Exploring the Automaticity of Speech Perception and Adaptation

Rachel Sabatello, Shawn Cummings, & Florian Jaeger

Despite being one of the most prevalent forms of human communication, spoken language is highly variable. Even talkers with similar language backgrounds tend to differ in how they produce speech sounds, for example how they may distinguish /s/ (the "Ss" sound) from /ʃ/ (the "Sh" sound). Yet listeners can often understand newly encountered talkers when hearing them speak for the very first time. Variation in speech 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 speech they encounter in the future. This process often occurs without the listener even noticing.

However, the degree of automaticity in speech perception adaptation remains unclear. While a large body of research has found that perceptual learning is not inhibited by distractions, lack of intention, or exposure to multiple talkers, 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. Furthermore, the utility of speech perception adaptation may also pose the possibility of listener's prioritizing adaptation when there is a need to understand speech, versus passively sponging information from background noise in their environment.

In this experiment, we will investigate the role of attention in speech perception and adaptation by limiting the participant's available attentional resources. To achieve this, we will expose participants to two simulated talkers simultaneously. These talkers have been engineered to have distinct voices, and also produce inversely atypical "accents" on the S-ʃ continuum in their speech (i.e., one talker produces their /s/ sounds more like "Sh", and the other talker produces their / ʃ/ sound smore like "Ss"). Participants will be instructed to attend to one of the two talkers throughout the experiment, and then select if that talker is saying a word or a nonword in a series of virtual forced-choice lexical decision tasks.

We will then test the effects of directing the participants' attention to *one talker* on the participants' ability to adapt their speech perception to *both talkers*. We can measure how participants categorize sounds on the S-ʃ spectrum using a set of categorization tasks that feature a S-ʃ test continuum, and then compare how participants respond to these tasks for each of the two simulated talkers. 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. No matter which of these theories is supported by our results, this study has the potential to pioneer future paradigms for studying speech perception adaptation, and develops a foundation for further research on how human cognition prioritizes speech processing.