

Relationships Between Auditory Discrimination and Adaptation of Vocal Features of Speech

Madeline Lis Knutson

A thesis

submitted in partial fulfillment of the  
requirements for the degree of

Master of Science

University of Washington

2023

Committee:

Tanya Eadie

Mara Kapsner-Smith

Gabriel Cler

Program Authorized to Offer Degree:

Department of Speech and Hearing Sciences

©Copyright 2023

Madeline Lis Knutson

University of Washington

**Abstract**

Relationships Between Auditory Discrimination and Adaptation of Vocal Features of Speech

Madeline Lis Knutson

Chair of the Supervisory Committee:

Tanya Eadie

Department of Speech and Hearing Sciences

The current study evaluated relationships between auditory discrimination and adaptation of vocal features of speech. Previous research suggests relationships between auditory discrimination and adaptation of articulatory features of speech. However, there is conflicting evidence relating auditory discrimination ability and adaptation responses of one specific vocal feature of speech, fundamental frequency ( $f_0$ ). Additionally, there are no published studies evaluating the relationship of auditory discrimination ability and adaptation of intensity. Therefore, the purpose of this study was to investigate relationships between auditory discrimination and adaptation of *vocal* features of speech, including fundamental frequency ( $f_0$ ) and intensity. Twenty typical speakers underwent auditory discrimination tasks for  $f_0$  and intensity. They also participated in speech adaptation experiments for  $f_0$  and intensity. Results indicated no to weak relationship between auditory discrimination and adaptation of  $f_0$  and no to

weak relationship between auditory discrimination and adaptation of intensity. The results therefore support the use and further development of voice-specific models of sensorimotor control for vocal parameters, such as  $f_0$  and intensity, when analyzing auditory-motor control of vocal features of speech. Implications for future research and clinical studies will be discussed.

## Table of Contents

<b>I. Overview</b>	6
<b>II. Models of Sensorimotor Control of Speech</b>	7
<b>III. Auditory Discrimination and Speech Production Goal Regions</b>	11
<b>IV. Auditory Discrimination and Adaptation of F1</b>	12
<b>V. Auditory Discrimination and Adaptation of Vocal Features of Speech</b>	14
Auditory Discrimination and Adaptation of fo	14
<b>VI. Research Questions</b>	19
<b>VII. Methods</b>	20
Participants	20
Hearing Screen	20
Voice Screen	21
Equipment	21
Procedures	22
Descriptive Analysis	24
Statistical Analysis	25
<b>VIII. Results</b>	25
Descriptive Results	25
Relationships between Auditory Discrimination Values and Adaptation Responses	27
<b>IX. Discussion</b>	28
Auditory Discrimination and Adaptation of fo	29
Auditory Discrimination and Adaptation of Intensity	31
<b>X. Additional Research and Clinical Implications</b>	33
<b>XI. Limitations</b>	34

### I. Overview

Fluent speech production by neurologically healthy speakers is a highly skilled motor ability. Changes and updates to speech motor plans are necessary during development, and are also necessary to maintain accurate production and intelligibility throughout the lifespan. These changes rely on an integration of sensory feedback and feedforward control. The feedback

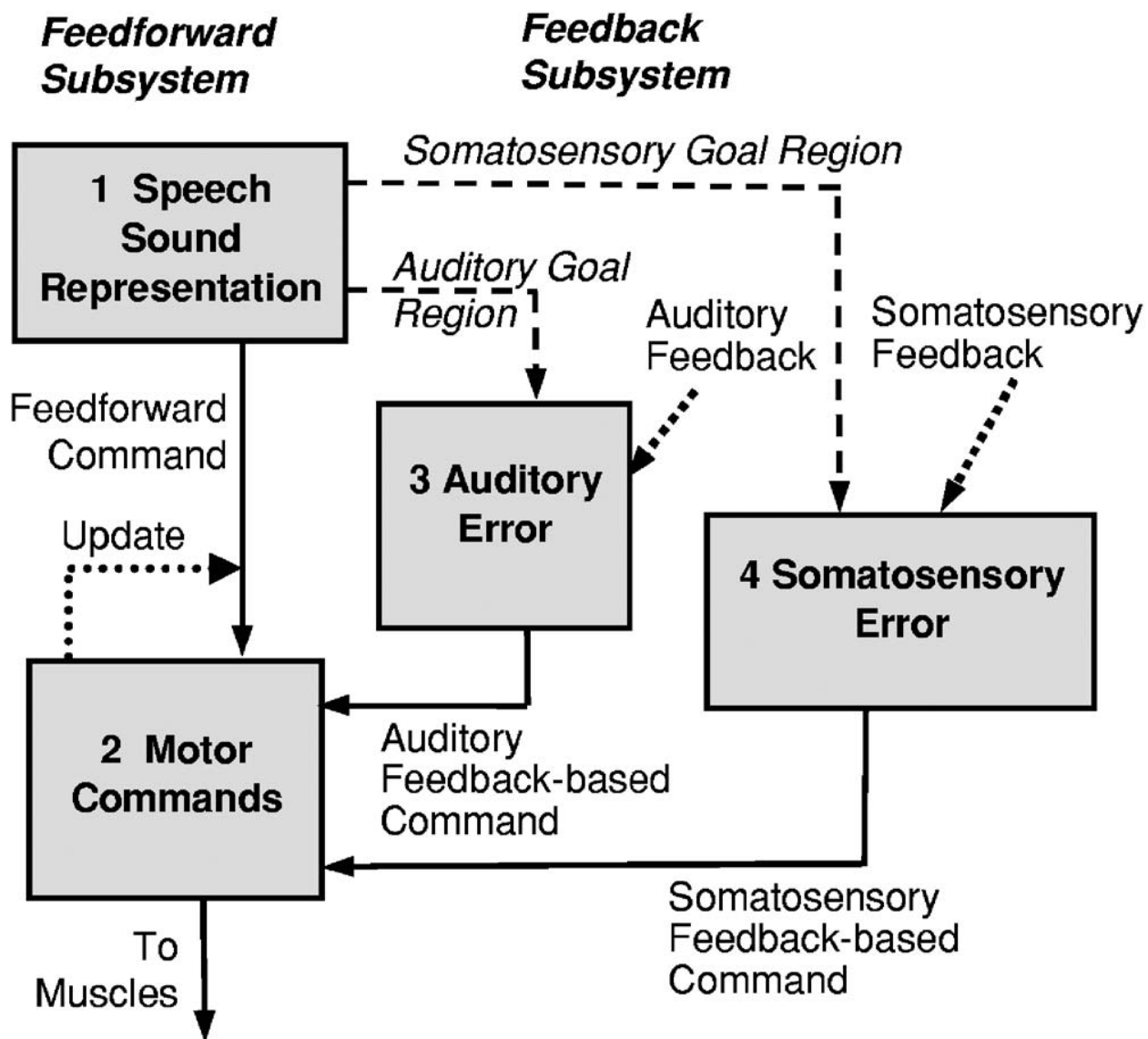
sensorimotor control system utilizes auditory and somatosensory feedback of speech production to monitor and adjust motor plans. The feedforward sensorimotor control system sends motor plans (and expected sensory feedback of speech) to various areas within the central nervous system. Over time, consistent mismatches in actual vs. expected sensory feedback, such as when a speaker detects the difference between the sound they are producing and what the target is, updates the stored motor plans for speech (Perkell, 2013). This process is called adaptation.

It is thought that adaptation, or the changes in the stored motor plans for speech, could be related to auditory discrimination abilities (Perkell, 2013). Auditory discrimination can be defined as the extent of a listener's ability to differentiate auditory stimuli. There is mixed evidence suggesting that auditory discrimination ability is correlated with the strength of an adaptation response when auditory feedback is perturbed. Although there is some evidence that auditory discrimination ability is correlated with strength of adaptation responses of the first formant ( $F_1$ ) (Martin et al., 2018; Villacorta et al., 2007), there is conflicting evidence relating auditory discrimination ability and adaptation responses of fundamental frequency ( $f_0$ ) (Abur et al., 2021; Alemi et al., 2020; Lester-Smith et al., 2020). Additionally, there are no published studies evaluating the relationship of auditory discrimination ability and adaptation of intensity. The aim of the current study was to evaluate the relationship between auditory discrimination of  $f_0$  and adaptation of  $f_0$ , and the relationship between auditory discrimination of intensity and adaptation of intensity in healthy, typical speakers of English. The results will provide further information on sensorimotor control of vocal features of speech. Background information supporting this experiment includes models of sensorimotor integration of speech, auditory discrimination and speech production goal regions, auditory discrimination and adaptation of  $F_1$ , and auditory discrimination and adaptation of  $f_0$ .

## II. Models of Sensorimotor Control of Speech

Numerous models attempt to capture sensorimotor control of speech (Houde & Nagarajan, 2011; Parrell et al., 2019; Tourville & Guenther, 2011). The Directions Into Velocities of Articulators (DIVA) model is an evidence-based neurocomputational model that incorporates the neuroanatomy and physiology involved in producing speech. Two main components of this model include a feedback control system and a feedforward control system. Within this framework, a motor plan is produced, and phonemic goals are sent from the premotor cortex to the sensory cortex to encode expected sensory patterns. In the feedback control system, auditory and somatosensory information of the actual production is compared to the expected sensory plans. Somatosensory information is compared in the ventral supramarginal gyrus (Hickok & Poeppel, 2004) and auditory information is compared along the posterior superior temporal gyrus (Buchsbaum et al., 2001; Hickok & Poeppel, 2004). If there is a mismatch between the actual and expected feedback, an error signal is sent to the right ventral primary motor cortex to immediately adjust the motor plan in a closed loop circuit (Tourville & Guenther, 2011). These adjustments are known as reflex responses.

