

Online Language Processing and Listening Effort in the Face of Unexpected Talker Information

Sarah Colby & Bob McMurray

Department of Psychological & Brain Sciences, University of Iowa

Auditory Perception, Cognition, & Action Meeting

November 16, 2023

Why do we care about talker information during speech processing?

Two ways to take advantage of talker identity (Luthra, 2023):

- Streaming
- Normalization



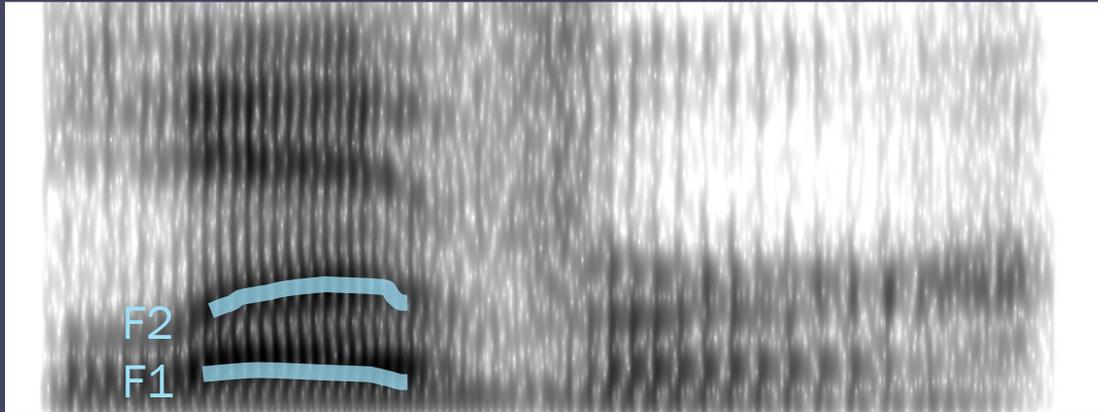
Talker Streaming

Listeners need to be able to separate speech they're listening to from background speech or noise

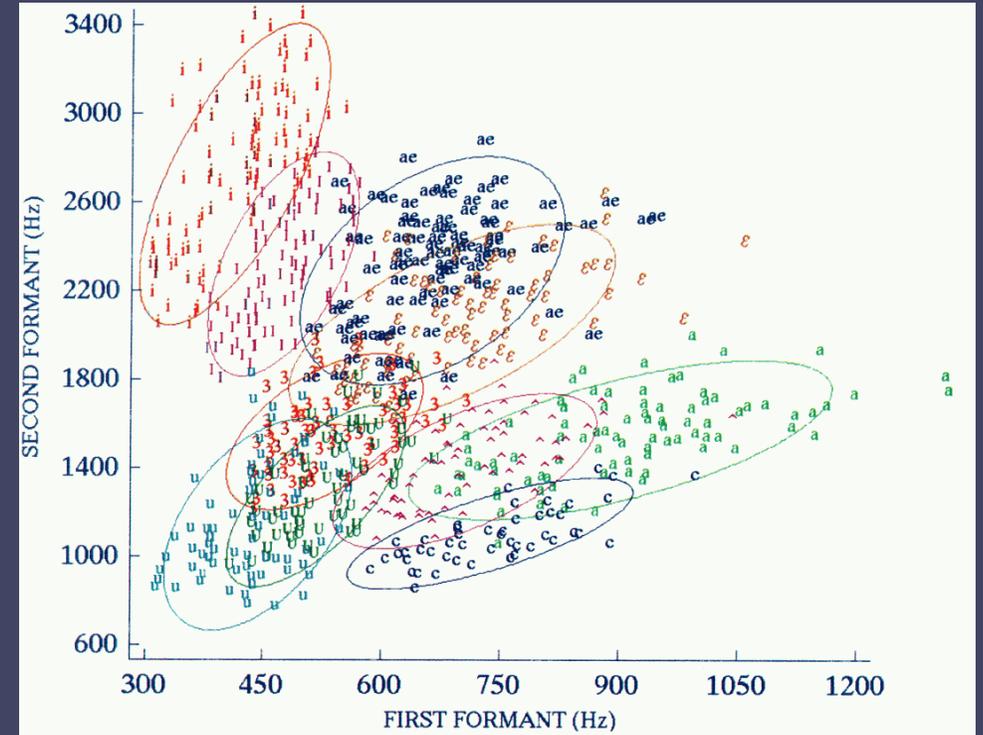
- Can use spatial location, but also cues to talker identity (F0, vowel formants)
- Easier to separate streams when talkers are acoustically dissimilar

But what about when you're not in noise?

