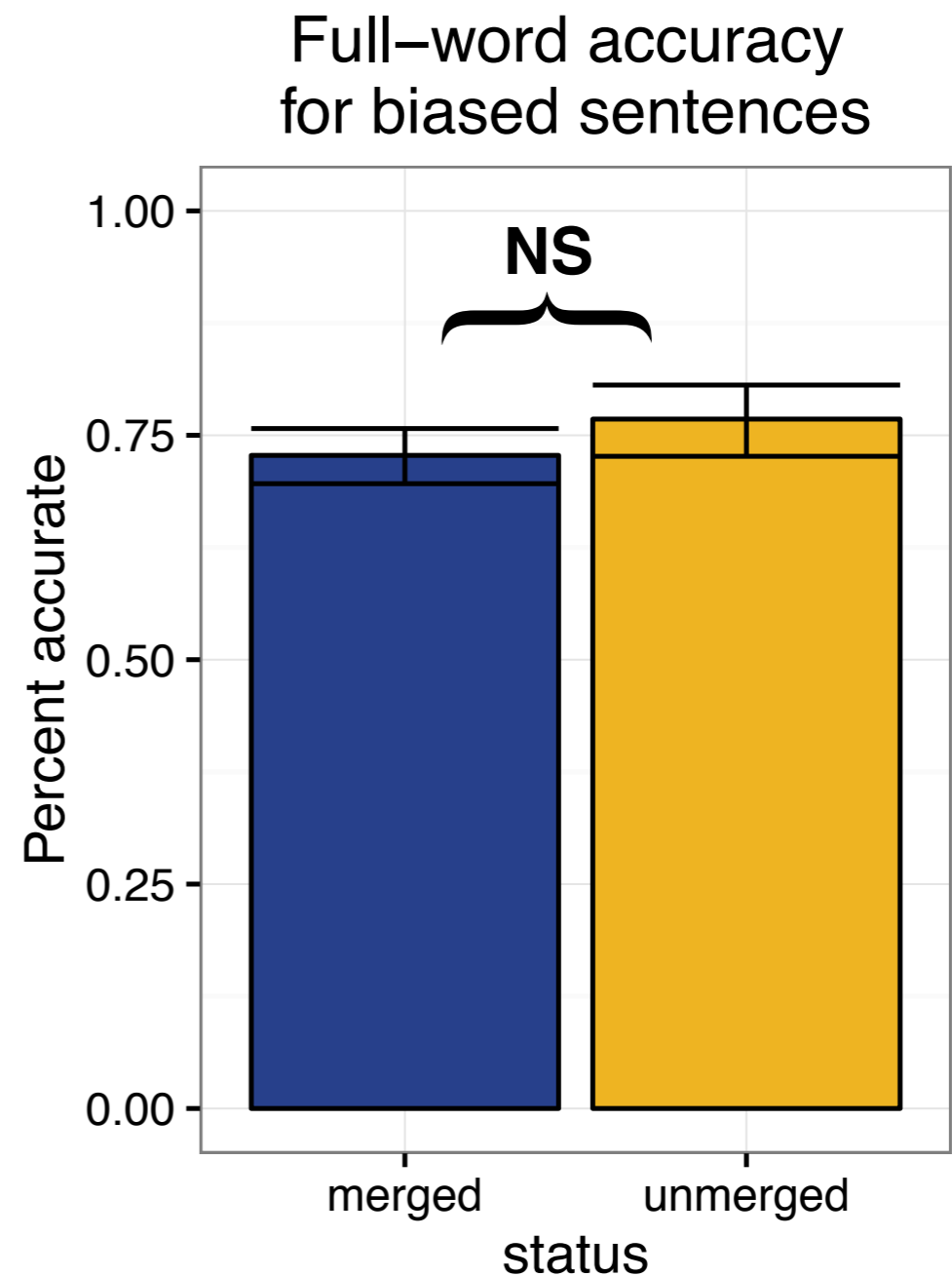
Are merged speakers more accurate?

Prediction

Unmerged speakers should have overall lower accuracy.

Unmerged speakers may be misled by phonetic information that conflicts with contextual information.

Merged speakers should rely primarily on context and should be more accurate.



Are merged speakers more accurate?

A lack of differences in accuracy might result from two opposite effects:

1. **Unmerged speakers might be more accurate for LOT words** because they are getting helpful phonetic information that merged speakers might not be using.
2. **Unmerged speakers might be less accurate for THOUGHT words** because they are receiving misleading phonetic information that merged speakers might not be using

What about for THOUGHT-biased sentences?