In contrast, the feedforward control system relies on stored motor plans that include auditory and somatosensory targets to execute desired articulatory movements, independent of auditory or somatosensory feedback. Information from reflex responses, or information from the comparison of actual and expected sensory feedback (i.e., error signals), is thought to update the feedforward control system, particularly when errors occur repeatedly. Changes to stored motor plans within the feedforward system due to feedback-based error signals are known as adaptation (Perkell, 2013). (See Fig. 1).



**Figure 1.** Representation of the DIVA model from Ménard et al. (2007)

Adaptation of speech can be captured through experiments measuring changes to speech over time in response to auditory or somatosensory errors (Perkell, 2013). For the purposes of this study, however, we will focus on the auditory domain. Studies measuring adaptation are typically performed in several stages. Baseline measurements of speech production (e.g., measuring specific acoustic parameters of habitual speech production) are typically taken of a participant's speech. Next, during a ramp phase, auditory or somatosensory feedback of the

participants is gradually altered in near-real time across many trials of speech tasks, and participants typically demonstrate a compensatory response, or a change in speech behavior due to perturbation of sensory feedback. It is important to note that some studies do not have a ramp phase. During a hold phase, the maximum perturbation of auditory or somatosensory feedback is held over a number of trials to induce a compensatory response and changes to motor plans over time. A typical compensatory response includes speech behavior that opposes the perturbation. Atypical compensatory responses can include following responses, i.e., speech behavior that follows the direction of the perturbation, or inconsistent responses that do not have a clear pattern of direction (Alemi et al., 2020). After perturbation conditions are removed, researchers measure any lasting compensatory responses remaining in the participant's speech to determine if their feedforward motor commands have been updated, thus demonstrating an adaptation response. It is important to note that when auditory information is perturbed, somatosensory feedback remains unperturbed. A mismatch in auditory feedback and somatosensory feedback may partially explain why participants do not fully compensate to auditory feedback, although there is wide interspeaker variation (Perkell, 2013). In the area of speech motor control, studies have often focused on adaptation of perturbations of  $f_0$ , the first formant ( $F_1$ ) of speech, and the second formant ( $F_2$ ) of speech. These represent responses in the control of the larynx as well as other articulators such as the tongue and jaw, respectively.

Reflexive responses can be measured experimentally through unpredictable perturbation of sensory feedback while a participant performs speech tasks. These perturbations are performed on a limited number of trials, usually spaced apart, to prevent adaptation of speech motor plans due to expected exposure to the altered condition. Most speakers make quick changes to their speech opposing the perturbation within the same utterance. The changes in a

participant's speech due to the sudden perturbation of their feedback is known as a reflexive response (Perkell, 2013). These responses are usually measured as the percentage of a perturbation that is corrected for in a trial. Thus, these measurements help quantify the use of auditory and somatosensory feedback to compensate for perturbations (Kearney et al., 2020).

Several neurocomputational models have been developed based on the basic framework of the DIVA model for adaptation and reflexive responses. SimpleDIVA, a mathematical model developed by Kearney et al. (2020), focuses on adaptation of speech motor control. It distills the classic DIVA model to focus on three parameters necessary for adaptation of speech: auditory feedback control, somatosensory feedback control, and feedforward control. SimpleDIVA is a mathematically validated model that has been tested through computer simulations and compared to previously recorded data from auditory-motor adaptation studies of  $f_0$ ,  $F_1$ , and  $F_2$ . It can model auditory feedback control and sensory feedback control gains, or a theoretical quantitative measurement of a speaker's use of auditory and somatosensory feedback to compensate for a perturbation. This model can also estimate feedforward adaptation learning rates (Kearney et al., 2020), or how quickly the feedforward motor plan is updated due to auditory and somatosensory feedback. Kearney et. al (2020) states that this model will likely be used in the future to estimate differences of sensorimotor integration of typical speakers and individuals with communication disorders.

Computational modeling of speech adaptation has traditionally focused on articulatory features of speech (e.g.,  $F_1$ , and  $F_2$ ) and has neglected vocal features of speech other than  $f_0$ , such as intensity. Articulatory features of speech provide segmental information which convey phonemic differences and are crucial for intelligibility. In contrast, in non-tonal languages, vocal

features of speech, such as  $f_0$  and intensity, provide suprasegmental cues that convey paralinguistic information (Anderson, 2018).

LaDIVA, also based on the DIVA model, is the only neurocomputational model that incorporates neural control of speech and physiologically and biomechanically relevant aspects of *laryngeal* control for phonation. This model accurately predicts reflexive and adaptation responses to  $f_0$  when auditory feedback is perturbed, compared to previously collected behavioral data (Weerathunge et al., 2022). LaDIVA is significant for the future of voice research, as the authors intend to use it to model neural motor control of phonation. It may also be used to help conceptualize the results of the present study, which further examines vocal features of speech and how they may or may not relate to auditory discrimination.

### III. Auditory Discrimination and Speech Production Goal Regions

One theory of auditory sensorimotor learning is that speech production variability and the strength of responses to auditory perturbation are correlated with auditory discrimination abilities (Perkell, 2013). Auditory discrimination can be defined as the extent of a listener's ability to differentiate auditory stimuli. It is often captured as just noticeable difference (JND), a quantitative measurement of the lowest value required for participants to differentiate between slightly different sounds. JND values are specific for various acoustic parameters such as  $F_1$ ,  $f_0$  and intensity. Evidence has suggested that those with more acute auditory discrimination have more distinct speech goal regions, i.e., a more narrowly defined range of acceptable acoustic outcomes, and therefore produce speech sounds with increased contrast (Perkell, 2012). For example, Matthies et al. (2008) has shown that speakers who produce the most contrast between /s/ and /ʃ/ have better than average auditory discrimination scores of these sibilants. In another study, Perkell et al. (2004) similarly discovered that participants with better auditory

discrimination of vowels and had increased contrast in the production of vowels. These studies provide evidence of a relationship between auditory discrimination and speech production for articulatory features of speech. This relationship between auditory discrimination and speech production targets also suggests that an individual's ability to discriminate acoustic parameters of speech may relate to the magnitude of their response to perturbations of auditory feedback.

#### IV. Auditory Discrimination and Adaptation of $F_1$

Many experiments on sensorimotor integration of speech have examined the articulatory feature  $F_1$  in speech tasks. Several have specifically investigated the relationship between auditory discrimination and adaptation of  $F_1$  (Lester-Smith et al., 2020; Martin et al., 2018; Villacorta et al., 2007). Villacorta and colleagues (2007) analyzed this relationship in a study including 20 American English-speaking participants. They measured adaptation responses to auditory feedback perturbation of  $F_1$  during speech in an adaptation experiment. They found that compensatory responses to the auditory perturbation of  $F_1$  were maintained when auditory feedback was later masked by noise, indicating change in motor speech plans. In a separate task, the researchers measured auditory discrimination of  $F_1$  in synthetically produced voices. They found that individuals with better auditory discrimination abilities produced larger adaptation responses to auditory perturbation of  $F_1$ . They theorized that individuals with better auditory discrimination abilities have smaller auditory goal regions of  $F_1$ , or a range in which the vowel produced is deemed acceptable by their feedback control system, and no error signal is produced. They posited that responses outside of auditory goal regions required stronger correctional responses, and this would thereby lead to a greater degree of adaptation.

Similarly, Lester-Smith et al. (2020) conducted a study to assess relationships between auditory discrimination, reflex responses and adaptation responses to auditory perturbation of

$F_1$ . The study included 20 participants who were speakers of American English without speech disorders. Passive auditory discrimination measures (i.e. the participant listens to stimuli and does not speak) of  $F_1$  were determined by speakers selecting differences in pre-recorded versions of their own speech. In contrast, active auditory discrimination measures (i.e. the participant listens for differences in production while speaking during the task) of  $F_1$  were collected while the speaker's auditory feedback was perturbed online. In an adaptation experiment, auditory feedback of  $F_1$  was perturbed to a maximum increase of 30% relative to the produced signal during the hold phase. Additionally, the researchers measured reflex responses to sudden and unexpected perturbations to auditory feedback (an increase of 30% of  $F_1$ ). Relationships were analyzed between the reflex responses, adaptation responses, active auditory discrimination measures and passive auditory discrimination measures. In contrast to Villacorta et al. (2007), they found that adaptive responses to  $F_1$  were not correlated with either auditory discrimination measure. They also found no relationship between reflex and adaptation responses to  $F_1$  perturbation. However, they did find that speakers with better passive auditory discrimination of  $F_1$  produced larger reflex responses to sudden  $F_1$  perturbation. These findings could suggest a relationship between auditory discrimination and the size of the auditory goal region of  $F_1$ , assuming the reflex response is caused by an error outside of the goal region. However, the study does not support the idea that auditory discrimination of  $F_1$  influences sensorimotor learning of speech. This is due to the lack of relationship between auditory discrimination and adaptation of  $F_1$ . In contrast to previous studies, these results could imply that auditory discrimination of  $F_1$  does not have a significant role in motor learning.