Talker Normalization



Listeners need to accommodate for variation in production across talkers



Effects of Talker Variability

Talker variability is sometimes beneficial, but there is likely a processing cost

- Listeners will respond slower when the talker changes unexpectedly (Magnuson & Nusbaum, 2007; Magnuson, Nusbaum, & Akahane-Yamada, 2021)

Language processing system is robust

- Listeners are primed by words produced by two talkers (e.g., 'my' 'great' primes 'goose') (Newman, 2016)

Experiments

1. Sentence processing in noise (babble)

How does unexpected talker information affect talker streaming?

- Introducing an unexpected talker will disrupt real-time processing

2. Sentence processing with acoustic ambiguity

- How does unexpected talker information affect normalization?

3. Effort engaged for sentence processing with unexpected talker information

- Does listening effort increase in the face of unexpected talker information?

Experiment 1: Talker Streaming

How does unexpected talker information affect talker streaming?

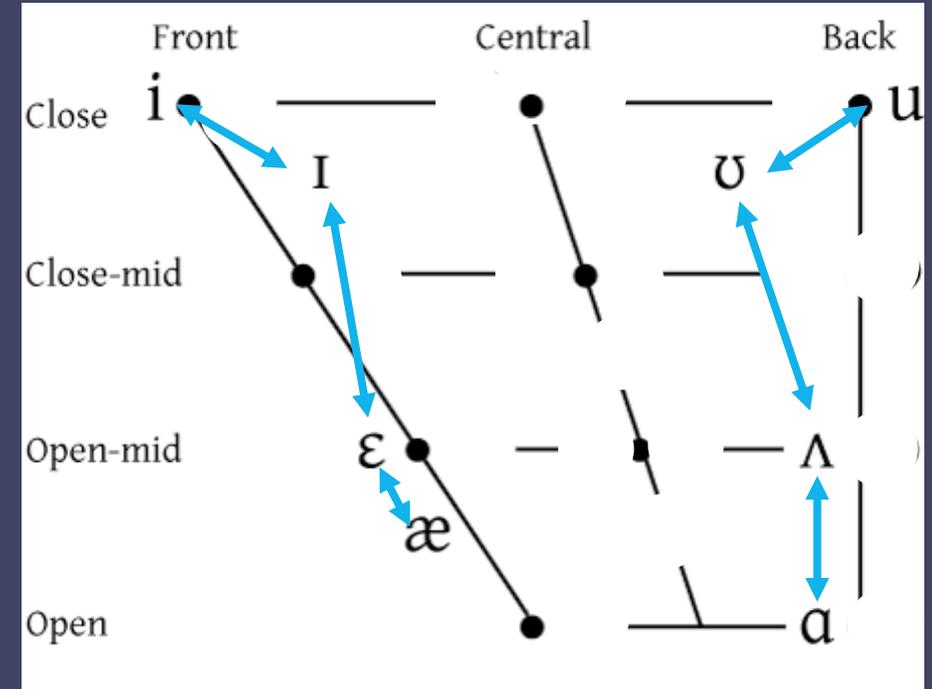
22 minimal pairs that differed in vowel height

- chick-check, vet-vat, shook-shuck, etc.

Spliced into sentences

- “This time, choose _____ [from the display]”

Embedded in multi-talker babble at 0 SNR

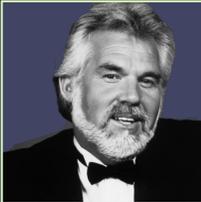
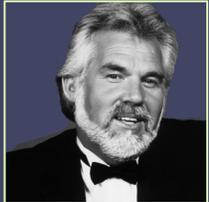


Experiment 1: Talker Streaming

Sentence Voice Target Voice



Look at how cute my cat is, what we are
This time, I selected the book
Islands in the sky
I heard that on the screen



Look at how cute my cat is, what we are
This time, I selected the book
Islands in the sky
I heard that on the screen

