Talker Interference in Speech Perception Adaptation

Rachel Sabatello, Shawn Cummings, & Florian Jaeger

University of Rochester, Dept. of Brain & Cognitive Sciences

Background

Spoken language is highly variable by nature. Talkers differ in how they produce speech sounds, even when they share similar language backgrounds Still, listeners understand newly encountered talkers when hearing them speak for the first time. To overcome speech variation, the brain actively learns how talkers speak, and constructs expectations about that talker will produce speech in the future. Though this process often occurs without the listener noticing, it still requires cognitive resources. In this experiment, we limit the available resources for speech perception by exposing a listener to two talkers speaking simultaneously. We then test the effects of directing the listener's attention to one talker on the listener's ability adapt to both talkers.

Hypothesis

When listening to two talkers speak simultaneously, listeners will change their perceived categorical boundary more for the talker they are instructed to attend to compared to the unattended second talker.

Predictions

Example /S/-/ʃ/ Perception (Full Adaptation to Both Talkers)

Example /S/-/ʃ/ Perception (Attending to Talker A)
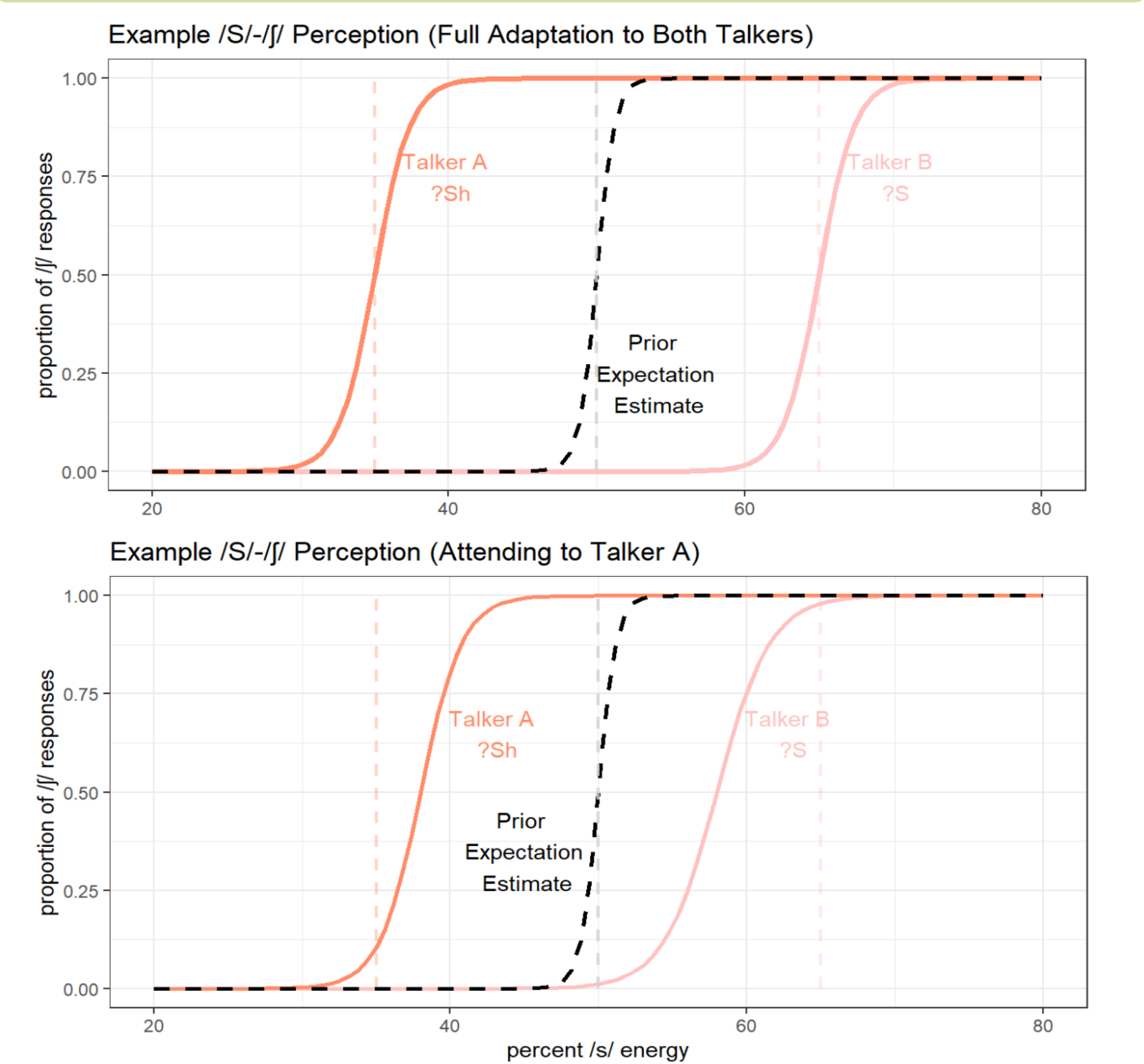
Implications

The results of this experiment begin to explore the role of attention in speech perception adaptation. A listener's perceptual boundary changing more to fit the attended talker's speech than the unattended talker's speech in this experiment would suggest a difference between passive and active attention in speech processing and give insight into how our brains allocates resources under higher cognitive loads.

Design

In this study, we will be measuring listeners' perceptual adaptation to 2 simulated talkers' S-ʃ production.