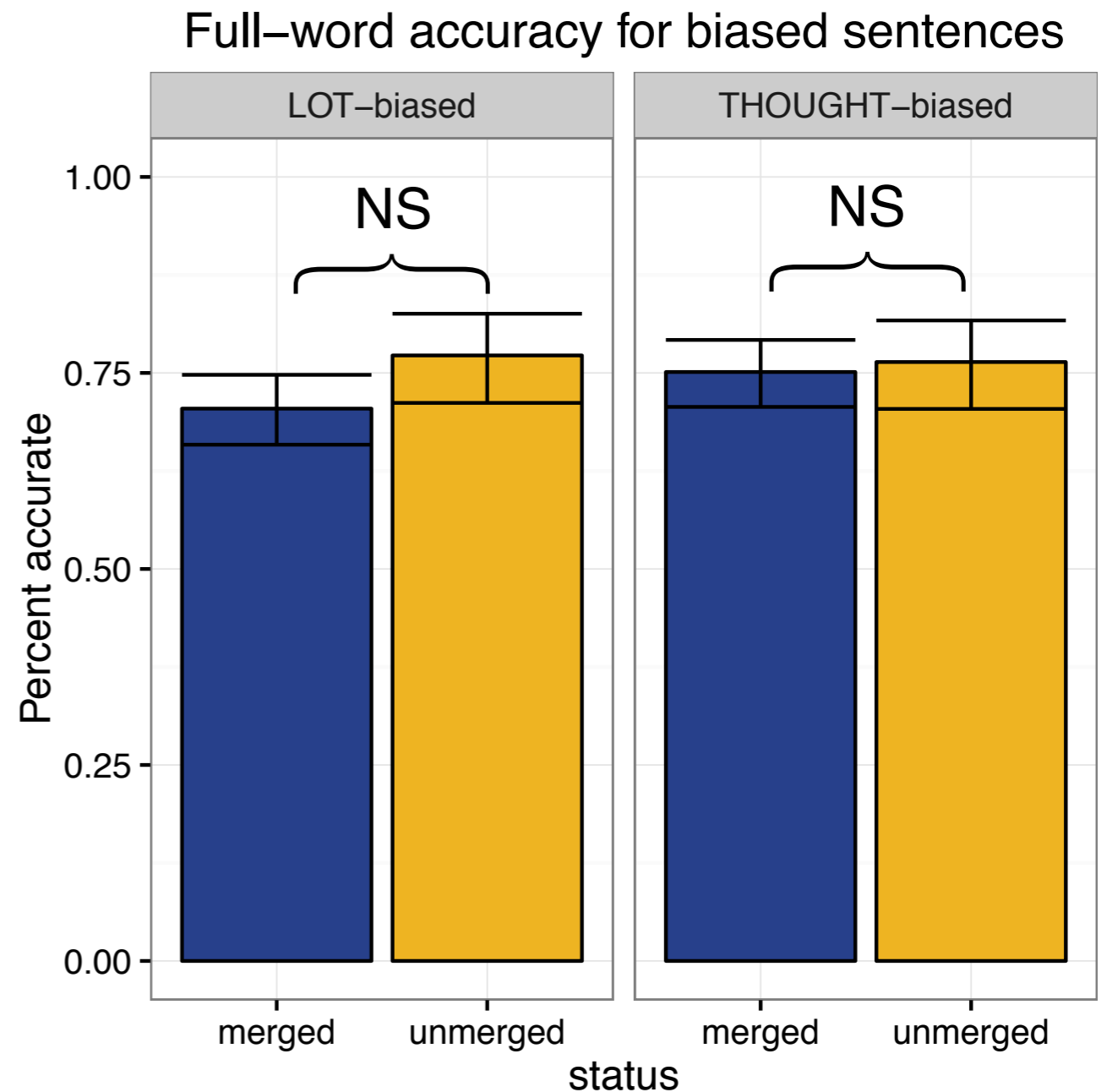
Prediction

This is the context in which phonetic information and contextual information conflict for an unmerged speaker. Competing conflicting cues might result in lower accuracy only in this context.

What about for THOUGHT-biased sentences?

Prediction

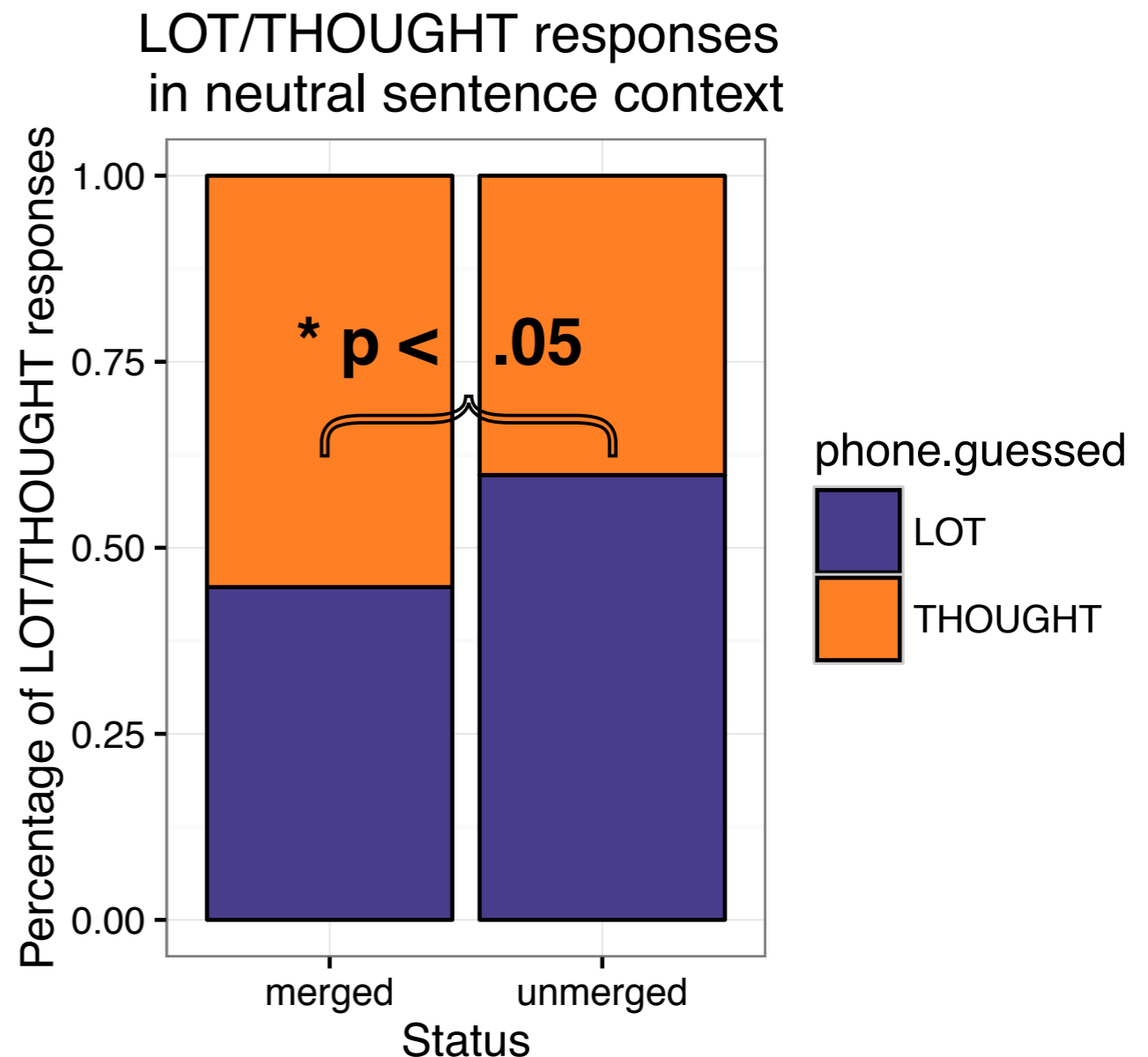
This is the context in which phonetic information and contextual information conflict for an unmerged speaker. Competing conflicting cues might result in lower accuracy only in this context.