Match



Experiment 1: Visual World Paradigm

Visual displays contain 2 minimal pairs

- Target, Competitor, 2 unrelated

2/3 match trials, 1/3 mismatch trials

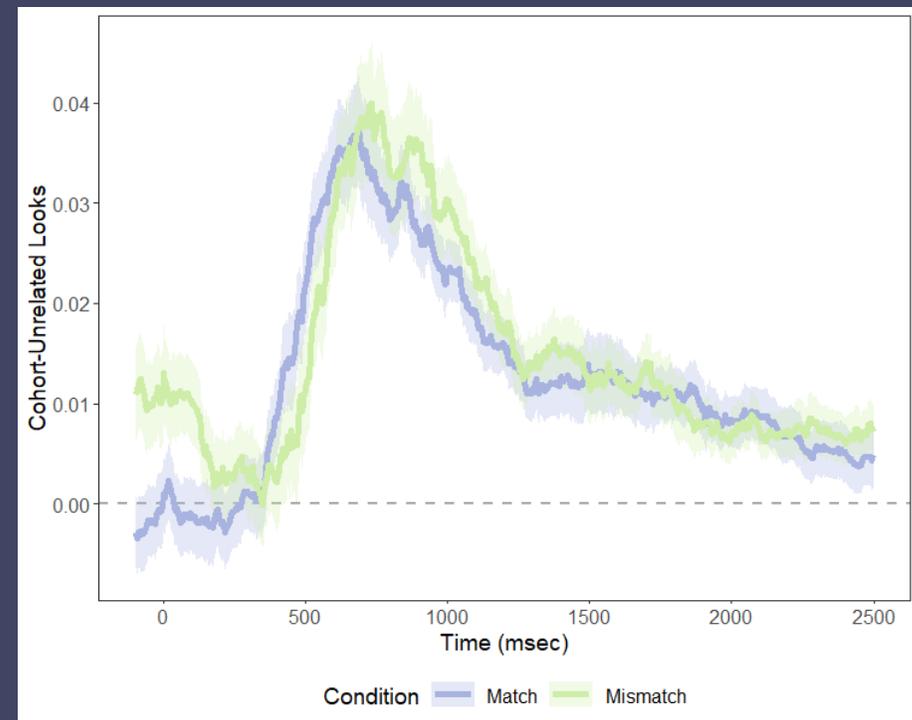
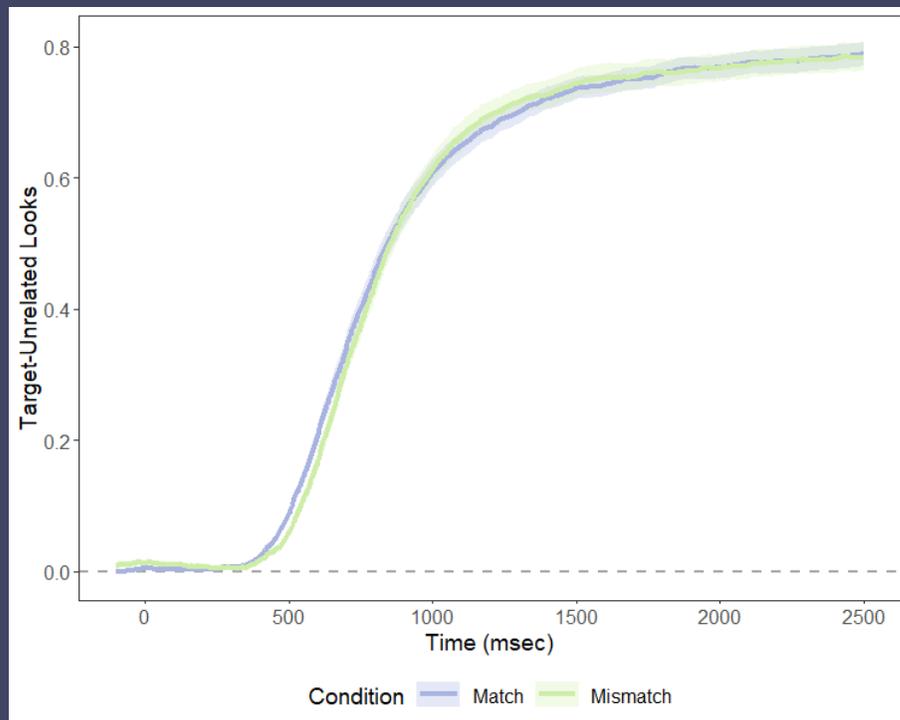
- Start with a block of match trials, then trials are mixed



Results – Sentences in Babble

Condition	Sentence Voice	Target Voice	Target	Cohort
Match	Female	Female	0.96	0.03
	Male	Male	0.93	0.06
Mismatch	Male	Female	0.95	0.03
	Female	Male	0.94	0.04

N=40



Experiments

1. Sentence processing in noise (babble)

How does unexpected talker information affect talker streaming?

- Maybe it doesn't! No evidence that switching disrupts processing

2. Sentence processing with acoustic ambiguity

How does unexpected talker information affect normalization?