Another study examined the role of auditory discrimination of vocal features of speech (i.e.  $f_0$  and intensity) in auditory sensorimotor adaptation of  $F_1$ . Martin et al. (2018) aimed to

determine if auditory discrimination of these vocal features of speech, as opposed to articulatory features, are correlated with differences in speech motor control of  $F_1$ , an articulatory feature of speech. The study included 36 native Spanish speakers without speech disorders. In contrast to previous studies assessing auditory discrimination of  $F_1$ , the first task measured auditory discrimination levels of  $f_0$  and intensity using non-speech stimuli. Next, the participants engaged in an adaptation experiment, and the researchers gradually decreased each speaker's auditory feedback of  $F_1$  until it reached -150 Hz below their production in the hold phase. They found that those with better auditory discrimination measurements of  $f_0$  and intensity showed larger adaptation responses to  $F_1$  perturbations (Martin et al., 2018). The results from this study support the idea that auditory discrimination of  $f_0$  and intensity could play a role in adaptation of articulatory features of speech.

Across studies relating auditory discrimination and auditory sensorimotor adaptation of  $F_1$ , results are mixed. There is some evidence supporting the idea that auditory discrimination of  $F_1$ ,  $f_0$  and intensity could impact motor learning and adaptation of articulatory measures, such as  $F_1$  (Martin et al., 2018; Villacorta et al., 2007). However, other results do not find a relationship between auditory discrimination of  $F_1$  and adaptation of  $F_1$  (Lester-Smith et al., 2020). More research is needed to determine the relationship between auditory discrimination and adaptation of articulatory features of speech. How these results might also extend to vocal features warrants investigation.

## V. Auditory Discrimination and Adaptation of Vocal Features of Speech

### **Auditory Discrimination and Adaptation of $f_0$**

Several studies have investigated the relationship between auditory discrimination and auditory sensorimotor adaptation of vocal features of speech. Some have specifically examined

the relationship between auditory discrimination and  $f_0$  sensorimotor control. For example, Lester-Smith et al. (2020) assessed relationships between active auditory discrimination, passive auditory discrimination, and reflex and adaptation responses of the vocal parameter of  $f_0$  within the same study investigating these relationships of the articulatory parameter of  $F_1$ . Participants engaged in both active and passive auditory discrimination tasks of  $f_0$ , as well as an adaptation experiment with a maximum perturbation of +100 cents of  $f_0$  relative to the participants' speech production. To measure reflex responses, the participant's auditory signal of  $f_0$  was increased unexpectedly by 100 cents. The researchers found no relationship between reflex and adaptation responses to  $f_0$ . Additionally, reflex and adaptation responses to  $f_0$  were not associated with either active or passive auditory acuity. These results contrast with their findings that speakers with better passive  $F_1$  auditory acuity often produced larger reflex responses to sudden  $F_1$  perturbation, indicating that there are potential differences between underlying mechanisms of auditory-motor control of  $F_1$  and  $f_0$ .

Some authors have hypothesized that a failure to find a relationship between auditory discrimination and adaptation of  $f_0$  may have been due to the insensitivity of the methods. Some have theorized that tailoring the degree of perturbation relative to each participant's JND value might be a more sensitive approach for examining potential relationships. In contrast to Lester-Smith et al. (2020) who altered all participants' auditory feedback by the same amount, Alemi et al. (2020) sought to determine if participants adapt in a more homogenous manner when the magnitude of adaptation is personalized to each participant's auditory sensitivity. This was based on the theory that individuals with more sensitive auditory acuity abilities would not recognize excessively perturbed auditory feedback as their own voice. Sixty speakers without voice disorders participated in this study. The first condition was an adaptation experiment with a

maximum perturbation of 100 cents relative to the speaker's production. This condition served to replicate a standard adaptation study. The second condition was an adaptation experiment in which the researchers gradually increased or decreased each participant's auditory signal of  $f_0$  to a maximum of their JND value of  $f_0$ . This condition served to reduce the likelihood that participants would not recognize perturbed auditory feedback as their own. In both experiments, on average participants demonstrated significant opposing responses to perturbation during the ramp and hold phases. However, the participants' auditory discrimination of  $f_0$  did not correlate with adaptation response magnitude or direction of  $f_0$  for either experiment. These results suggest that the amount of adaptation to an  $f_0$  perturbation is not dependent on an individual's sensitivity to  $f_0$ .

The relationship between  $f_0$  discrimination and adaptation has also been studied in people with and without voice disorders. These types of studies are important because it is thought that one contributing factor to some types of voice disorders may be poor auditory discrimination and atypical sensorimotor control. Abur et al. (2021) conducted a study with a sizable sample of 124 individuals (62 with and 62 without hyperfunctional voice disorders). That is, the study included a control group that was age- and sex-matched to the group with hyperfunctional voice disorders. Similar to Lester-Smith et al. (2020), the participants completed an adaptation experiment shifted to a maximum of 100 cents relative to their production. To measure reflex responses in a separate task, the participant's auditory signal of  $f_0$  was increased unexpectedly by 100 cents. Auditory discrimination of  $f_0$  was measured in a JND task. Compared to the control group, individuals with hyperfunctional voice disorders showed worse auditory discrimination of  $f_0$ , similar reflex responses, and a greater frequency of atypical adaptation responses. Unlike findings from Alemi et al. (2020), high JND values were correlated with atypical adaptation

responses. This suggests a relationship between  $f_0$  discrimination and  $f_0$  adaptation specifically for individuals with atypical adaptive behavior.

However, important differences in data analysis exist between the two studies. Alemi et al. (2020) investigated statistical relationships between auditory discrimination of  $f_0$  and amount of adaptation during each experiment of all participants and found no significant correlations. Alternatively, Abur et al. (2021) used z-scores to classify participants' adaptation responses as typical or atypical. Participants with adaptation that deviated outside of the 90th percentile (relative to the control group mean) were classified as atypical. Although previous research has suggested no relationship between auditory discrimination and adaptation of  $f_0$  (Lester-Smith et al., 2020; Alemi et al., 2020), this variation in data analysis provided novel evidence of auditory discrimination, and perhaps sensorimotor, differences of individuals with atypical adaptation responses to  $f_0$  (Abur et al., 2021).

### **Auditory Discrimination and Adaptation of Intensity**

A number of studies also have specifically examined sensorimotor control of intensity during speech. One branch of literature explores changes in intensity of speech with sidetone amplification. Sidetone amplification can be defined as increasing the intensity of a speaker's auditory feedback during speech tasks, such as conversation. This typically results in a reduction of vocal intensity (H. Lane & Tranel, 1971; Lombard, 1911; Tomassi et al., 2021). This is believed to be due to sidetone regulation, or the regulation of vocal intensity to meet conversational demands. When signal-to-noise ratio of conversational speech is altered, speakers typically adjust their vocal intensity to improve intelligibility for their listener (Chang-Yit et al., 1975; H. L. Lane et al., 1961; H. Lane & Tranel, 1971). Side-tone amplification studies provide

evidence of integration of auditory feedback of intensity in speech production seen in changes to the intensity of speech that oppose changes in the intensity of auditory feedback.

Other literature has investigated not only the regulation of intensity to increase intelligibility, but to also convey prosodic demands of a speech task (Hilger et al., 2022). Hilger and colleagues (2022) aimed to determine if reflexive responses to intensity perturbations are universal or task-dependent. They examined reflex responses of intensity in sustained vowel and sentence tasks when perturbing auditory feedback by 3 and 6 dB SPL relative to the speaker's vocal signal. They found a significantly larger magnitude and faster rate of reflexive response of intensity in sentences as compared to vowels (Hilger et al., 2022). These data indicate that the vocal parameter of intensity responds flexibly to auditory feedback to meet the prosodic demands of a variety of speech tasks.

Literature on sidetone amplification and reflex responses to changes of intensity provide evidence that adaptation of intensity can occur. As flexibility of intensity in these studies has been shown to meet prosodic demands of speech, some researchers have explored whether vocal parameters of speech can be used interchangeably to convey meaning. The theory of cue-trading suggests that vocal features of speech can be “traded” or substituted to create desired emphatic meaning (Howell, 1993, Lieberman, 1960). In an adaptation study, with perturbation of  $f_0$  during a task in which participants were instructed to put emphasis on either the first or second word in a sentence, Patel et al. (2011) found that participants opposed the perturbation. They used both  $f_0$  and intensity as a compensatory strategy to alter emphatic stress. These data support the model of integrated vocal control, that the vocal features of speech  $f_0$  and intensity are controlled jointly in an integrated channel (Patel et al., 2011). Patel et al. (2015) performed a similar adaptation study, with auditory perturbation of intensity. The researchers found that participants produced

compensatory responses using only intensity, and changes to  $f_0$  did not reach statistical significance. These data provide further evidence of adaptation of intensity. However, contrary to findings from Patel et al. (2011), these findings suggest that intensity and  $f_0$  could be controlled separately during speech production. Additionally, the authors theorize that prosodic modulation could be threshold dependent. In other words, perturbation of a vocal feature of speech might reach a certain threshold before other vocal features of speech are recruited to create desired emphatic stress.

