ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Test–Retest Reliability of Behavioral Assays of Feedforward and Feedback Auditory–Motor Control of Voice and Articulation

Mara R. Kapsner-Smith,[a] Defne Abur,[b,c,d] Tanya L. Eadie,[a] and Cara E. Stepp[b,e,f]

[a] Department of Speech and Hearing Sciences, University of Washington, Seattle [b] Department of Speech, Language, and Hearing Sciences, Boston University, MA [c] Department of Computational Linguistics, Center for Language and Cognition, University of Groningen, the Netherlands [d] Research School of Behavioral and Cognitive Neurosciences, University of Groningen, the Netherlands [e] Department of Biomedical Engineering, Boston University, MA [f] Department of Otolaryngology–Head and Neck Surgery, Boston University School of Medicine, MA

ARTICLE INFO

Article History:
Received January 16, 2023
Revision received July 24, 2023
Accepted September 25, 2023

Editor: Raymond D. Kent

https://doi.org/10.1044/2023_JSLHR-23-00038

ABSTRACT

Purpose: Behavioral assays of feedforward and feedback auditory–motor control of voice and articulation frequently are used to make inferences about underlying neural mechanisms and to study speech development and disorders. However, no studies have examined the test–retest reliability of such measures, which is critical for rigorous study of auditory–motor control. Thus, the purpose of the present study was to assess the reliability of assays of feedforward and feedback control in voice versus articulation domains.

Method: Twenty-eight participants (14 cisgender women, 12 cisgender men, one transgender man, one transmasculine/nonbinary) who denied any history of speech, hearing, or neurological impairment were measured for responses to predictable versus unexpected auditory feedback perturbations of vocal (fundamental frequency, f_o) and articulatory (first formant, F_1) acoustic parameters twice, with 3–6 weeks between sessions. Reliability was measured with intraclass correlations.

Results: Opposite patterns of reliability were observed for f_o and F_1; f_o reflexive responses showed good reliability and f_o adaptive responses showed poor reliability, whereas F_1 reflexive responses showed poor reliability and F_1 adaptive responses showed moderate reliability. However, a criterion-referenced categorical measurement of f_o adaptive responses as typical versus atypical showed substantial test–retest agreement.

Conclusions: Individual responses to some behavioral assays of auditory–motor control of speech should be interpreted with caution, which has implications for several fields of research. Additional research is needed to establish reliable criterion-referenced measures of F_1 adaptive responses as well as f_o and F_1 reflexive responses. Furthermore, the opposite patterns of test–retest reliability observed for voice versus articulation add to growing evidence for differences in underlying neural control mechanisms.

Humans learn and maintain movement patterns for intelligible speech by monitoring sensory feedback, especially in the auditory domain. By comparing auditory feedback of speech with internal predictions, speakers can detect errors and produce online corrections as well as updates to stored motor programs (Tourville & Guenther, 2011). In studies of sensorimotor control of voice and articulation, experimenters often measure behavioral responses to altered auditory feedback to assess auditory–motor control of speech (e.g., Burnett et al., 1997; Cai et al., 2011; Daliri et al., 2020; Houde & Jordan, 1998; Jones & Munhall, 2000; Larson & Robin, 2016; Lester-Smith et al., 2020; Villacorta et al., 2007). These behavioral assays provide experimental data that have been used to develop and test neurocomputational models of speech

Correspondence to Mara R. Kapsner-Smith: mkapsner@uw.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

sensorimotor control and are used in combination with neurophysiological methods to investigate the neurobiology of speech (e.g., Houde & Chang, 2015; Tourville & Guenther, 2011). Studies of children and aging adults have contributed to our understanding of speech development across the life span (Caudrelier & Rochet-Capellan, 2019; H. Liu, Russo, & Larson, 2010; Scheerer et al., 2020). Recently, an increasing number of studies use these paradigms to examine impaired sensorimotor function in clinical populations (Caudrelier & Rochet-Capellan, 2019; Weerathunge, Tomassi, & Stepp, 2022). Such experiments can elucidate the pathophysiology of speech disorders as well as test hypotheses about sensorimotor control of speech using clinical populations as a model. However, to date, no studies have directly examined the reliability of behavioral assays of auditory–motor control of voice and articulation. This information is crucial to appropriately interpret differences between and within experimental groups in participants' responses to altered auditory feedback.