- Introducing an unexpected talker change will disrupt real-time processing

3. Effort engaged for sentence processing with unexpected talker information

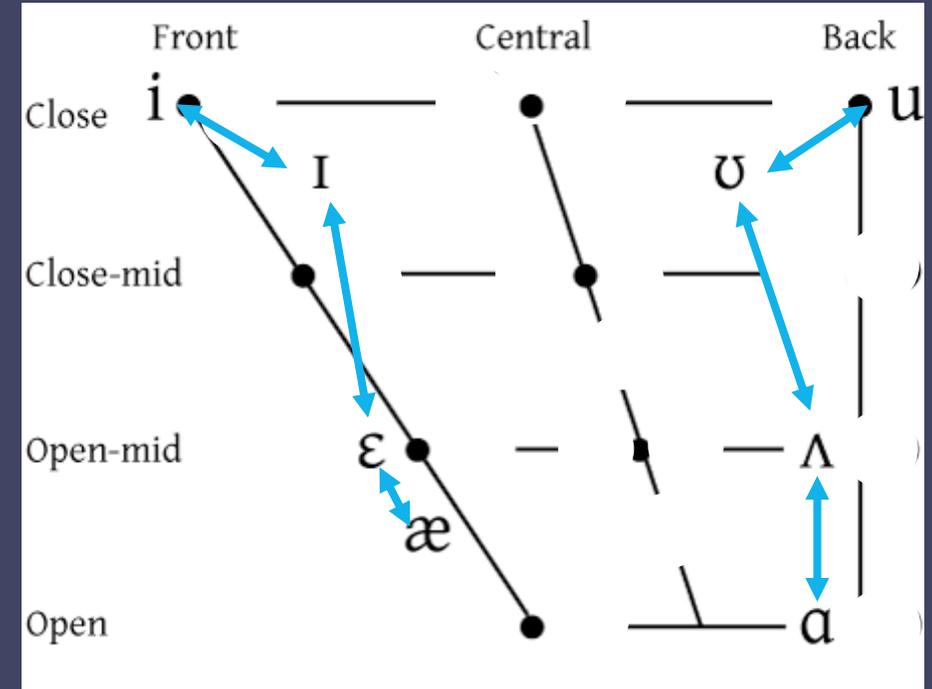
- Does listening effort increase in the face of unexpected talker information?

Experiment 2: Vowel Normalization

Same minimal pairs → 11-step vowel continua

- Found ambiguous steps (~90/10 response rate)

Slightly ambiguous targets spliced into sentences



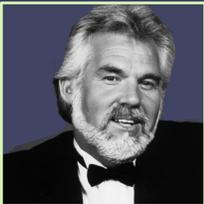
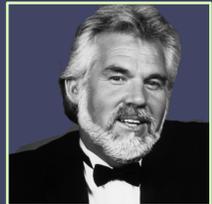
Experiment 2: Vowel Normalization

Sentence
Voice

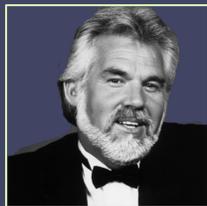
Target
Voice



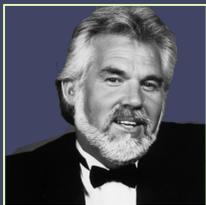
This time, select sp?ll on the screen



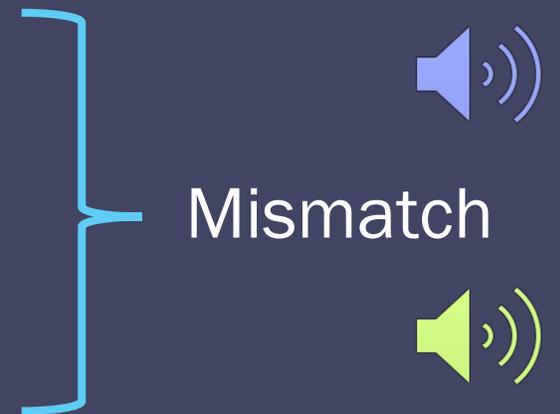
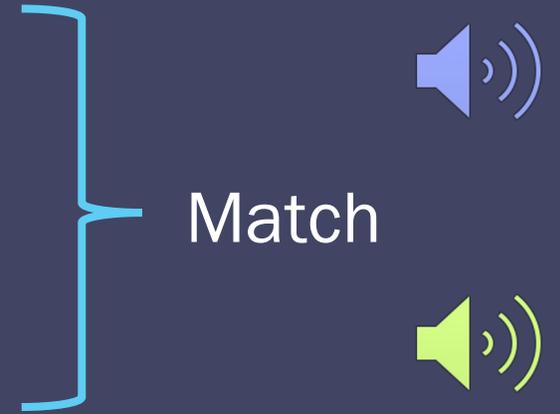
This time, select sp?ll on the screen



