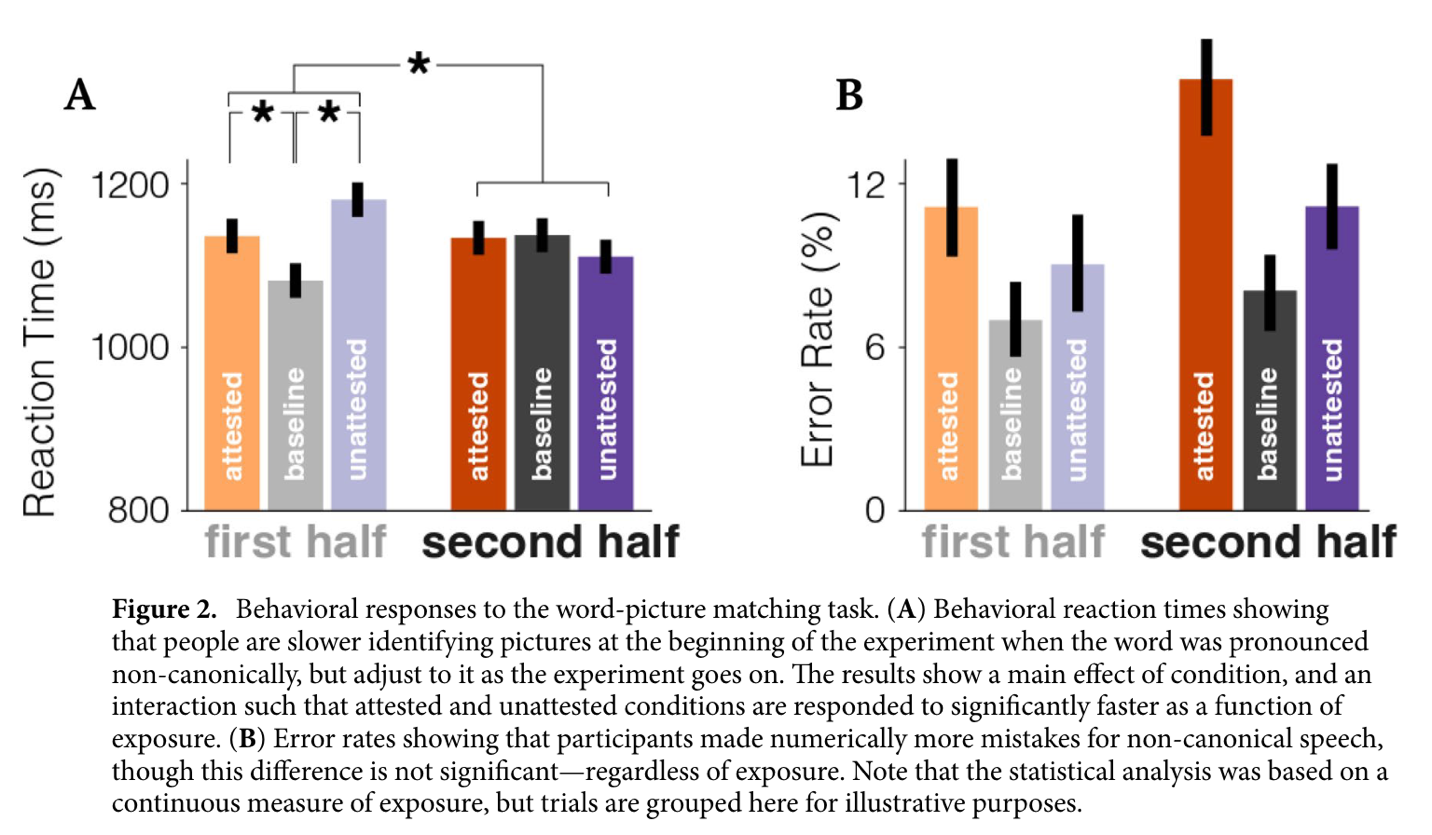
**Recent work that I might want to cite:**

1) Blanco-Elorrieta, E., Gwilliams, L., Marantz, A. *et al.* Adaptation to mis-pronounced speech: evidence for a prefrontal-cortex repair mechanism. *Sci Rep* **11,**97 (2021). https://doi.org/10.1038/s41598-020-79640-0