Although evidence of adaptation of intensity exists (Patel et al., 2015; Tomassi et al., 2021) no studies have yet examined the relationship between auditory discrimination of intensity and adaptation response to intensity. If the trading relations theory is correct, broader auditory targets could be permissible for  $f_0$  and intensity as they are suprasegmental features of speech and add emphatic meaning to language, but are not crucial for intelligibility in English speakers. Therefore, unlike articulatory parameters, or segmental features of speech that carry semantic meaning, the amount of adaptation that occurs for these prosodic features of speech may not be related to differences in an individual's auditory discrimination ability. The demonstrated flexibility of  $f_0$  and intensity (Hilger et al., 2022, Patel et al., 2011, Patel et al., 2015) as well as studies showing no evidence of a relationship between auditory discrimination and adaptation of  $f_0$  (Abur et al., 2021; Alemi et al., 2020; Lester-Smith et al., 2020) suggest that there is likely no relationship between auditory discrimination and adaptation of intensity. Further investigation of this topic could provide evidence of differences in sensorimotor control of vocal vs. articulatory features of speech, and whether such potential differences are due to linguistic demands or physiological limitations.

## VI. Research Questions

The purpose of the study is to investigate the relationship between auditory discrimination and adaptation of vocal features of speech in healthy, typical speakers to gain further insight on sensorimotor control of vocal features of speech. The study intends to answer the following questions: 1) Is there a relationship between  $f_0$  auditory discrimination values and adaptation responses? 2) Is there a relationship between intensity discrimination values and adaptation responses? We hypothesized finding no relationship between auditory discrimination and adaptation of  $f_0$ , consistent with results of previous studies (Lester-Smith et al., 2020; Alemi et al., 2020). Additionally, we expected to find no relationship between auditory discrimination and adaptation of intensity due to the flexible nature of vocal features of speech to meet prosodic demands (Chang-Yit et al., 1975; H. L. Lane et al., 196; H. Lane & Tranel, 1971; Hilger et al., 2022).

## VII. Methods

### **Participants**

Twenty English-speaking participants between the ages of 18-58 years were included in this study. The sample size provided 80% power to detect moderate to large effect sizes ( $r > .58$ ). The mean age of participants was 28.7 years, with a standard deviation of 9.6 years. Participants included 12 women, 7 men, and 1 nonbinary individual. The participant's sex assigned at birth included 14 females and 6 males. Thirteen participants reported their race as white, 5 as Asian, 1 as Black or African American, and 1 participant's ethnicity was unknown. Three individuals reported their ethnicity as Hispanic or Latino, and 17 reported their ethnicity as not Hispanic or Latino. Speakers of tonal languages, and musicians with post-secondary musical training were excluded due to potential for superior abilities of auditory discrimination of  $f_0$  (Liu et al., 2010;

Micheyl et al., 2006). All participants were self-reported non-smokers, and reported no history of any type of communication disorder, neurological disorders, hearing disorders, or a diagnosis affecting attention.

### **Hearing Screen**

Due to the potential effect of hearing loss on auditory discrimination, all participants underwent a hearing screen with a Garson-Stadler GSI 17 portable audiometer. The passing threshold for individuals under the age of 50 was 25 dB at 125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz. For individuals 50 years of age and older, the passing threshold was 25 dB at frequencies 1000 Hz and below and 40 dB at 2000 Hz and above (*Scope of Practice in Audiology*, 2018).

### **Voice Screen**

All participants were screened for voice disorders based on self-reported history, auditory-perceptual judgment of voice quality by the supervising researcher for characteristics of dysphonia, and a reported history from the participant. The supervising researcher has been a speech-language pathologist with more than 14 years working with individuals with voice disorders. All participants had a Voice Handicap Index Scores (VHI-10) less than 11, which is the established cut-off differentiating those with a self-reported voice handicap from those with typical voice (Rosen et al., 2004).

### **Equipment**

An omni-directional ear-set microphone (Shure MX153) was placed 7 cm from the corner of the participant's mouth, at a 45-degree angle to capture the sound. The microphone signal was amplified by an RME QuadMic II Microphone Preamp and digitized at 44,100 Hz by

an RME Fireface sound card. A custom MATLAB 2019b program initiated trials and controlled Eventide Eclipse perturbations (Mathworks, 2019). The processed signal was amplified by a Behringer Xenyx Q802USB amplifier and played back to the participant through binaural Etymotic ER-2 insert earphones, which provided approximately 30 dB of attenuation of air-conducted sound.

## **Procedures**

Auditory Discrimination All participants began the experiment with an auditory discrimination task measuring auditory discrimination of  $f_0$  and intensity counterbalanced in order. Auditory discrimination values were quantified through just noticeable difference (JND) values of participants through a two down, one up staircase procedure (Levitt, 1971). Each participant was instructed to stay the word “hob” extending their vowel production for 2-3 seconds. A steady 500 ms segment of the vowel was extracted and presented to participants in two AXB tasks. During these AXB tasks, each participant was presented with a reference tone in the second interval and asked if the tone in the first or third interval was the same as the reference tone.

During the auditory discrimination task for  $f_0$ , sound presented to participants through their headphones was normalized at 75 dB. In the  $f_0$  AXB task, the first or third tone was randomly presented at 50 cents above the reference tone. The task reversed in direction of step size when the participant identified or misidentified the correct tone two trials in a row. After two correct trials, the step size decreased by 20 cents. Then after two reversals, it went down to 10 cents, then 4 cents, and ended after a total of 10 reversals. The JND was calculated as the average of the interval size of the last 6 reversals (Levitt, 1971).

The intensity AXB task followed a similar procedure to the  $f_0$  AXB task. However, the loudness interval was initially presented at a 5 dB difference above the reference tone and decreased to 1.4 dB, 0.8 dB, and 0.2 dB. During the auditory discrimination task for intensity, sound presented to participants through their headphones started at a baseline of 75 dB, with intervals split above and below this sound level.

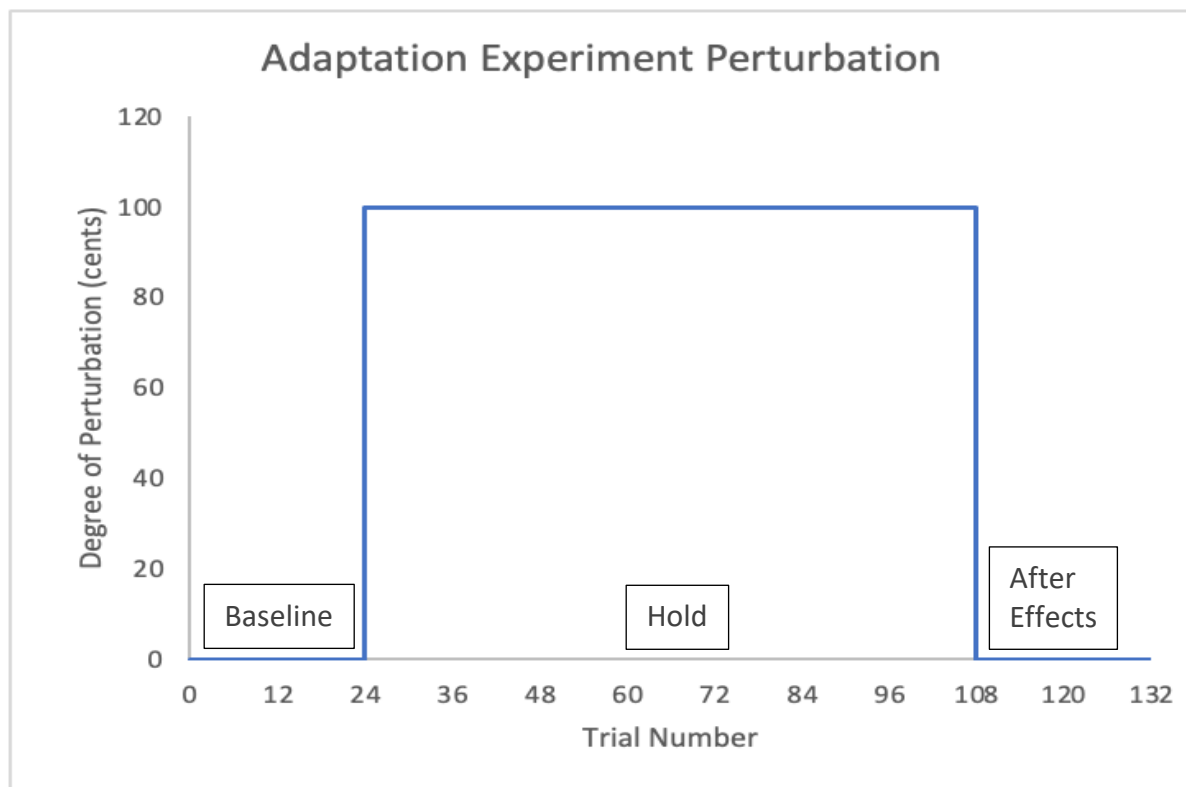
Adaptation of  $f_0$  In a  $f_0$  adaptation task, auditory feedback of participants' speech was manipulated while they were actively voicing. The insert headphones were calibrated to produce sound at a 5dB increase in sound pressure level (SPL) relative to the microphone level to mask ambient and bone-conducted sound. The participants were instructed to say one word per trial at a comfortable pitch and loudness, and using a naturalistic rate when it appears on a computer monitor. These target words, "hob", "hog" and "hod", were included in epochs. Each epoch contained twelve trials of these words, including four pseudo-randomly distributed trials of each word. The baseline and aftereffects phases each included two epochs (24 trials), and the hold phase included seven (84 trials).

This task included both a control condition and a shift condition. During the control phase, participants' feedback remained unperturbed during all phases. During the hold phase of the shift condition, the participant's voice was shifted by 100 cents relative to their baseline mean  $f_0$ . A ramp phase was not included in the study due to concurrent investigations that required the introduction of a large auditory error.