Behavioral Assays of Feedforward and Feedback Auditory–Motor Control

Sensorimotor control of speech is an incredibly complex task, requiring precise control of approximately 100 muscles to produce rapid movements and quickly changing acoustic features, on a scale of 50–300 ms (Parrell et al., 2019). At the same time, a speaker must be able to adapt to changing environmental conditions, such as background noise, to maintain intelligibility. This is achieved through a combination of feedforward and feedback control mechanisms. Feedforward control involves execution of preprogrammed motor commands for speech movements that allow for fast and fluent production. At the same time, an internal prediction of the sensory outcomes of the movements is activated. Comparison of this prediction to the actual sensory feedback allows the speaker to detect errors and generate corrective, feedback-based motor commands. Persistent error signals may lead to changes to feedforward motor commands, that is, sensorimotor learning.

An extensive body of literature has explored the role of auditory feedback in sensorimotor control of speech by applying near–real-time perturbations to parameters of auditory feedback. These include parameters related to voice, such as fundamental frequency (f_o; e.g., Burnett et al., 1997), and parameters related to articulation, such as formants (e.g., Houde & Jordan, 1998). When a speaker is exposed to persistent, predictable perturbations of their auditory feedback over many trials, they typically will adapt by progressively opposing the change across trials and will demonstrate a brief period of persistence of this change after the perturbation is removed (Daliri & Dittman, 2019; Houde & Jordan, 1998; Jones & Munhall, 2000; Purcell & Munhall, 2006a; Villacorta et al., 2007). This is interpreted as reflecting updates to underlying feedforward motor commands based on an error signal, often referred to as an adaptive response (Parrell & Houde, 2019; Tourville & Guenther, 2011). When a speaker is exposed to an unexpected, sudden-onset perturbation of their auditory feedback, they typically will respond with a rapid correction opposing the direction of the perturbation (Cai et al., 2011; Daliri et al., 2020; Larson & Robin, 2016; Niziolek & Guenther, 2013; Purcell & Munhall, 2006b). This is interpreted as a reflex-like correction produced by feedback control loops (Parrell & Houde, 2019; Tourville & Guenther, 2011). Measures such as the magnitude of the response to predictable or unexpected auditory feedback perturbations are thus used as behavioral assays of underlying feedforward and feedback control mechanisms, respectively, although both mechanisms are involved in these tasks.

Use of Assays to Study Neurophysiology of Speech

Behavioral assays of speech auditory–motor control have been combined with neurophysiological methods to elucidate underlying control mechanisms. Feedforward control assays combined with noninvasive neurostimulation methods including anodal transcranial direct stimulation and repetitive transcranial magnetic stimulation have provided information about the neurophysiological bases of speech sensorimotor learning. To date, these studies have focused on adaptive responses to predictable perturbations of the first formant (F_1), an acoustic parameter related to articulatory movements of the tongue and jaw. Changes in the magnitude and/or rate of adaptation in response to stimulation support the role of a sensorimotor corticocerebellar loop in integrating auditory feedback error signals into feedforward motor plans, that is, adaptive auditory–motor learning for speech (Deroche et al., 2017; Lametti et al., 2018; Scott et al., 2020; Shum et al., 2011; Tang et al., 2021).

Feedback control assays combined with methods including positron emission tomography, magnetoencephalography, electroencephalography, electrocorticography, and functional magnetic resonance imaging (fMRI) have characterized the role of the auditory cortex in comparison of predicted and actual auditory feedback for voice and articulation. Responses of the auditory cortex are significantly smaller when speakers hear their own speech during speech production (i.e., auditory feedback) than when they listen passively to playback of the same speech signal or when they hear auditory feedback that has been altered, a phenomenon known as speaking-induced suppression (Chang et al., 2013; Flinker et al., 2010; Greenlee et al., 2011; Hirano et al., 1997; Houde et al., 2002; Sato & Shiller, 2018). This evidence supports the hypothesis that auditory feedback is compared to an internal prediction to detect speech errors. Further evidence comes from studies of responses to altered auditory feedback, which reveal enhanced auditory cortical responses to altered auditory feedback during speech compared to passive listening to playback of the same altered signals, known as speech perturbation response enhancement (Behroozmand et al., 2009; Chang et al., 2013; Greenlee et al., 2013; Kort et al., 2014; Parkinson et al., 2012). Furthermore, studies using fMRI during feedback control assays show increased activity in and connectivity between bilateral posterior superior temporal cortices, supporting their role in error detection, and right and/or bilateral motor cortices, suggesting a role in translating error signals into corrective motor output (Behroozmand et al., 2015; Floegel et al., 2020; Niziolek & Guenther, 2013; Tourville et al., 2008). Although such studies have contributed substantially to understanding the neurobiology of speech sensorimotor control, information about the test–retest reliability of the behavioral measures is crucial to understand whether these responses reflect state (temporary, situational) versus trait (stable, long-lasting) characteristics of the speaker. Furthermore, information about reliability could provide future directions for mechanistic studies to determine what causes the behavior of an individual to change.