Are there *any* differences between merged and unmerged speakers?

There is a significant interaction between sentence type (biased vs. neutral) and word class guessed.

Unmerged speakers guess significantly more LOT in neutral contexts

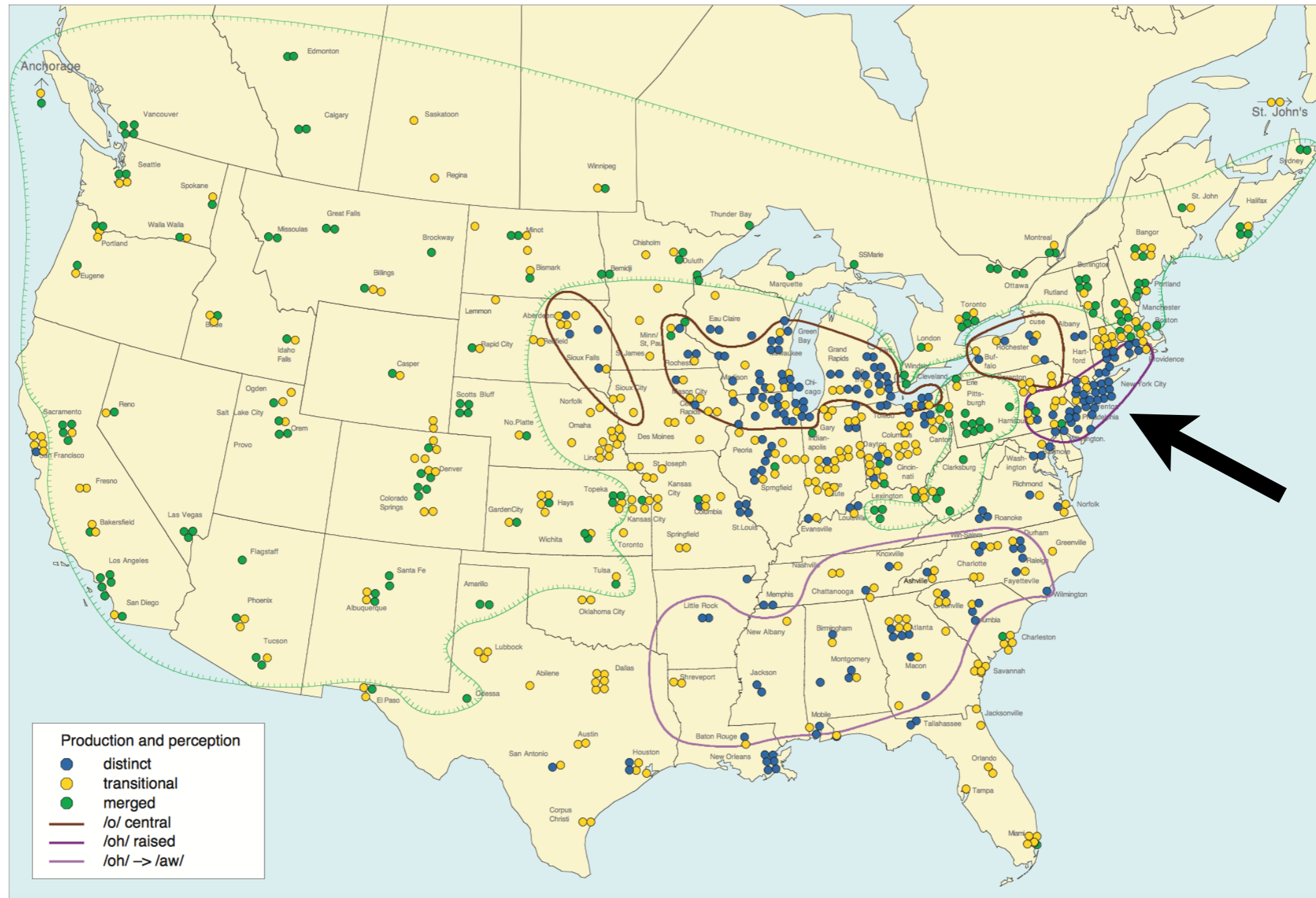


Limiting the geographic scope

Maybe unmerged speakers who have lived in merged areas are extremely accurate for sentences with contextual information because they have already abandoned phonetic cues

What if we control for the area in which participants have lived?

