**Introduction**

Despite spoken language being highly variable, listeners tend to understand most newly encountered talkers when hearing them speak for the first time. Variation in speech production presents a unique challenge for cognitive processing that is solved seemingly automatically: Our brains learn how talkers speak, and then apply this information to construct expectations about the speech they will encounter in the future [K&J 2015]. This is a lower-level cognitive process that often occurs without the listener noticing []. However, this also brings into question the potential limitations of this ability: Are listeners always passively sponging information about talker’s speech production from their environment, or must they be intentionally directing their attention towards a given talker? The goal of this experiment is to investigate the automaticity of speech perception adaptation; Specifically, how directing attention to one talker competes with adaption to a second talker when both are speaking simultaneously.

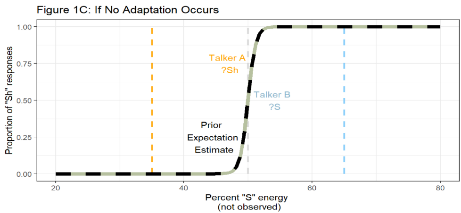
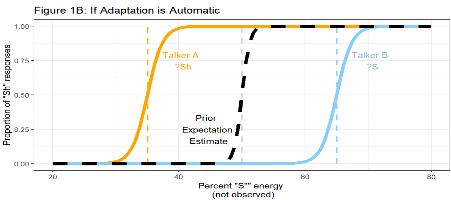
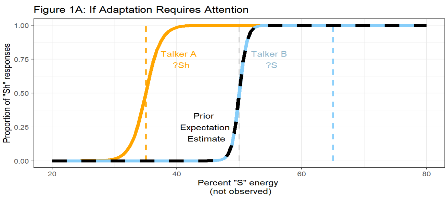
**Background**

While a large body of research has found that perceptual learning is not inhibited by distractions [], lack of intention [], or exposure to multiple talkers [Luthra et al., 2021], other research has found that listeners do consider contextual factors and causality (e.g., accommodating a talker visibly chewing while talking) when learning how talkers speak [Shawn’s paper]. The implications of the latter suggests that listeners do consider the context of an utterance when formulating expectations of how a talker produces speech sounds. In a similar fashion, we also see that listeners utilize stereotypes [] and labeling to adapting their expectations as a form of top-down processing []. Both processes indicate that listeners may store and use information based on its perceived relevance.

This assumption is echoed in the findings of Dr. Samuel’s 2016 paper. Rather than focusing on how talkers’ speech may compete in noisy environments, Samuel explores the dependency of adaptation on lexical processing. Participants were exposed to two simulated talkers, where one talker produced speech with a phonetically shifted s/sh and was always presented before the 2nd talker. The 2nd talker only produced filler words or filler nonwords. Samuel found that when the onset of the 2nd talker was presented <200ms after the 1st talker, the participant did not exhibit adaptation, suggesting that listeners need time to lexically process the stimulus in order to recalibrate their perception of the s/sh speech sound [Samuel 2016]. While this paradigm was used to investigate the length of the lexical processing window in word recognition, the implications for this proposed experiment support that a talker must be attended to for perceptual adaptation to occur.

**Specific Aims/Hypotheses**

We hypothesize that perceptual adaptation to speech is contingent on a listener’s attention being directed towards a given talker. In this experiment, we therefore aim to simulate two distinct talkers that a listener will hear speak simultaneously and compare participants’ adaptation to both. The participant will be given a task, where the correct response is contingent on the participant attending to only one of the talkers’ speech—e.g., a task where the participant is instructed to select if the female talker said a word or a nonword. We will then test if participant’s exhibit perceptual adaptation for both the attended talker and the unattended talker. If perceptual adaptation requires a listener to be attending to a talker, then we would expect a listener to only adapt to the attended talker and **not** the unattended talker *(figure 1a)*. If perceptual adaptation is not contingent on a listener paying attention to a given talker, then we would expect participants to exhibit adaptation for both talkers in the experiment *(figure 1b)*.