Adaptation of Intensity In the intensity adaptation task, auditory feedback of participants' speech was manipulated while they were actively voicing. During the control condition of this experiment, participants' feedback remained unperturbed across the 3 ordered phases. During the hold phase of the shift condition, the insert headphones were calibrated to produce sound at a

5dB increase in sound pressure level (SPL) relative to the microphone level. The structure of this task remained similar to the  $f_0$  adaptation task described above.

A general schematic of the experimental (shift) paradigm for both adaptation tasks is presented in Figure 2. For the purposes of this example, the degree of perturbation is illustrated specifically for the  $f_0$  adaptation task on the Y-axis. The baseline, hold, and after effects phases are labeled.



**Figure 2.** Perturbation of  $f_0$  relative to trial number during Adaptation of  $f_0$

### Descriptive Analysis

To analyze auditory discrimination, JND values were calculated as the average values of the last six reversals of each task in cents relating to  $f_0$  and dB (SPL) for intensity. Analysis of adaptation responses required the use of Praat (Boersma & Weenink, 2016) and MATLAB

2021b (Mathworks, 2021) scripts to track and analyze  $f_0$  and intensity traces. A portion of each vowel, 40 ms -120ms after onset, was extracted to exclude any instability from the onset of the vowel and conclude analysis before reflexive responses occurred (Tourville et al., 2008). Mean  $f_0$  of baseline trials was calculated in Hz, and mean intensity of baseline trials was calculated in dB (SPL). The average baseline value was subtracted from each trial so that all trials were normalized to the baseline average (Tourville et al., 2008). Additionally, each trial in the control condition was subtracted from the corresponding trial in the shift condition to control for natural changes in  $f_0$  and intensity while voicing over time (Jones & Munhall, 2000). Because adaptation tasks did not include a ramp phase, data exclusively from the second half of the hold phases (trials 66-108) were used to calculate adaptation of  $f_0$  and intensity.

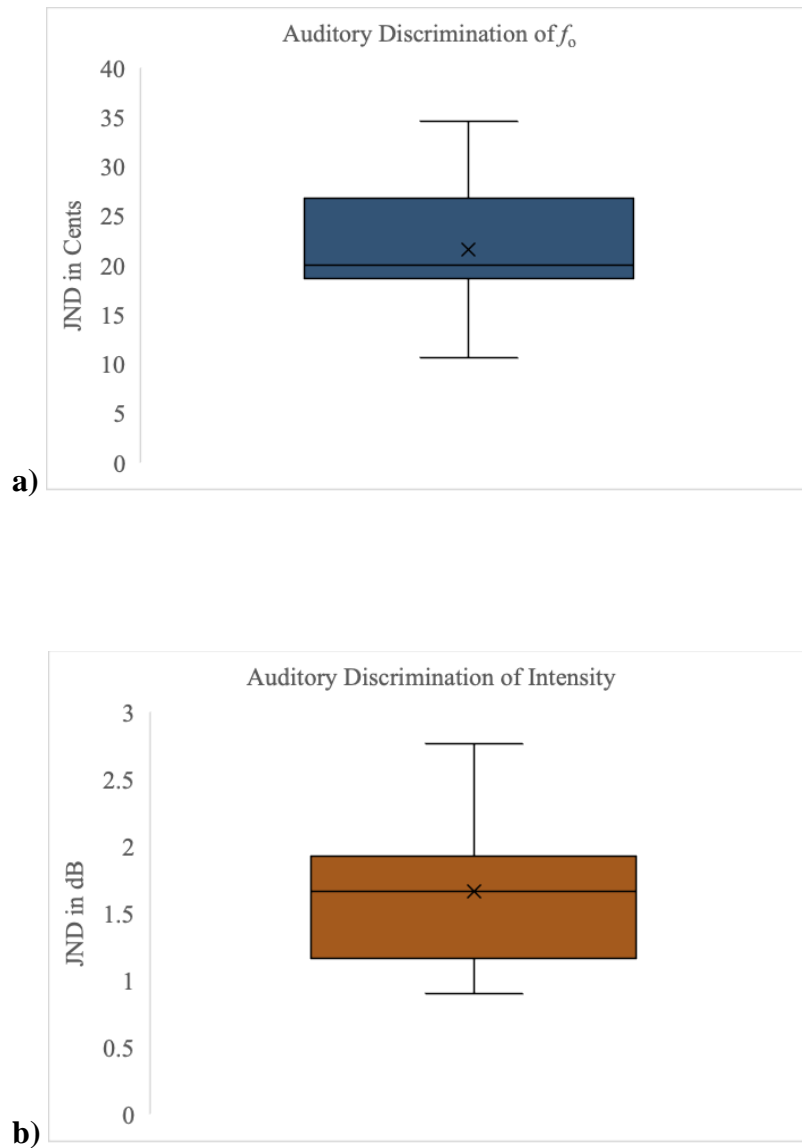
### **Statistical Analysis**

Scatter plots were first created to visually examine: relationships between  $f_0$  auditory discrimination values and  $f_0$  adaptation responses, and relationships between intensity discrimination values and intensity adaptation responses. Pearson's  $r$  correlation coefficients were used to statistically analyze these relationships. In this present study, the researchers defined  $r = 0.0-0.4$  as weak,  $r = 0.4-0.6$  as moderate, and  $r > 0.6$  as statistically strong (Akoglu, 2018).

## **VIII. Results**

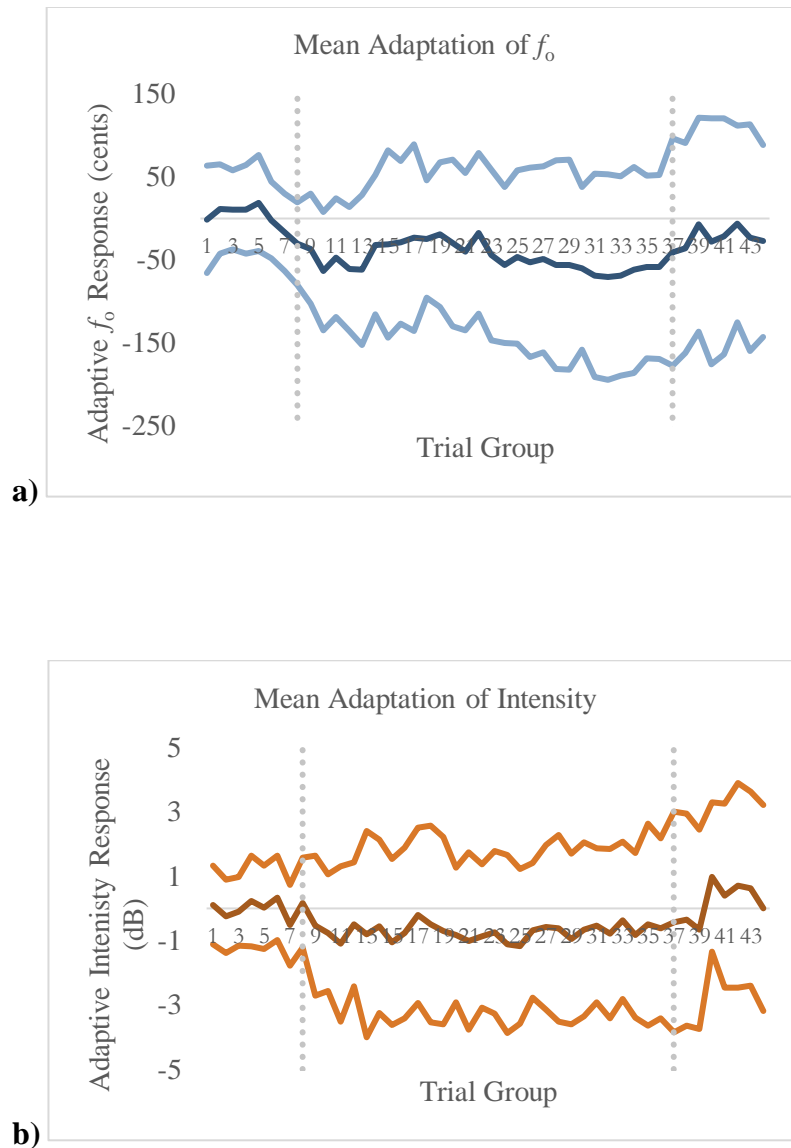
### **Descriptive Results**

On average, participants had JND values of  $f_0$  at 21.6 cents (SD = 6.3 cents, range = 10.7 - 34.7 cents). (See Fig 3a). The average JND value of intensity was 1.7 dB (SD = 0.6 dB, range = 0.9 - 2.8 dB). (See Fig 3b).



**Figure 3. a)** The mean and distribution of  $f_0$  JND depicted in a box plot. **b)** The mean and distribution of intensity JND depicted in a box plot.

The average adaptation response of  $f_0$  during the second-half of the hold phase was  $-57.7$  cents (SD = 98.4 cents, range =  $-269.6$  -  $147.4$  cents). (See Fig 4a). The average adaptation response of intensity during the second-half of the hold phase was  $-0.7$  dB (SD = 2.2 dB, range =  $5.5$  -  $3.2$  dB). (See Fig 4b).