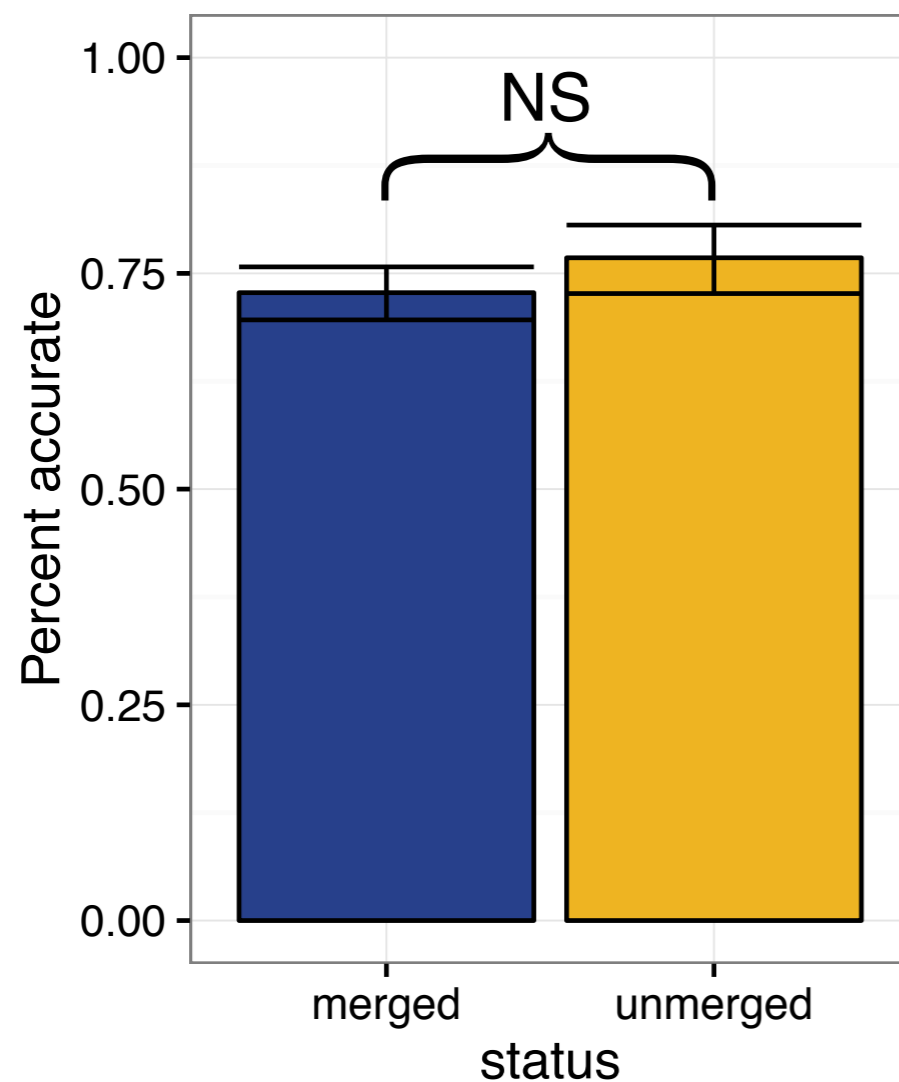
Limiting the geographic scope



Limiting the geographic scope

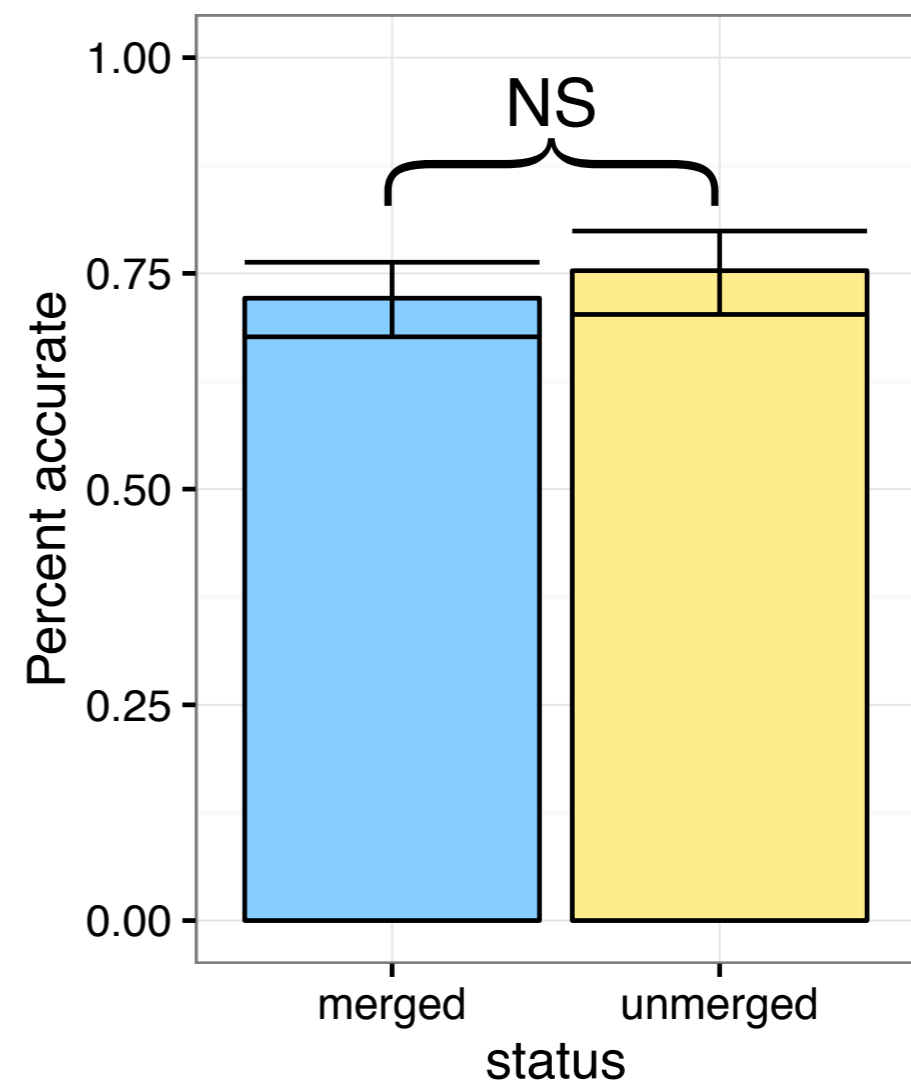
Total Dataset

Full-word accuracy
for biased sentences



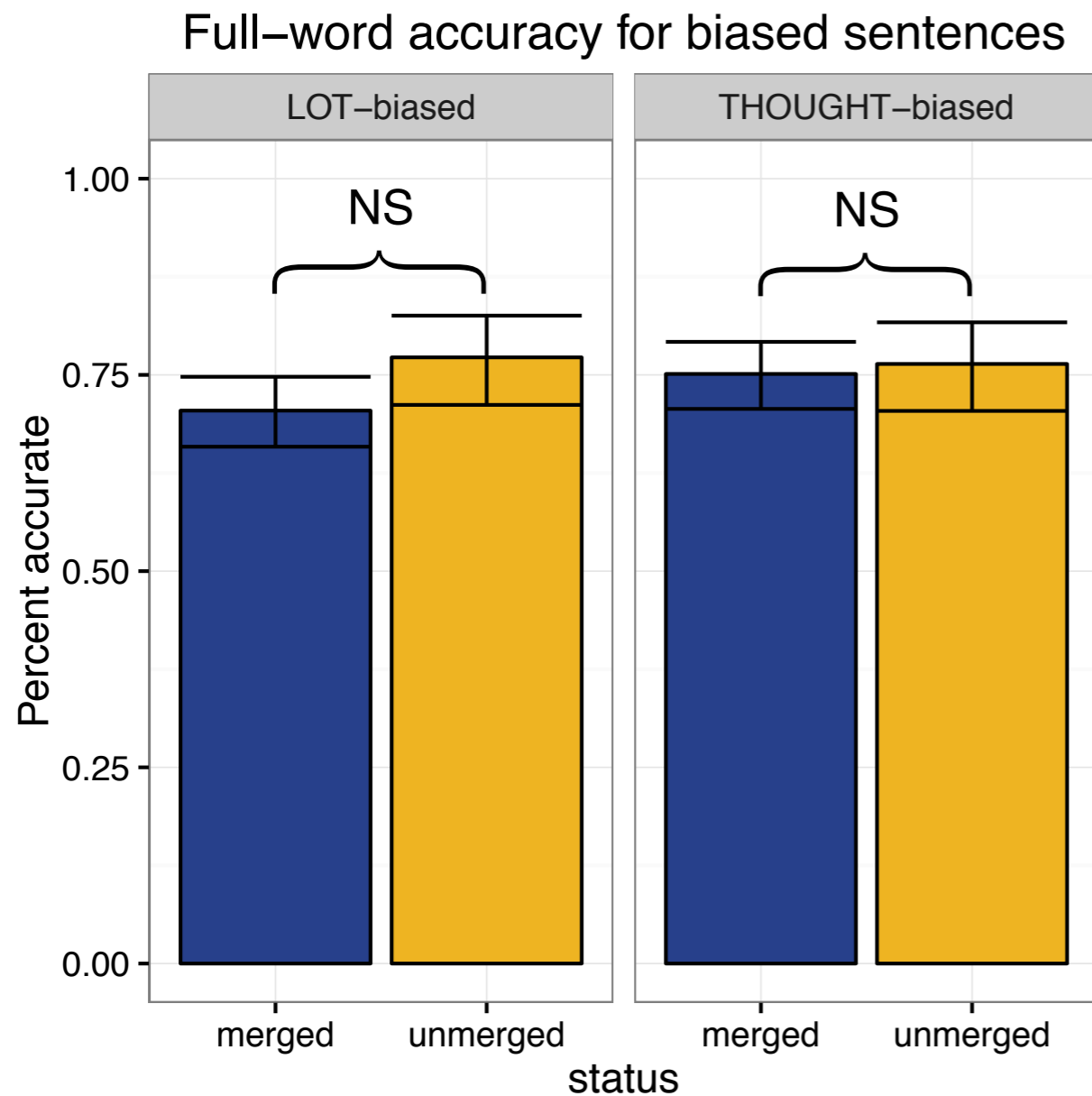