**Significance**

* Implications about the level of processing in adaptation at the phonemic level
* Ecologically validity in noisy (crowded) contexts
* Possibility of listener’s prioritizing adaptation when there is a need to understand speech
* **Paradigm would allow for testing listener preference factors related to talker perceived relevance**
* Language learning implications
* Social cohesion implications & role of identity on PR

**Methods**

In this study, we will be measuring participants’ perceptual adaptation to two simulated talkers’ s/sh production. s/sh sounds exist on a continuum, spanning from “s” as in “Sock” to “sh” as in “Shock.” Earlier research suggests that listeners’ adaptation to s/sh production is talker-specific, meaning that listeners adjust their perceived boundary between s/sh for each talker [K&S, 2005]. This quality prevents cross-talker contamination in adaptation, allowing us to simulate two distinct talkers with different s/sh productions during the same experimental exposure[Shawn and Dr. Theodore’s paper]*.*

Our stimuli were adapted from those developed in Kraljic & Samuel, 2005 []. Critical items included two versions a word that contained an s/sh sound: one with the typical s/sh production, and one where the s/sh sound was replaced by an ambiguous s/sh sound (e.g., “dino**s**aur” changed to “dino**sh**aur,” or “publi**sh**er” becoming “publi**ss**er”). Filler items did not contain any s/sh sounds and were either nonwords that followed the typical pattern of real words or real words. We then used the audio processing software Praat [] to change the gender of the voice [Luthra et al., 2021], and to pair both the critical items and the filler items into stereo files where one talker is heard in the left ear of a headset and the other in the right. This resulted in a set of audio files that simulates two distinct talkers: a female talker and a male talker in opposite spatial positions on either side of the participant. These factors will be counterbalanced across participants (*figure ?*).

Diagram

Description automatically generatedIn each experiment, participants first experience an exposure phase that will consist of 80 randomized trials. 20 of these trials will be critical trials, and 60 trials will be filler trials. Participants will be instructed to listen to either the female talker or the male talker, and to then perform a lexical recognition task *(figure 2).* Participants will then be exposed to a 72-trial test phase. The participant will hear both talkers independently produce a six-increment s/sh test continuum 6 times, during which they will perform 2FC lexical decision tasks to determine the categorization boundary for both talkers. We will then compare responses for both talkers within-subject. If there are limits to the automaticity of speech perception, then we can expect listeners will adapt their perceived categorical boundary to align better with the attended talker’s speech compared to their adjustment for the unattended talker. Conversely, complete adaptation to both talkers could suggest that speech perception adaptation is automatically shaped by any speech in a listener’s environment.

*Figure 2:*

**Roles**

* Florian is going to be inserting my lists into a java code to run on prolific
* Will probably need some additional help with analyses
* And revising the paper

**Timeline**

I began developing this project over this past summer, thanks to the support of the Wiesman Fellowship. Since that time, I have crafted my question, developed a methodological design, and created my stimuli. I am currently finishing up creating lists that will be used to call the stimuli in the web-based experiment. By the beginning of the spring, I aim to have run the initial pilot experiment. From that point, my target timeline is below:

* What has been done so far (since the summer)
* Run the pilot version (first 4 lists) when?
  + **Note that this should be enough data to report on even if we do run the next two lists**
  + Leave Exp. 2 out for now?
* Analyses 🡪 Early March
* **Write-up 🡪 Friday, April 14th**
  + Break down deadlines for sections in the write-up

**References**

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

limitations. *Attention, Perception, & Psychophysics*.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and

adapt to the novel. *Psychological Review*, 122(2), 148–203. <https://doi.org/10.1037/a0038695>

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

141–178. <https://doi.org/10.1016/j.cogpsych.2005.05.001>

Luthra, S., Mechtenberg, H., & Myers, E. B. (2021). Perceptual learning of multiple talkers requires additional exposure.

*Attention, Perception, & Psychophysics*, 83(5), 2217–2228. <https://doi.org/10.3758/s13414-021-02261-w>

Samuel, A. G. (2016). Lexical representations are malleable for about one second: Evidence for the non-automaticity of

perceptual recalibration. *Cognitive Psychology*, 88, 88–114. <https://doi.org/10.1016/j.cogpsych.2016.06.007>

Zhang, X., & Samuel, A. G. (2014). Perceptual learning of speech under optimal and adverse conditions. *Journal of*

*Experimental Psychology: Human Perception and Performance*, 40(1), 200–217. <https://doi.org/10.1037/a0033182>