This time, select **sp?ll** on the screen



This time, select **sp?ll** on the screen



Experiment 2 – Visual World Paradigm

Same visual displays as Experiment 1

Same distribution of match/mismatch trials

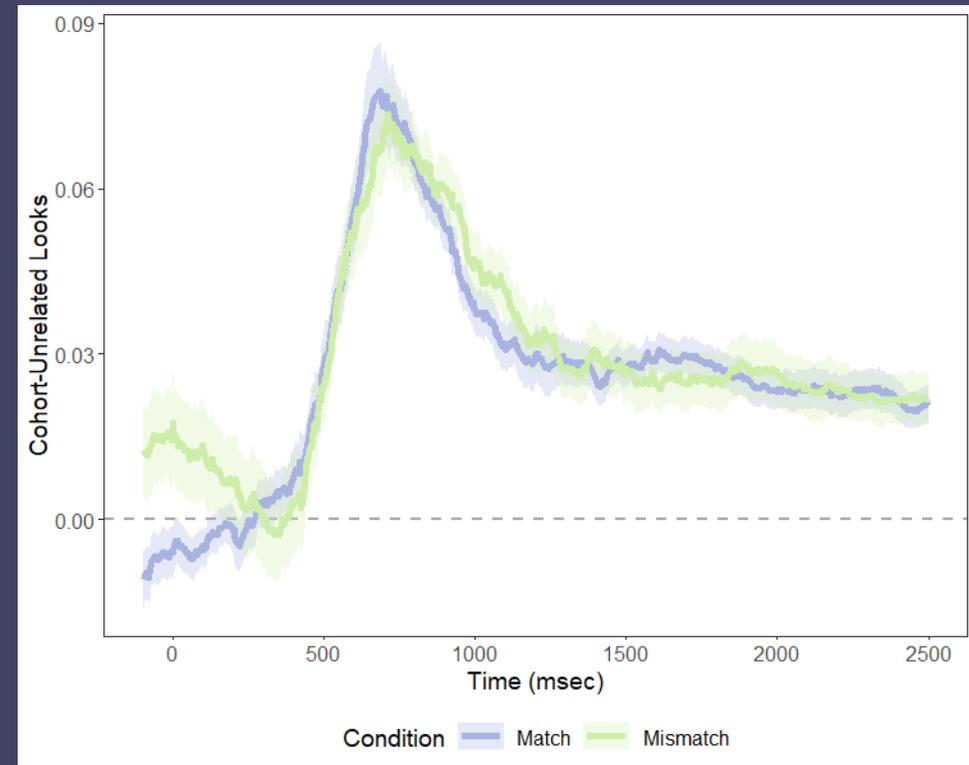
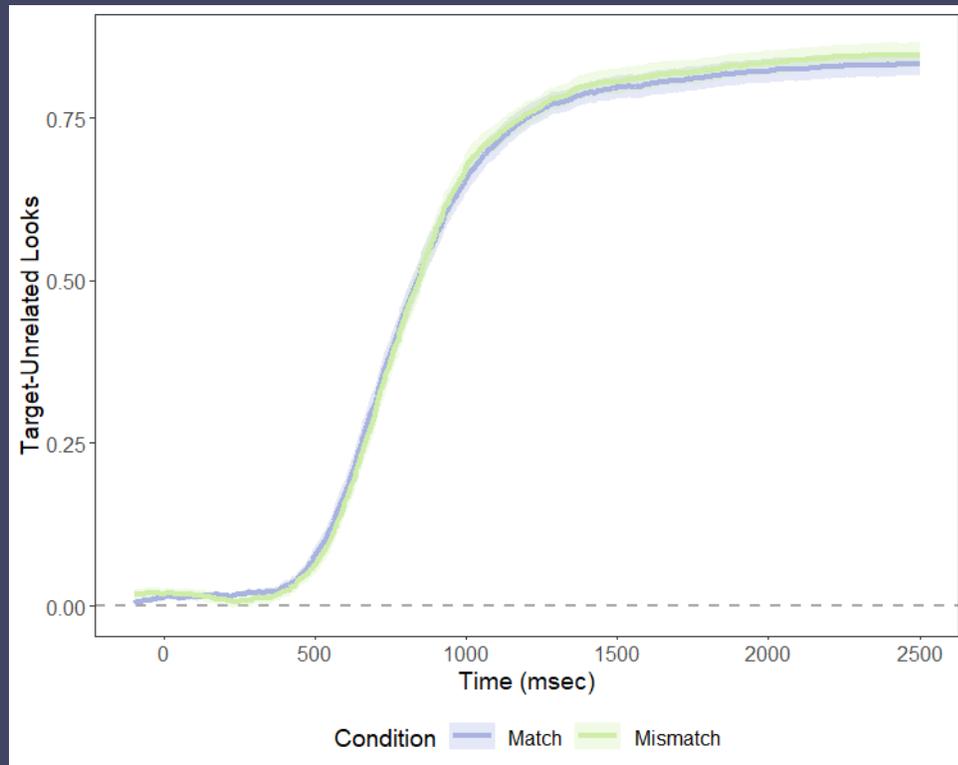
- 2/3 match trials, 1/3 mismatch trials
- Start with a block of match trials, then trials are mixed