Use of Assays to Study Speech Development

Studies of the developmental trajectory of auditory feedback perturbation responses suggest that young children use auditory feedback to guide sensorimotor control and learning for both vocal and articulatory parameters. Children as young as 24 months have been shown to compensate for perturbations of auditory f_o feedback, suggesting that toddlers use auditory feedback to facilitate motor control of vocal parameters of speech (Scheerer et al., 2020). It is possible that feedback-based control of vocal parameters matures earlier than articulatory control (MacDonald et al., 2012; Scheerer et al., 2020), but more research is needed to confirm these findings. Maturation decreases variability of both vocal and articulatory parameters and decreases the latency of feedback responses to unpredictable f_o perturbations (H. Liu, Russo, & Larson, 2010; P. Liu, Chen, et al., 2010; MacDonald et al., 2012; Scheerer et al., 2013, 2016, 2020; van Brenk & Terband, 2020). This may be due to neurodevelopmental factors such as myelination that increase neural transmission speed and reduce processing and response times for auditory feedback responses.

Two studies have investigated the influence of aging on responses to auditory feedback perturbations of f_o. Aging appears to increase the magnitude of compensatory responses to unpredictable f_o perturbations, but it is unknown whether this reflects changes in feedback control mechanisms or physiological pitch control more generally (H. Liu, Russo, & Larson, 2010; P. Liu et al., 2011). The influence of aging on responses to auditory feedback perturbations of articulatory parameters and on adaptive responses to predictable perturbations remains unexplored.

Use of Assays to Study Speech Disorders

An increasing number of auditory feedback perturbation studies examine clinical populations. One purpose of such studies is to elucidate the pathophysiology of voice and articulation disorders that previously were poorly understood, such as developmental stuttering and hyperfunctional voice disorders. Numerous studies document a broad, likely developmental impairment of sensorimotor control and learning in individuals who stutter, including articulatory (Cai et al., 2012, 2014; Daliri et al., 2018; Kim, Daliri, et al., 2020; Sengupta et al., 2016), vocal (Bauer et al., 2007; Loucks et al., 2012; Sares et aal., 2018, 2020), and limb control (Kim, Daliri, et al., 2020). Recent evidence points to impaired adaptive auditory–motor learning for voice in some people with hyperfunctional voice disorders, with likely preserved feedback auditory–motor control (Abur, Subaciute, Kapsner-Smith, et al., 2021; Stepp et al., 2017). Such studies may have substantial implications for diagnosis and treatment of speech disorders.

In contrast, patient populations with known neurogenic impairments can serve as models to understand underlying neural mechanisms of typical sensorimotor control. For example, in people with aphasia secondary to left-hemisphere stroke, lesion mapping, combined with feedback control assays and neurophysiological methods, has supported the role of frontotemporal networks in error detection and correction (Behroozmand et al., 2018, 2022). Behavioral and electroencephalographical data from people with cerebellar disease suggest cerebellar function is important for feedforward auditory–motor adaptive learning and may play a role in feedback auditory–motor control through interactions with cortical regions that modulate auditory–motor integration (Hilger, 2020; Houde et al., 2019; Li et al., 2019; Parrell et all., 2017). Finally, people with Parkinson's disease (PD), a neurodegenerative disease associated with impaired dopamine production, have shown increased responses to feedback control assays for the vocal parameters f_o and intensity while off dopaminergic medications (X. Chen et al., 2013; H. Liu et al., 2012; Mollaei et al., 2016) but decreased responses for the articulatory parameter F_1 (Mollaei et al., 2016). This dissociation of impaired feedback auditory–motor control of voice versus articulation suggests differences in underlying neural mechanisms. Although no studies to date have examined differences in medication state within the same individuals, studies of people with PD while on dopaminergic medications have shown typical feedback control responses for both vocal and articulatory parameters (Abur, Subaciute, Daliri, et al., 2021; Kiran & Larson, 2001), suggesting that dopaminergic signaling is involved in typical auditory–motor feedback control. Crucially, most of these studies of clinical populations have relied on group-level findings, despite substantial variability in behavior at the individual level (e.g., Abur, Subaciute, Daliri, et al., 2021). Information about the test–retest reliability of individual responses to behavioral assays of speech sensorimotor control will help determine whether they are rigorous measures capable of reliably detecting differences between groups or changes in individuals over time.