East Coast

Full-word accuracy
for biased sentences

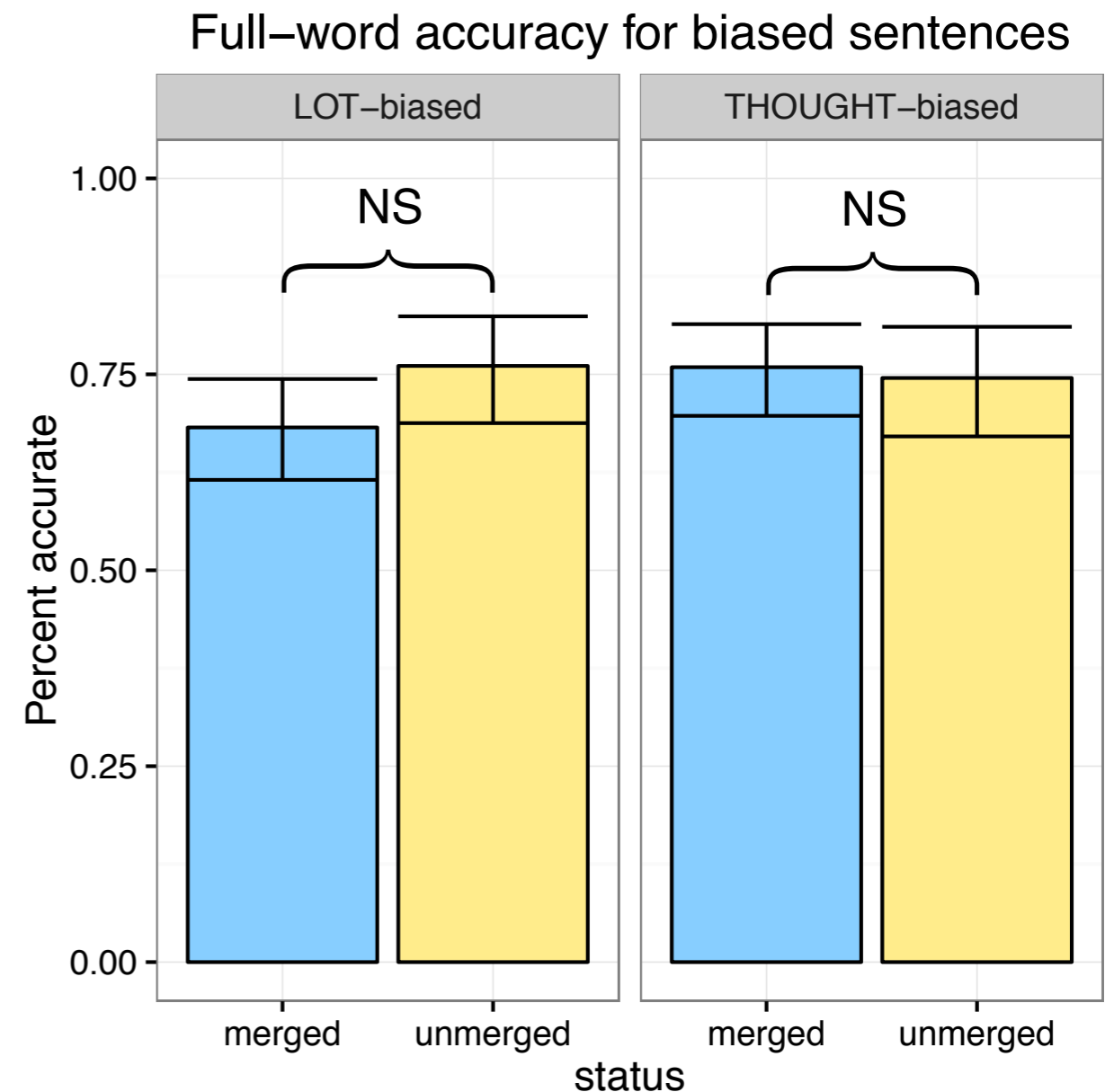


Limiting the geographic scope

Total Dataset



East Coast



Summary

Unmerged and merged speakers have equal accuracy rates

Unmerged speakers are biased towards LOT when contextual information is impoverished

Unmerged speakers perform similarly to merged speakers in sentences with contextual information, suggesting they use primarily contextual information and are not misled by phonetic cues