Results – Ambiguous Vowel

Condition	Sentence Voice	Target Voice	Target	Cohort
Match	Female	Female	0.93	0.07
	Male	Male	0.88	0.12
Mismatch	Male	Female	0.91	0.09
	Female	Male	0.92	0.08

N=43



Experiments

1. Sentence processing in noise (babble)

How does unexpected talker information affect talker streaming?

- Maybe it doesn't! No evidence that switching disrupts processing

2. Sentence processing with acoustic ambiguity

How does unexpected talker information affect normalization?

- No evidence that switching disrupts processing

3. Effort engaged for sentence processing

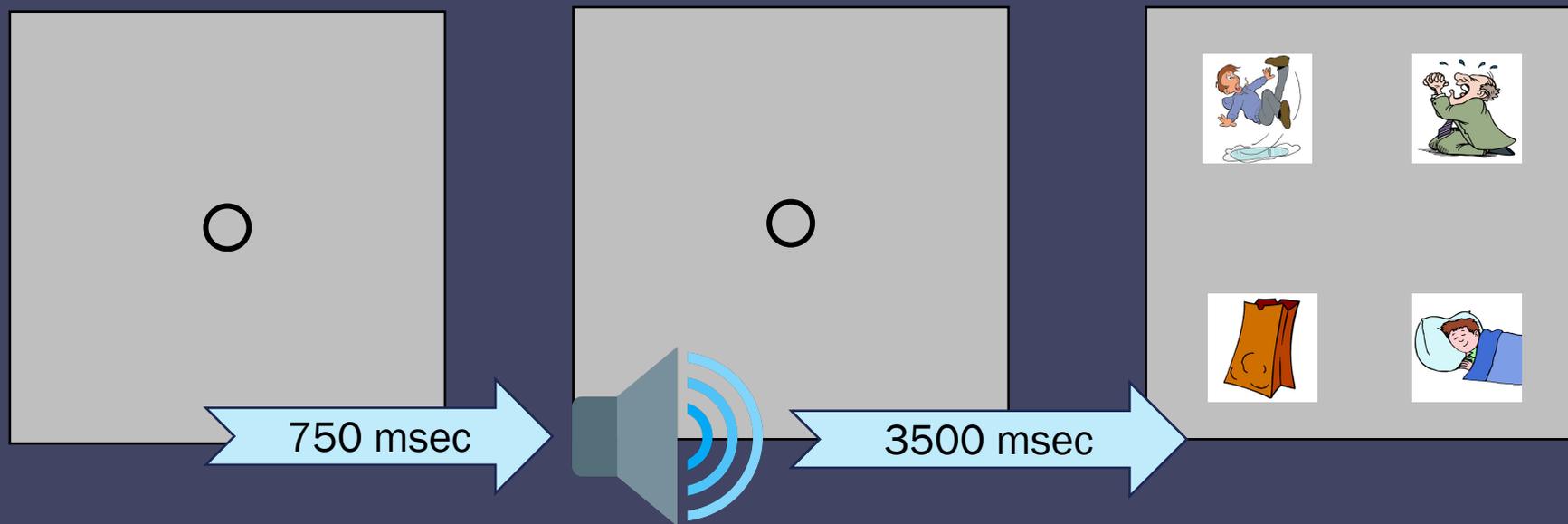
Does listening effort increase in the face of unexpected talker information?