* Neurobiological basis of attunment
* MEG
* 24 native English participants listened to words spoken by a “canonical” American speaker and 2 non‑canonical speakers, and performed a word‑picture matching task
* Non‑canonical speech was created by including systematic phonological substitutions within the word (e.g. [s] → [sh]).
* Activity in the auditory cortex (superior temporal gyrus) was greater in response to substituted phonemes, and, critically, this was not attenuated by exposure.
* By contrast, prefrontal regions showed an interaction between the presence of a substitution and the amount of exposure: activity decreased for canonical speech over time, whereas responses to non‑canonical speech remained consistently elevated.
* Grainger causality analyses further revealed that prefrontal responses serve to modulate activity in auditory regions, suggesting the recruitment of top‑down processing to decode non‑canonical pronunciations. In sum, our results suggest that the behavioral deficit in processing mispronounced phonemes may be due to a disruption to the typical exchange of information between the prefrontal and auditory cortices as observed for canonical speech



* But note that the stimuli here include a full-scale “phoneme switch” rather than acoustic/phonetic shifts. So it’s entirely possible that the task biased the regions that can be activated.

2) Guediche, S., Holt, L. L., Laurent, P., Lim, S. J., & Fiez, J. A. (2015). Evidence for Cerebellar Contributions to Adaptive Plasticity in Speech Perception. *Cerebral cortex (New York, N.Y. : 1991)*, *25*(7), 1867–1877. https://doi.org/10.1093/cercor/bht428

* Any involvement of cerebrum?
* The fMRI study investigated whether adaptation in speech perception also involves the cerebellum. Acoustic stimuli were distorted using a vocoding plus spectral-shift manipulation and presented in a word recognition task.
* the right Crus I region of the cerebellum in adaptive changes in speech perception.
* The results provided evidence of a functional network between the cerebellum and language-related regions in the temporal and parietal lobes of the cerebral cortex.
* “Since adaptive changes in speech perception generalize to new items, it is thought that the locus of adaptive change must be relatively early within the speech processing pathway”.
* The temporal area that emerged in our functional correlation analysis may be a target area that represents sensory prediction error signals. This interpretation is based on neurobiological models of speech perception, which typically posit engagement of primary auditory cortex and a belt of surrounding auditory association areas located along the superior temporal gyrus in prelexical speech processing (Rauschecker and Tian 2000; Rauschecker and Scott 2009;Okada et al. 2010;Peelle et al.2010). Consistent with this interpretation, recent studies have shown modulation of activity in the superior temporal cortex as a function of predictive contexts (e.g., Davis 2011;Sohogluet al. 2012;Wild et al. 2012) as well as the predictability of a sensory consequence associated with motor planning during speech production (Chang et al. 2013)

Timeline

Description automatically generated with medium confidence

Chart

Description automatically generated with low confidence

Introduction   
  
[53] We do not, at this stage, test the framework against novel experiments but rather aim to demonstrate why such a framework is critical in moving the field forward, how it works, and what it can offer.

🡪 “More specifically, here we demonstrate how the framework works, what it offers, and why a computational approach like this is needed. In so doing, we discuss how to design a decisive behavioral experiment that can delineate contributions of the underlying mechanisms.”

[66] Adaptive adjustment in response to systematic phonetic variation–even that present in starkly different pronunciations from accented talkers—results in faster processing and more accurate recognition [ref].

🡪 I am not sure what you mean to say?

[71] In short, there is now a rich literature on the effects of recent exposure on subsequent

perception of speech from the same talker.1 Listeners’ ability to adapt based on recent input is

now considered a central part of human speech perception, and acknowledged by all major

theories of speech perception.

🡪 Can go?

[80] As illustrated in Figure 2, speech perception is generally thought to begin with (A) the extraction and normalization of acoustic/phonetic cues.

🡪 “While different in important details, existing theories in speech perception share general assumptions about how the acoustic input supports perception of a speech category (Figure 2). The process generally begins with (A)….”

[84] While it is uncontroversial that all these processes occur during speech perception, We do not know whether and how these three mechanisms interact to influence adaptive changes in speech perception.