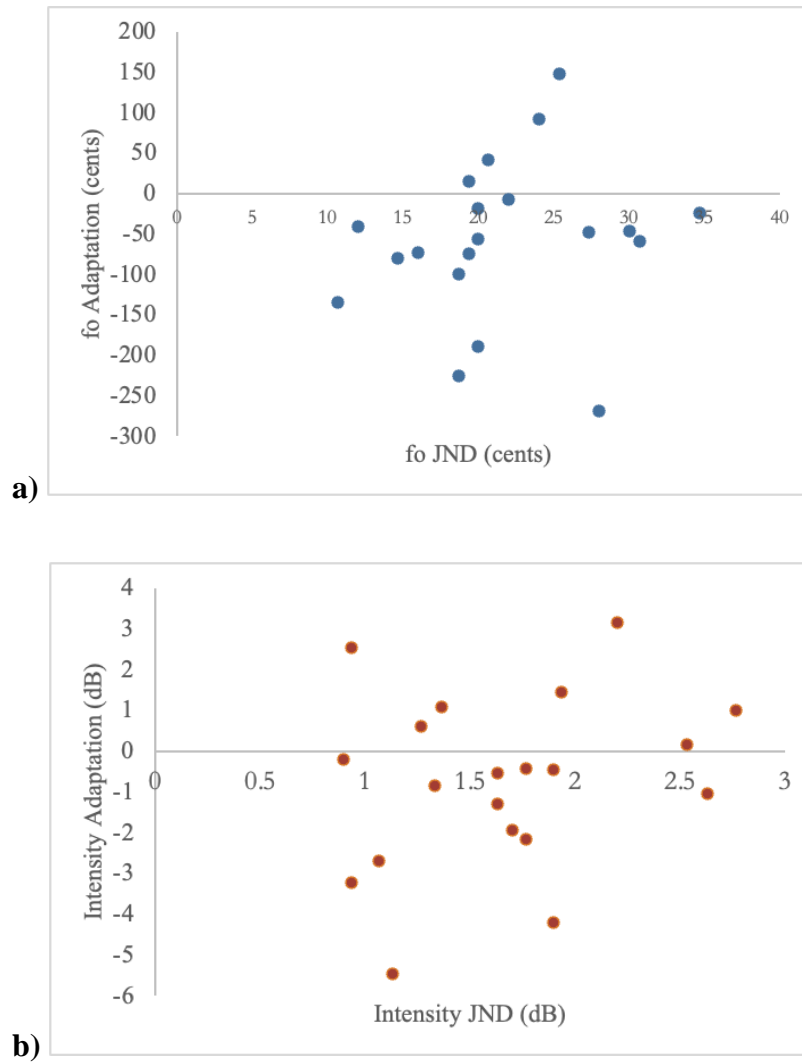
**Figure 4.** a) Average adaptation response of  $f_0$  ( $\pm$  1 standard deviation). b) Average adaptation response of intensity ( $\pm$  1 standard deviation).

### Relationships between Auditory Discrimination Values and Adaptation Responses

The average JND of  $f_0$  was not significantly correlated with the average adaptation response of  $f_0$  ( $r = .157$ ,  $p = .509$ ) (See Fig 5a). Additionally, the average JND of intensity was

not significantly correlated with the average adaptation response of intensity ( $r = .252, p = .284$ )

(See Fig 5b).



**Figure 5.** a) Participant values of JND of  $f_0$  in relation to adaptation response of  $f_0$ . b) Participant values of JND of intensity in relation to adaptation response of intensity.

## IX. Discussion

The purpose of this study was to investigate the relationships between auditory discrimination and adaptation of vocal features of speech in typical speakers. We set out to

answer the following questions: 1) Is there a relationship between  $f_0$  auditory discrimination values and adaptation responses? 2) Is there a relationship between intensity discrimination values and adaptation responses? We hypothesized no relationship between auditory discrimination and adaptation of  $f_0$ , due to results from previous studies finding no significant relationship (Lester-Smith et al., 2020; Alemi et al., 2020). Additionally, we expected no relationship between auditory discrimination and adaptation of intensity due to the flexible nature of vocal features of speech to meet prosodic demands (Chang-Yit et al., 1975; H. L. Lane et al., 1961; H. Lane & Tranel, 1971; Hilger et al., 2022). The results of this study provide insight on sensorimotor control of vocal features of speech.

### **Auditory Discrimination and Adaptation of $f_0$**

Before analyzing relationships between any discrimination values and adaptation responses, descriptive values of the current study relating to  $f_0$  may be compared to results from previous studies. Notably, the average JND value of  $f_0$  from the current study differs from results from Lester-Smith et al. (2020) and Abur et al. (2021). The average  $f_0$  JND was 21.6 cents (SD = 6.3 cents) in the current study. Results from Lester-Smith et al. (2020) indicate an average  $f_0$  JND of 37 cents (SD = 14 cents) and results from Abur et al. (2021) show an average  $f_0$  JND of 35 cents (SD = 20 cents). Possible reasons for a relatively lower  $f_0$  JND in our study when compared to other studies could be due to differences between the studies with regard to sample sizes as well as differences in data collection methods. First, our study included 20 typical speakers, who may not have had as much variability (and slightly better thresholds) as the larger sample of control speakers included in the study by Abur et al. (2021). Second, the current study used an AXB procedure to assess auditory discrimination whereas Lester-Smith et al. (2020) and Abur et al. (2021) utilized AX procedures. With an AXB procedure, participants are presented with a

reference tone in the second interval and asked if the tone in the first or third interval was the same as the reference tone. With an AX procedure, participants are presented with two tones and asked if they are the “same” or “different”. The disadvantage of AX procedures is that participants may be more likely to respond with a “same” response, which may be susceptible to bias. In contrast, the unique advantage of AXB procedures, which was used in this study, is that listeners compare a stimulus to two possibilities and know on each trial that either A or B is the correct answer, removing some of the bias problems in AX discrimination (Hautus et al., 2021).

Similarly, adaptation of  $f_0$  can be compared to results from Lester-Smith et al. (2020). Both studies shifted  $f_0$  100 cents relative to baseline. In the current study, the average adaptation response of  $f_0$  during the second half of the hold phase was -57.7 cents (SD = 98.4 cents). Lester-Smith et al. (2020) found an average  $f_0$  adaptation response of -61.2 cents (SD = 79 cents). Results from the two studies appear comparable. Of note, Lester-Smith et al. (2020) found an average adaptation response of  $F_1$  of -8.6% (SD = 6.2%). The difference in amount of average adaptation between  $F_1$  from Lester-Smith et al. (2020) and the average adaptation of  $f_0$  from the Lester-Smith et al. (2020) (approximately 61.2%) and the present study (approximately 57.7%) highlights potential differences in auditory motor control between  $F_1$  and  $f_0$ . Additionally, out of a sample size of 20 participants, both studies each identified four participants who exhibited a “following response” (or positive response) and one participant with a “non-response” (or minimal response) to upward perturbation of  $f_0$ . Although small sample sizes were present in both studies, similarity between the data strengthens our understanding of average adaptation of  $f_0$ . However, it should be noted that differences in study design could impact adaptation values between the two studies. The most salient difference is the absence of a ramp phase during the adaptation experiment and the use of the second half of the hold phase in the current study. Thus,

although the results for  $f_0$  JND in the current study are not consistent with previous studies, current findings of adaptation of  $f_0$  are consistent with results from Lester-Smith et al. (2020).

This study questioned if there is a relationship between  $f_0$  auditory discrimination values and adaptation responses. A very weak, non-significant relationship was found between the average  $f_0$  JND and the average  $f_0$  adaptation response ( $r = .157, p = .509$ ). These results are consistent with our hypothesis as well as previous studies indicating no significant relationship between these parameters (Alemi et al., 2020; Lester-Smith et al., 2020). Like the current study, Lester-Smith et al. (2020) found that adaptation of  $f_0$  was not associated with active or passive measures of  $f_0$  JND ( $r^2 = .175, p = .195$ ). Additionally, Alemi et al. (2020) found that auditory discrimination of  $f_0$  did not correlate with  $f_0$  adaptation response magnitude or direction when shifting  $f_0$  100 cents relative to baseline ( $r^2 = 0.01, p = 0.478$ ), or when personalizing the shift relative to each participants  $f_0$  JND ( $r^2 = 0.01, p = 0.406$ ).

Unlike Lester-Smith et al. (2020) and Alemi et al. (2020), Abur et al. (2021) found that high  $f_0$  JND values were correlated with atypical adaptation responses. Although these data suggest a relationship between  $f_0$  discrimination and adaptation, specifically for individuals with atypical adaptive behavior, Abur et al. (2021) included data from both control participants and individuals with hyperfunctional voice disorders in their analysis. Differences in statistical analysis methods and sample characteristics (i.e. exclusion of individuals with hyperfunctional voice disorders in this study) could explain variation of our results when compared to Abur et al. (2021). For example, participants in our study demonstrated JNDs within a typical range, and therefore, relationships between perception and production may have been limited.

### **Auditory Discrimination and Adaptation of Intensity**

Data on auditory discrimination of intensity from the present study can be compared to results from a previous study. Sinnott & Aslin (1985) have previously found auditory discrimination levels of intensity between 1-2 dB in six young adults when listening to a standard 1kHz tone. In the present study, the average JND value of intensity was 1.7 dB (SD = 0.6 dB, range = 0.9 - 2.8 dB) in twenty typical speakers. The results between the studies are comparable, yet differences in sample size and characteristics as well as methodological differences (such as using a speech vowel stimuli vs. a standard tone) may explain minor differences in results.