Experiment 3: Vowel Normalization

Pupillometry task: 132 trials

- Ambiguous target words that mismatch voice on 1/3 trials

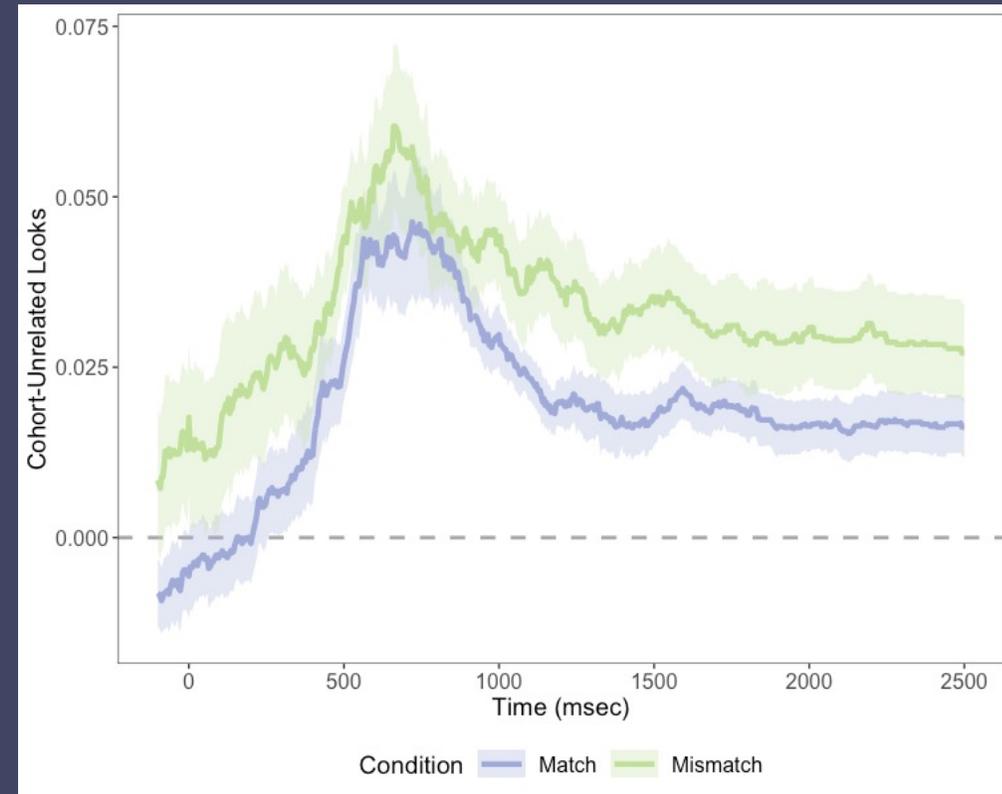
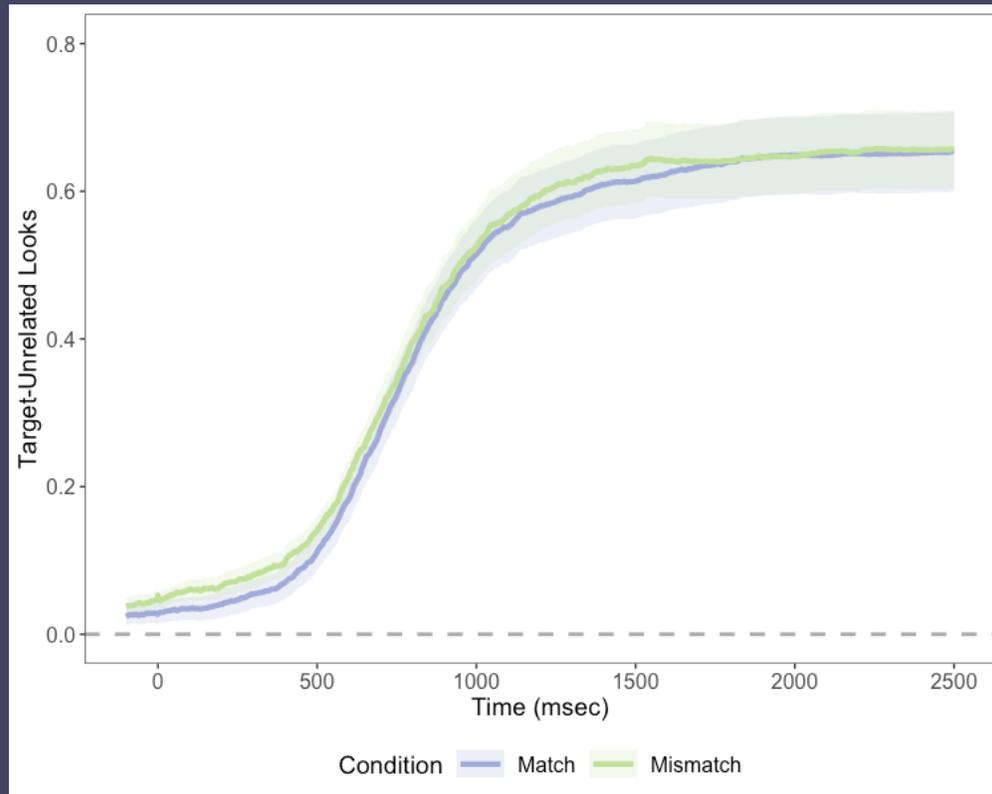
VWP task: 264 trials