🡪 We talked about this as we finalized our SA for the R01 proposal

🡪 It is a bit risky to assume that all of them are at work because a part of our core argument is that we do not even have solid evidence that each of them actually matters

🡪 “While there is now substantial work on *each* of the mechanisms, little is known about whether, and if so how, they may interact or trade off with one another. This absence of comprehensive knowledge, we argue, is largely because the vast majority of empirical investigations have tested one mechanism at a time. A very few have directly contrasted predictions of two mechanisms; None that we know contrasts all three. Results are consequentially interpreted as evidence in support of a particular mechanism without explicit rejection of alternative explanations. Resultantly, as we detail below, there is currently no strong evidence that any of the mechanisms is in fact needed to explain existing behavioral results in the literature.”

[92] The majority of research on talker-related adaptation has focused on the middle layer,

changes in linguistic representations.

🡪 “This empirical indeterminacy has far-reaching theoretical consequences. Most critically, a large body of the recent work on talker-dependent adaptation has focused on (B) in Fig.2, i.e., changes in linguistic representations.”

[117] There are, however, alternative explanations that have been pointed out in separate lines of work.One possibility is that pre-linguistic mechanisms underlie the effects (bottom of Figure 2). Low-level, automatic (involuntary) normalization processes might transform the speech input during the early stages of auditory processing.

🡪 “The absence of contrastive tests, however, means that the same behavioral results could in principle be explained by an alternative mechanism that *assumes no change of linguistic representations*. For instance, if is possible that adaptive changes of responses can be due to low-level, automatic (involuntary) normalization processes might transform the speech input the early stages of auditory processing.”

[148] Recent work using neuroimaging techniques suggests that multiple mechanisms might be at play and in fact interact with each other to support accommodation of talker variation. Auditory and superior temporal cortices associated with acoustic-phonetic mapping have shown sensitivity to exposure to nonstandard speech tokens [ref].

Recent reviews highlight this empirical and theoretical indeterminacy as an important unresolved gap in our understanding of speech perception (e.g., Bent & Baese-Berk, 2021; Schertz & Clare, 2020; Weatherholtz & Jaeger, 2016; Zheng & Samuel, 2020). Identifying the exact mechanisms of adaptive changes perception not only contributes to our basic understanding of speech perception, it is also important for developing treatments for impaired adaptation. Which may arise from multiple auditory, linguistic and cognitive sources.

🡪 “In the neuroimaging literature, in contrast, it is commonplace to separate post-perceptual decision-making processes from lower-level perceptual adjustments, each involving distinct cortical areas and networks [ref]. In fact, a recent work on adaptation to phoneme category substitutions (e.g., /s/ pronounced as /sh/) found evidence that adaptive processing of such “mispronunciations” engages prefrontal regions, responsible for post-perceptual ‘repair” mechanisms (Blanco-Elorrietta et al., 2021): Simply put, the listener seems to initially map the acoustic input onto the wrong category and subsequently corrects the mapping. This was contrasted with the hypothesis that adaptation occurs via recalibration of the acoustic-phonetic mapping in the auditory cortices, to which they found no support in their results.

Across the behavioral and neuroimaging work, the past research has thus split the pie in distinct ways. Behavioral studies have begun to distinguish between the (A) normalization and (B) representational changes while treating (C) post-perceptual decision-making as a nuisance factor. In contrast, the neuro-imaging work tends to group (A) and (B) together as a low-level perceptual process, functionally distinct from (C) i.e., a higher-level, decision-making process. To meaningfully link these lines of work, and to better elucidate how human speech perception operates over the variable acoustic input, we need a framework that can encompass all (A)-(C). More specifically, we need a framework that can isolate when*—in what environment and to what types of inputs—* the three mechanisms may support diverging predictions about human perceptual responses.

[168] Yet, which of the mechanisms in Figure 2….

🡪 This paragraph has a lot of information that is overlapping with what we have already discussed above. (e.g., “The majority of research on…”

🡪 Can we shorten this and say

1. We are not assuming that there is always one winning mechanism
2. Even if we were to assume that the three mechanisms jointly underlie the effect of recent exposure, we would still need to explain “how”
3. Answers to this question will have important implications for our future investigations and theorizing.