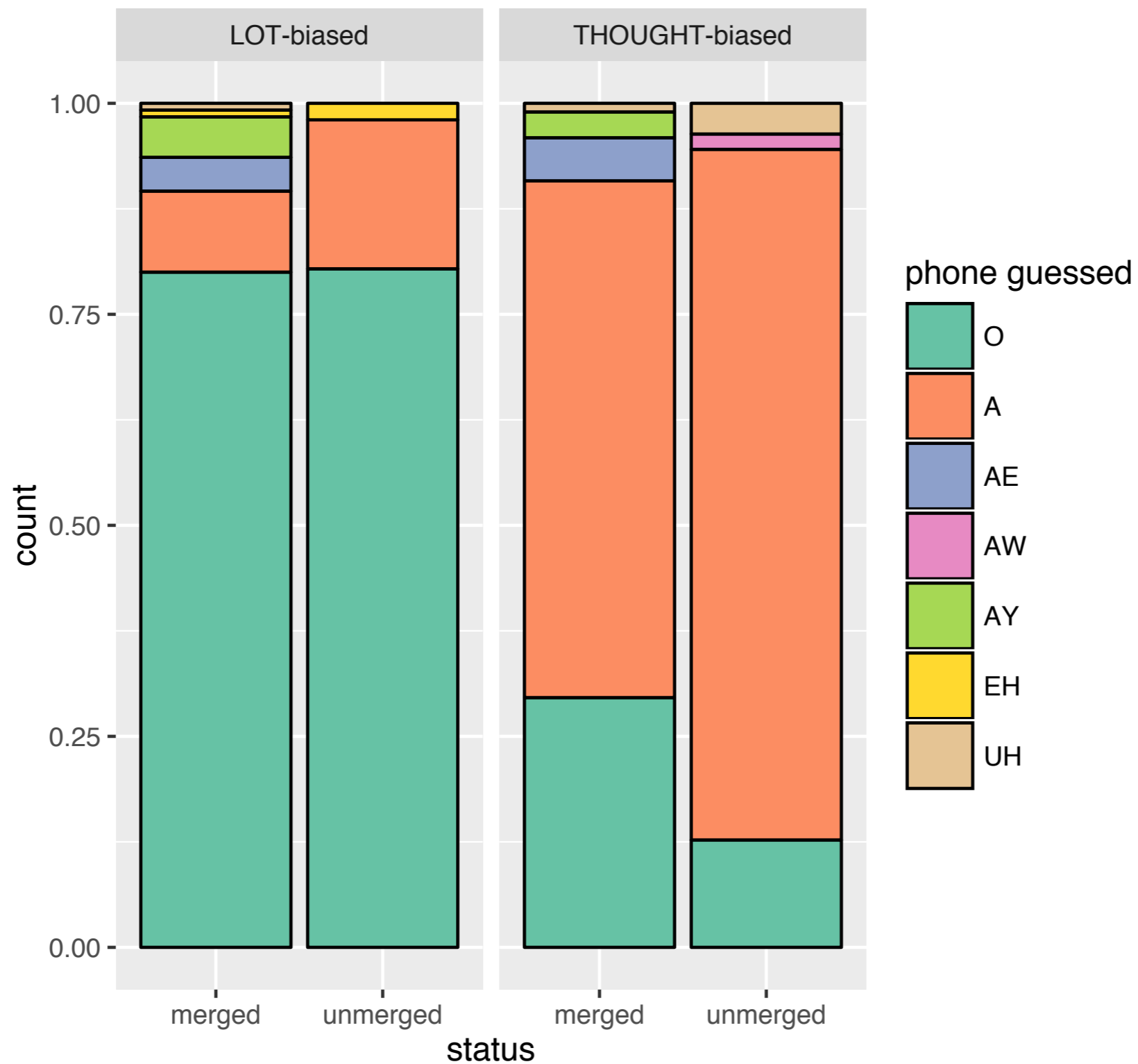
Discussion

There are other ways that unmerged speakers might be at a disadvantage than accuracy alone (speed)

Unmerged speakers might rely more on context in adverse listening conditions (pink noise)