S-ʃ sounds exist on a continuum, spanning from /s/ as in "Solid" to /sh/ as in "Shore." Earlier research suggests that listener adaptation to talker S-ʃ production is **talker-specific.** This means that listeners adjust their perceived boundary between S-ʃ for each talker regardless of other talkers the listener may also hear (Kraljic & Samuel, 2005). **This quality Could allow us to simulate two distinct talkers with different S-ʃ productions during the same experimental exposure** (Cummings & Theodore, *in press*).

Exposure Phase

Critical Trials

Our critical stimuli are created from 40 recordings of S/ʃ words, each **spoken typically (S, Sh)** and **accented (?S, ?Sh)** (Kraljic & Samuel, 2005).

All recordings were processed using Praat (Boersma, 2002) to simulate the words being **spoken by a male talker** and **a female talker** (Luthra et al., 2021).

These words were split in half to **create two sets of words representing two talkers**: 10 unique S words and 10 unique ʃ words were allocated to each talker (*see below*).

Talker A			Talker B	
ʃ	**S**		**ʃ**	**S**
Ambition	Pregnancy		Initial	Parasite
Machinery	Democracy		Beneficial	Obscene
Brochure	Embassy		Neogtiate	Medicine
Official	Legacy		Commercial	Tennessee
Crucial	Reconcile		Parachute	Peninsula
Pediatrician	Personal		Efficient	Hallucinate
Flourishing	Eraser		Publisher	Arkansas
Reassure	Episode		Glacier	Compensate
Graduation	Literacy		Refreshing	Dinosaur
Vacation	Coliseum		Impatient	Rehersal

Each experiment will have a male and a female talker:
If Talker A is female, then Talker B is male.
If Talker A is male, then Talker B is female.

Filler Trials

Each experiment consist **of 80 total exposure trials**, including **20 critical** trials and **60 filler** trials. During filler trials, one talker will say a word, and the other talker will say a nonword.

Paradigm

Participants will be instructed to attend to either the female talker or the male talker. They will then perform a series of 2-option forced-choice lexical decision tasks where they hear a recording and then select on their screen if this talker said a word or a nonword (*see below*).

🔊 Listen → Word 🟢 Nonword
Select

Talker A and Talker B recordings were paired to create Materials A and Materials B. Half of the participants will hear the words in Materials A with the simulated accent (?S, ?Sh), and the words in Materials B without the accent (S, Sh). The other half of the participants will hear the inverse, meaning Material B will be accented and Materials A will not be *(See below)*.

The word pairings shown horizontally across in Materials A & B were then spliced together to **create stereo audio files** where one talker is played in the left ear, and the other in the right. Like talker gender, ear assignment was counterbalanced across participants.

Materials A

Talker A		
Ambition		Parasite
Machinery		Obscene
Brochure		Medicine
Official		Tennessee
Crucial		Peninsula
Pediatrician		Hallucinate
Flourishing		Arkansas
Reassure		Compensate
Graduation		Dinosaur
Vacation		Rehersal

Materials B

Talker B		
Pregnancy		Initial
Democracy		Beneficial
Embassy		Neogtiate
Legacy		Commercial
Reconcile		Parachute
Personal		Efficient
Eraser		Publisher
Episode		Glacier
Literacy		Refreshing
Coliseum		Impatient

Test Phase

After the Exposure Phase, participants will hear the asi-ashi test continuum across trials in both talkers' voices. Each trials will only play a recording from a single voice at once. Participants will select if the audio they heard was "asi" or "ashi" (*see left*) for each trial, to produce results like the predictions, shown to the left.

Asi 🟢 **Ashi**
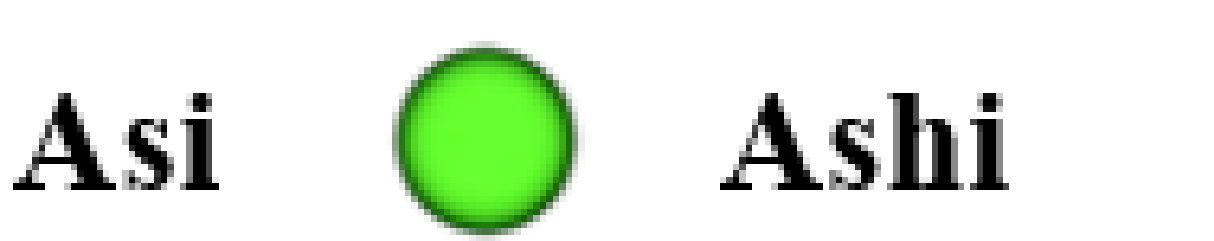
References

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

Cummings, S. N. & Theodore, R. M. (in press). Perceptual learning of multiple talkers: Detriments, characteristics, and limitations. *Attention, Perception, & Psychophysics*.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal?. *Cognitive psychology*, *51*(2), 141-178.

Luthra, S., Mechtenberg, H., & Myers, E. B. (2021). Perceptual learning of multiple talkers requires additional exposure. *Attention, Perception, & Psychophysics*, *83*, 2217–2228.

Tzeng, C. Y., Nygaard, L. C., & Theodore, R. M. (2021). A second chance for a first impression: Sensitivity to cumulative input statistics for lexically guided perceptual learning. *Psychonomic Bulletin & Review, 28*, 1003–1014.

Acknowledgements