Results – Ambiguous Vowel

Condition	Sentence Voice	Target Voice	Target	Cohort
Match	Female	Female	0.96	0.04
	Male	Male	0.92	0.07
Mismatch	Male	Female	0.95	0.05
	Female	Male	0.93	0.07

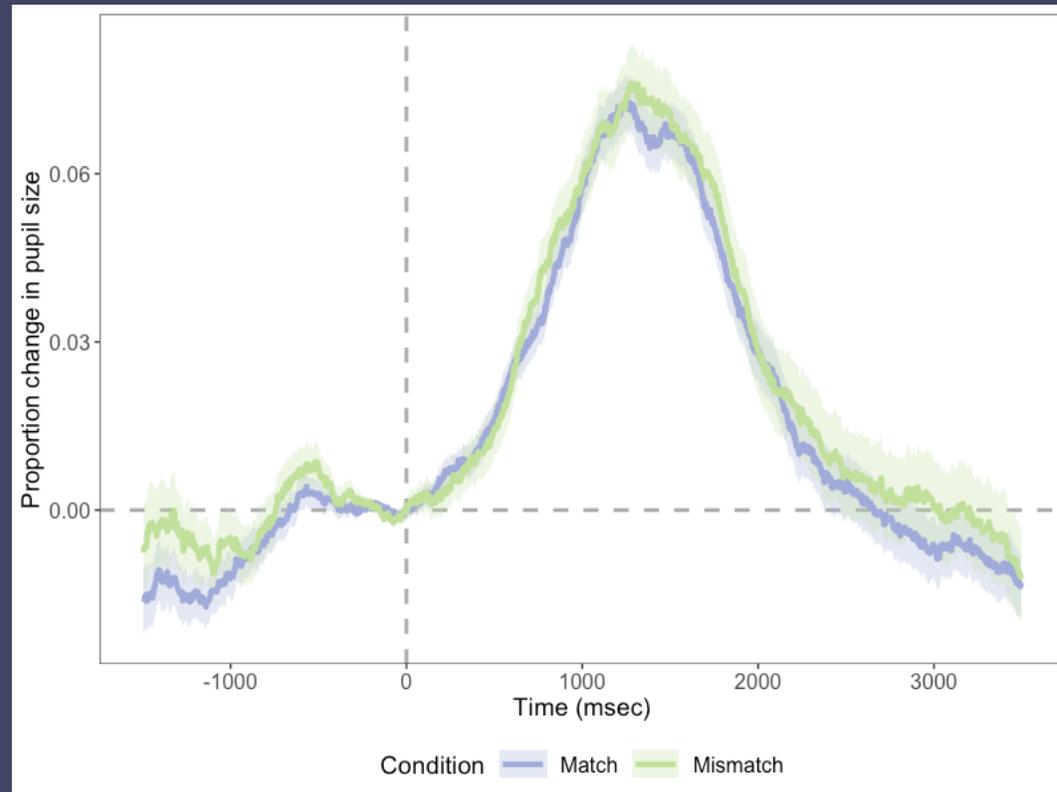
N=20



Pupil Results – Ambiguous Vowel

Condition	Sentence Voice	Target Voice	Target	Cohort
Match	Female	Female	0.95	0.05
	Male	Male	0.89	0.10
Mismatch	Male	Female	0.94	0.06
	Female	Male	0.91	0.09

N=20



Summary

- Robust word recognition in the face of talker mismatches
 - Supports previous evidence that switching talkers mid-word doesn't affect lexical access (Newman, 2016)
- Talker cues are not sufficient to impair word recognition
 - It doesn't even seem to require more effort (although data collection is ongoing)
- Remaining questions
 - Is increased variability necessary to impact performance?
 - Are listeners becoming familiar with two voices over the experiments?

Acknowledgements

Thanks to Marissa Huffman and the rest of MAClab team!

Funding Sources:

NIH NIDCD P50 DC000242 and R01 DC0008089