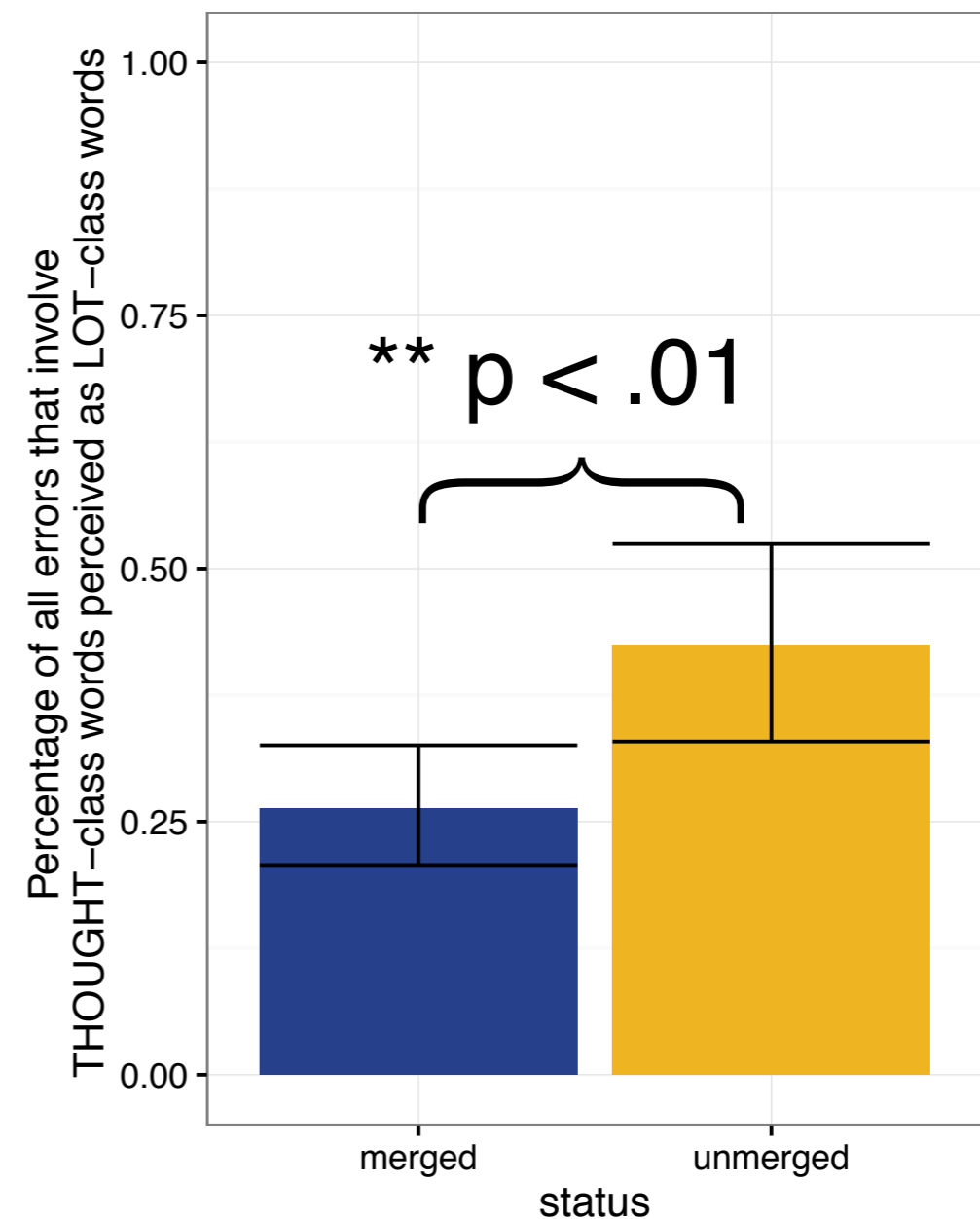
Supplementary Slides

Breakdown of errors by vowel guessed



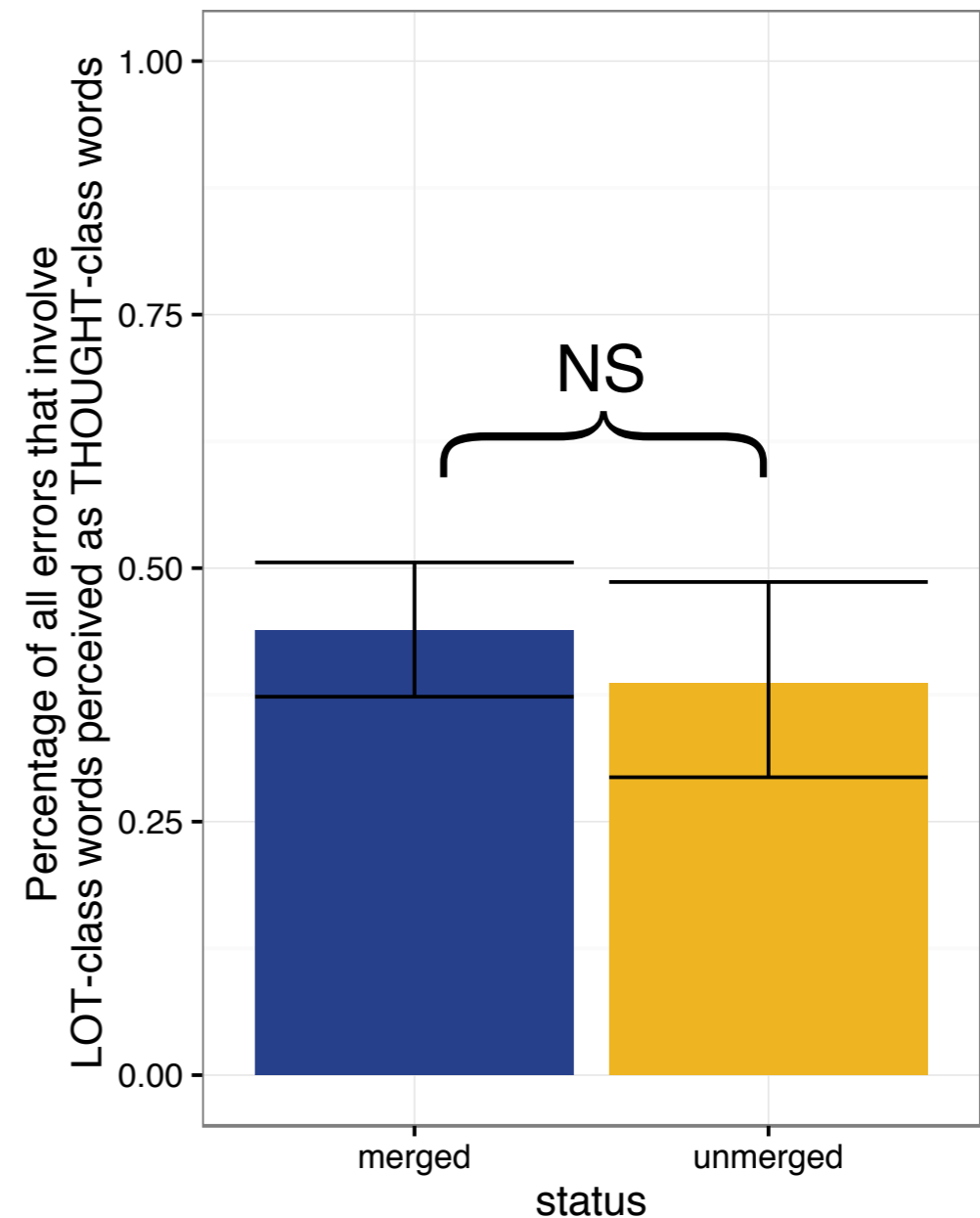
Are there *any* differences between merged and unmerged speakers?

What percentage of all errors involve misinterpretation of THOUGHT as LOT?



Are there *any* differences between merged and unmerged speakers?

What percentage of all errors involve misinterpretation of LOT as THOUGHT?



Effect of frequency