The novel part of the present study aimed to discover if there was a relationship between intensity discrimination values and adaptation responses. Consistent with our hypothesis, a weak, non-significant relationship was found between the intensity JND and the intensity adaptation response ( $r = .252, p = .284$ ). This is the first study that has examined the relationship between intensity discrimination and adaptation behavior. Previous research on side-tone amplification, an adaptation-like condition, has demonstrated flexibility of vocal intensity to meet conversational demands of speech in varying acoustic conditions (Chang-Yit et al., 1975; H. L. Lane et al., 1961; H. Lane & Tranel, 1971). This research highlights the need for flexibility in modulation of vocal intensity. Additionally, data from Patel et al. (2011) supports a trading relations theory. This theory suggests that there is an integration of control of  $f_0$  and intensity, such that when one feature is perturbed, the other feature adapts to compensate and meet the prosodic demands of speech. This theory suggests that broader auditory targets are permissible for  $f_0$  and intensity to create desired emphasis on certain syllables in conversation. Therefore, unlike articulatory parameters of speech, the amount of adaptation that occurs for these vocal features of speech may not be related to differences in an individual's auditory discrimination

ability (McRoberts, et. al, 1995; Williams Stevens, 1972). The necessity for flexibility of  $f_0$  and intensity for meeting prosodic speech demands offers an explanation for the lack of relationships between JND values and adaptation responses for these vocal parameters.

Results from this study highlight differences between sensorimotor integration of vocal features of speech and articulatory features of speech. In the present study, relationships between auditory discrimination and adaptation of vocal features of speech were not found. In contrast, Lester-Smith (2020) found a positive correlation between passive  $F_1$  auditory acuity and reflex responses to sudden  $F_1$  perturbation. The finding from Lester-Smith et al. (2020)'s study indicates that the articulatory feature of  $F_1$  is likely more sensitive to sudden perturbation than  $f_0$ , as errors in  $F_1$  must be corrected quickly to avoid vowel distortion. Similarly, Matthies et al. (2008) and Perkell et al. (2004) found relationships between perception and production for articulatory features of speech. Unlike vocal features speech like  $f_0$  or intensity, these findings suggest that sensorimotor control of articulatory features of speech likely rely on auditory discrimination abilities.

The differences in sensorimotor control between vocal features of speech and articulatory features of speech endorse the need for a model of sensorimotor integration of speech that addresses vocal features of speech separately. Computational models of speech adaptation such as DIVA and SimpleDIVA have largely focused on articulatory features of speech (Tourville & Guenther, 2011; Kearney, 2020). LaDIVA is the only neurocomputational model that incorporates neural control of speech and physiologically and biomechanically relevant aspects of laryngeal control for phonation (Weerathunge et al., 2022). The lack of relationships between auditory discrimination and adaptation of vocal features of speech in the current study highlight

the need for LaDIVA and other models of speech that conceptualize these features as distinct from articulatory features of speech.

## X. Additional Research and Clinical Implications

The current study investigated auditory-motor integration of vocal features of speech in individuals without voice or other communication disorders. Evidence from previous studies suggests differences in auditory discrimination abilities in individuals with hyperfunctional voice disorders (Hillman, 2020; Nguyen, 2022) and in auditory-motor integration of  $f_0$ . To determine if differences exist in auditory-motor integration of speech for both the vocal parameters  $f_0$  and intensity, results from this study will be used as a control group to enable a comparison with individuals diagnosed with hyperfunctional voice disorders.

Results from this current study could also inform future research with additional populations with voice and other communication disorders, as well as those who have other differences in vocal production and perception, such as singers. For example, future studies could consider comparing the current data to data from individuals with laryngeal dystonias (Thomas et al., 2021), unilateral vocal fold paralysis (Naunheim et al., 2019), Parkinson's disease (Abur et al., 2018; Senthinathan et al., 2021), and cerebellar disease (Hilger, 2020; Houde et al., 2019; Li et al., 2019) and singers (Abur, 2021). Studies of this nature could provide new information on potential risk factors or pathophysiology relating to auditory-motor integration of vocal features of speech. If causal relationships are found between auditory discrimination and speech motor control in these clinical populations, results might also provide future directions for rehabilitation or training.

## XI. Limitations

Limitations of this study include limited power because of small sample size. The sample size provided 80% power to detect moderate to large effect sizes ( $r > .58$ ). An increase in sample size would likely provide results that more accurately represent the population studied. As a result, future studies should be replicated in larger samples. Additionally, the age of participants in this study was disproportionately young as compared to the general population, with a mean age of 28.7 years (SD = 9.6 years) across the 20 participants. Results in the current study also showed little variability in  $f_0$  and intensity JND values as compared to those reported for participants with hyperfunctional voice disorders in previous studies (Abur et al., 2021). However, less variability is expected in typical, healthy participants than those with disruptions in speech or vocal sensorimotor control. The lack of variability in JND values of the present study could have prevented finding of significant relationships between auditory discrimination and adaptation of vocal features of speech. If participants with a wider range of JNDs were included, then it is possible that such relationships might emerge. Finally, variability in  $f_0$  adaptation behavior also observed in previous studies (Alemi, 2020; Abur et al., 2021) could have impacted the test-retest reliability of their (and our) results. Future studies should investigate the reliability of such approaches to ensure the results can be replicated and interpreted.

## XII. Conclusion

The purpose of this study was to investigate relationships between auditory discrimination and adaptation of  $f_0$ , and auditory discrimination and adaptation of intensity in typical speakers. Results indicated no to weak relationship was found ( $r = .157$ ) between auditory discrimination and adaptation of  $f_0$  in typical speakers. Additionally, no to weak relationship ( $r = .252$ ) was

found between auditory discrimination and adaptation of intensity in typical speakers. These findings differ from some studies that have found significant relationships between auditory discrimination and adaptation of articulatory features of speech (Martin et al., 2018; Villacorta et al., 2007). The results therefore support the use and further development of voice-specific models of sensorimotor control for vocal parameters (e.g., LaDIVA) as opposed to general models of sensorimotor control of speech (e.g., DIVA, SimpleDIVA) when analyzing auditory-motor control of vocal features of speech, such as  $f_0$  and intensity. Additionally, these findings could have implications for future directions of research for singers and various populations with voice disorders (e.g. hyperfunctional voice disorders, Laryngeal dystonia, Parkinson's disease, and cerebellar disease), who may have underlying sensorimotor deficits or differences than typical speakers. Studies of this kind could provide information on potential risk factors for certain disorders or pathophysiology relating to auditory-motor integration of vocal features of speech.

## References

- Abur, D., Lester-Smith, R. A., Daliri, A., Lupiani, A. A., Guenther, F. H., & Stepp, C. E. (2018). Sensorimotor adaptation of voice fundamental frequency in Parkinson's disease. *PLOS ONE*, 13(1), e0191839. <https://doi.org/10.1371/journal.pone.0191839>
- Abur, D., Subaciute, A., Kapsner-Smith, M., Segina, R. K., Tracy, L. F., Noordzij, J. P., & Stepp, C. E. (2021). Impaired auditory discrimination and auditory-motor integration in hyperfunctional voice disorders. *Scientific Reports*, 11(1), Article 1. <https://doi.org/10.1038/s41598-021-92250-8>
- Akoglu H. (2018). User's guide to correlation coefficients. *Turkish journal of emergency medicine*, 18(3), 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- Alemi, R., Lehmann, A., & Deroche, M. L. D. (2020). Adaptation to pitch-altered feedback is independent of one's own voice pitch sensitivity. *Scientific Reports*, 10(1), Article 1. <https://doi.org/10.1038/s41598-020-73932-1>
- Anderson, C. (2018). *Essentials of linguistics*. McMaster University.
- Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* (5.3-6.0).
- Buchsbaum, B. R., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*, 25(5), 663–678. [https://doi.org/10.1207/s15516709cog2505\\_2](https://doi.org/10.1207/s15516709cog2505_2)
- Chang-Yit, R., Pick, H. L., & Siegel, G. M. (1975). Reliability of sidetone amplification effect in vocal intensity. *Journal of Communication Disorders*, 8(4), 317–324. [https://doi.org/10.1016/0021-9924\(75\)90032-5](https://doi.org/10.1016/0021-9924(75)90032-5)
- Hautus, M. J., Macmillan, N. A., & Creelman, C. D. (2021). *Detection theory: A user's guide*. Routledge.

- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1–2), 67–99.  
<https://doi.org/10.1016/j.cognition.2003.10.011>
- Hillman, R. E., Stepp, C. E., Van Stan, J. H., Zañartu, M., & Mehta, D. D. (2020). An Updated Theoretical Framework for Vocal Hyperfunction. *American Journal of Speech-Language Pathology*, *29*(4), 2254–2260. [https://doi.org/10.1044/2020\\_AJSLP-20-00104](https://doi.org/10.1044/2020_AJSLP-20-00104)
- Hilger, A. I., Levant, S., Kim, J. H., Lester-Smith, R. A., & Larson, C. (2022). Task-Dependent Modulation of Auditory Feedback Control of Vocal Intensity. *Journal of Voice*, S0892199722002375. <https://doi.org/10.1016/j.jvoice.2022.08.004>
- Houde, J. F., Gill, J. S., Agnew, Z., Kothare, H., Hickok, G., Parrell, B., Ivry, R. B., & Nagarajan, S. S. (2019). Abnormally increased vocal responses to pitch feedback perturbations in patients with cerebellar degeneration. *The Journal of the Acoustical Society of America*, *145*(5), EL372–EL378. <https://doi.org/10.1121/1.5100910>
- Houde, J., & Nagarajan, S. (2011). Speech Production as State Feedback Control. *Frontiers in Human Neuroscience*, *5*. <https://www.frontiersin.org/articles/10.3389/fnhum.2011.00082>
- Howell, P. (1993). Cue trading in the production and perception of vowel stress. *The Journal of the Acoustical Society of America*, *94*(4), 2063–2073. <https://doi.org/10.1121/1.407479>
- Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of F0 production: Evidence from feedback perturbation. *The Journal of the Acoustical Society of America*, *108*(3), 1246.  
<https://doi.org/10.1121/1.1288414>
- Kearney, E., Nieto-Castañón, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., Ballard, K. J., Chang, S.-E., Chao, S.-C., Heller Murray, E. S., Scott, T. L., & Guenther, F. H. (2020). A Simple 3-Parameter Model for Examining Adaptation in Speech and Voice

- Production. *Frontiers in Psychology*, 10.  
<https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02995>
- Lane, H. L., Catania, A. C., & Stevens, S. S. (1961). Voice Level: Autophonic Scale, Perceived Loudness, and Effects of Sidetone. *The Journal of the Acoustical Society of America*, 33(2), 160–167. <https://doi.org/10.1121/1.1908608>
- Lane, H., & Tranel, B. (1971). The Lombard Sign and the Role of Hearing in Speech. *Journal of Speech and Hearing Research*, 14(4), 677–709. <https://doi.org/10.1044/jshr.1404.677>
- Lester-Smith, R. A., Daliri, A., Enos, N., Abur, D., Lupiani, A. A., Letcher, S., & Stepp, C. E. (2020). The Relation of Articulatory and Vocal Auditory—Motor Control in Typical Speakers. *Journal of Speech, Language, and Hearing Research*, 63(11), 3628–3643. [https://doi.org/10.1044/2020\\_JSLHR-20-00192](https://doi.org/10.1044/2020_JSLHR-20-00192)
- Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2B), 467–477. <https://doi.org/10.1121/1.1912375>
- Li, W., Zhuang, J., Guo, Z., Jones, J. A., Xu, Z., & Liu, H. (2019). Cerebellar contribution to auditory feedback control of speech production: Evidence from patients with spinocerebellar ataxia. *Human Brain Mapping*, 40(16), 4748–4758. <https://doi.org/10.1002/hbm.24734>
- Lieberman, P. (1960). Some Acoustic Correlates of Word Stress in American English. *The Journal of the Acoustical Society of America*, 32(4), 451–454. <https://doi.org/10.1121/1.1908095>
- Liu, H., Wang, E. Q., Chen, Z., Liu, P., Larson, C. R., & Huang, D. (2010). Effect of tonal native language on voice fundamental frequency responses to pitch feedback perturbations during sustained vocalizations. *The Journal of the Acoustical Society of America*, 128(6),

3739–3746. <https://doi.org/10.1121/1.3500675>

Lombard, E. (1911). Le signe de l'élevation de la voix. *Ann Mal de L'Oreille et Du Larynx*, 101–119.

MathWorks. (2019). MATLAB 2019 (Version b). Natick, Massachusetts, United States.

MathWorks. (2021). MATLAB 2021 (Version b). Natick, Massachusetts, United States.

Matthies, M. L., Guenther, F. H., & Denny, M. (2008). *Perception and production of /r/ allophones improve with hearing from a cochlear implant*. 13.

Martin, C. D., Niziolek, C. A., Duñabeitia, J. A., Perez, A., Hernandez, D., Carreiras, M., & Houde, J. F. (2018). Online Adaptation to Altered Auditory Feedback Is Predicted by Auditory Acuity and Not by Domain-General Executive Control Resources. *Frontiers in Human Neuroscience*, 12.

<https://www.frontiersin.org/articles/10.3389/fnhum.2018.00091>

McRoberts, G. W., Studdert-Kennedy, M., & Shankweiler, D. P. (1995). The role of fundamental frequency in signaling linguistic stress and affect: Evidence for a dissociation. *Perception & Psychophysics*, 57(2), 159–174. <https://doi.org/10.3758/BF03206502>

Ménard, L., Polak, M., Denny, M., Burton, E., Lane, H., Matthies, M. L., Marrone, N., Perkell, J. S., Tiede, M., & Vick, J. (2007). Interactions of speaking condition and auditory feedback on vowel production in postlingually deaf adults with cochlear implants. *The Journal of the Acoustical Society of America*, 121(6), 3790.

<https://doi.org/10.1121/1.2710963>

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, 219(1–2), 36–47. <https://doi.org/10.1016/j.heares.2006.05.004>

- Naunheim, M. L., Yung, K. C., Schneider, S. L., HendersonSabes, J., Kothare, H., Mizuiri, D., Klein, D. J., Houde, J. F., Nagarajan, S. S., & Cheung, S. W. (2019). Vocal motor control and central auditory impairments in unilateral vocal fold paralysis. *The Laryngoscope*, 129(9), 2112–2117. <https://doi.org/10.1002/lary.27680>
- Nguyen, D. D., Chacon, A. M., Novakovic, D., Hodges, N. J., Carding, P. N., & Madill, C. (2022). Pitch Discrimination Testing in Patients with a Voice Disorder. *Journal of Clinical Medicine*, 11(3), 584. <https://doi.org/10.3390/jcm11030584>
- Parrell, B., Ramanarayanan, V., Nagarajan, S., & Houde, J. (2019). The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLoS Computational Biology*, 15(9), e1007321–e1007321. <https://doi.org/10.1371/journal.pcbi.1007321>
- Patel, R., Niziolek, C., Reilly, K., & Guenther, F. H. (2011). Prosodic adaptations to pitch perturbation in running speech. *Journal of Speech, Language, and Hearing Research*, 54(4), 1051–1060. [https://doi.org/10.1044/1092-4388\(2010/10-0162\)](https://doi.org/10.1044/1092-4388(2010/10-0162))
- Patel, R., Reilly, K. J., Archibald, E., Cai, S., & Guenther, F. H. (2015). Responses to intensity-shifted auditory feedback during running speech. *Journal of Speech, Language, and Hearing Research*, 58(6), 1687–1695. [https://doi.org/10.1044/2015\\_JSLHR-S-15-0164](https://doi.org/10.1044/2015_JSLHR-S-15-0164)
- Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics*, 25(5), 382–407. <https://doi.org/10.1016/j.jneuroling.2010.02.011>
- Perkell, J. S. (2013). Five decades of research in speech motor control: What have we learned, and where should we go from here? *Journal of Speech, Language, and Hearing Research*, 56(6), S1857–S1857. [https://doi.org/10.1044/1092-4388\(2013/12-0382\)](https://doi.org/10.1044/1092-4388(2013/12-0382))
- Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., &

- Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America*, *116*(4), 2338–2344. <https://doi.org/10.1121/1.1787524>
- Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., Stockmann, E., & Guenther, F. H. (2004). The distinctness of speakers' /s/-/[integral]/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research*, *47*(6), 1259–1270.
- Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and Validation of the Voice Handicap Index-10: *The Laryngoscope*, *114*(9), 1549–1556. <https://doi.org/10.1097/00005537-200409000-00009>
- Senthinathan, A., Adams, S., Page, A. D., & Jog, M. (2021). Speech intensity response to altered intensity feedback in individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research*, *64*(6S), 2261–2275. [https://doi.org/10.1044/2021\\_JSLHR-20-00278](https://doi.org/10.1044/2021_JSLHR-20-00278)
- Scope of Practice in Audiology*. (2018). American Speech-Language-Hearing Association; American Speech-Language-Hearing Association. <https://www.asha.org/policy/sp2018-00353/>
- Sinnott, J. M., & Aslin, R. N. (1985). Frequency and intensity discrimination in human infants and adults. *The Journal of the Acoustical Society of America*, *78*(6), 1986–1992. <https://doi.org/10.1121/1.392655>
- Thomas, A., Mirza, N., & Eliades, S. J. (2021). Auditory Feedback Control of Vocal Pitch in Spasmodic Dysphonia. *The Laryngoscope*, *131*(9), 2070–2075. <https://doi.org/10.1002/lary.29254>

- Tomassi, N. E., Castro, M. E., Timmons Sund, L., Díaz-Cádiz, M. E., Buckley†, D. P., & Stepp, C. E. (2021). Effects of Sidetone Amplification on Vocal Function During Telecommunication. *Journal of Voice*, S0892199721001247.  
<https://doi.org/10.1016/j.jvoice.2021.03.027>
- Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952–981.  
<https://doi.org/10.1080/01690960903498424>
- Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, 39(3), 1429–1443.  
<https://doi.org/10.1016/j.neuroimage.2007.09.054>
- Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, 122(4), 2306–2319. <https://doi.org/10.1121/1.2773966>
- Weerathunge, H. R., Alzamendi, G. A., Cler, G. J., Guenther, F. H., Stepp, C. E., & Zañartu, M. (2022). LaDIVA: A neurocomputational model providing laryngeal motor control for speech acquisition and production. *PLOS Computational Biology*, 18(6), e1010159.  
<https://doi.org/10.1371/journal.pcbi.1010159>
- Williams, C. E., & Stevens, K. N. (1972). Emotions and Speech: Some Acoustical Correlates. *The Journal of the Acoustical Society of America*, 52(4B), 1238–1250.  
<https://doi.org/10.1121/1.1913238>