Reliability of Behavioral Assays of Auditory–Motor Control

Reliability refers to "the extent to which an experiment, test, or measuring procedure yields the same results on repeated trials" (Merriam-Webster, n.d.). Good test–retest reliability suggests that a measure is an accurate representation of an individual's performance rather than artifact related to irrelevant aspects of the testing session or environment. Stability of a measure over time may also suggest that it is sensitive to trait, rather than state, characteristics of the individual being measured, for example, a stable capacity of the speech sensorimotor control system versus one that fluctuates in response to internal and/or external conditions. Despite the importance of reliability for interpretation of results, to date, no studies have explicitly examined the test–retest reliability of behavioral assays of auditory–motor control of speech. One study examined adaptive auditory–motor learning for voice in response to two differently sized perturbations of f_o—one standard perturbation and the other that was personalized to the individual's f_o discrimination threshold (Alemi et al., 2020). Participants' responses were only weakly related across the two experiments, with substantial variability in the magnitude and the direction of f_o changes. Another study examined feedback control responses to perturbations of F_1 with and without the addition of noise (i.e., artificially increased variability) in participants' auditory feedback prior to the experiment (Niziolek & Parrell, 2021). The authors noted that interpretation of results in this study was limited by a high degree of intra-individual variability, potentially due to "limited stability of auditory feedback compensation measures. . .across sessions" (Niziolek & Parrell, 2021, p. 2169). In addition, it is unknown whether learning effects occur with repeated exposure to identical tasks. Although learning effects have been observed in studies of reaching movements, these studies often involve longer exposure to perturbations (e.g., Shadmehr & Holcomb, 1997), and important differences have been documented between learning in limb and speech systems, such as the presence (limb) and absence (speech) of explicit learning mechanisms (Kim & Max, 2020; Lametti et al., 2020). These studies highlight the importance of directly investigating the test–retest reliability of measures of feedforward and feedback auditory–motor control of voice and articulation.

To interpret within-group and between-groups differences in auditory perturbation experiments, investigators must know how reliable responses to auditory perturbation are in individuals with typical speech. To date, no studies have examined the test–retest reliability of common assays of feedforward and feedback control of voice and articulation. Thus, the purpose of the present study was to measure the test–retest reliability of adaptive responses to predictable perturbations and reflexive responses to unexpected perturbations in both the voice (f_o) and articulation (F_1) domains. Given the variability of responses to auditory feedback perturbations between individuals and across similar experiments (Alemi et al., 2020; Niziolek & Parrell, 2021), we hypothesized that test–retest reliability of these assays may be only moderate. As a secondary analysis, we also examined whether systematic learning effects occurred with repeated exposure to typical experimental tasks measured several weeks apart.

Method

All procedures were approved by the institutional review board of the University of Washington.

Participants

Twenty-eight adults reporting normal speech, language, and hearing function were recruited to participate in the study (14 cisgender women, 12 cisgender men, one transgender man, one transmasculine/nonbinary; 16 assigned female at birth, 12 assigned male at birth; M_{age} = 25 years, *SD* = 4.5, range: 18–35 years). The sample size provided 80% power to detect medium effect sizes (intraclass correlation coefficient [ICC] ≥ .6). Individuals with a self-reported history of neurological disease, stroke, or any health conditions affecting communication or the ability to pay attention were excluded. Individuals who were smokers were excluded due to potential effects on vocal function. Individuals who reported formal singing training or were active vocal performers were excluded. Speakers of tonal languages were excluded. All participants passed a hearing screening with thresholds at or below 25 dB HL at 125, 250, 500, 1000, 4000, and 8000 Hz (American Speech-Language-Hearing Association, 2018). All participants learned English in infancy.

Participants were screened for voice concerns via interview and administration of the Voice-Related Quality of Life (Hogikyan & Sethuraman, 1999). Participants reported no voice overuse or upper respiratory symptoms for at least 48 hr prior to each recording session. Participants reporting illness or voice overuse were rescheduled to allow for vocal recovery (Hunter & Titze, 2009).

Procedure

Adaptive and reflexive responses to f_o and F_1 perturbations (Lester-Smith et al., 2020) were each measured twice for all participants. Baseline measures of f_o and F_1 control occurred in separate sessions due to additional data collection for a different study and were conducted within a span of 2 weeks (hereafter referred to as Time 1). Retest measures of f_o and F_1 control occurred in a single session, which was conducted 3–6 weeks after the second baseline session (hereafter referred to as Time 2). Each session lasted 1–2 hr. We chose a timeline that we expected was unlikely to induce learning and might occur in a longitudinal study design, while also avoiding changes in vocal function over time due to seasonal or other factors. The order of the first two sessions and the order of f_o and F_1 tasks in the third session were counterbalanced across participants to control for order effects.

Participants were seated comfortably in front of a monitor that presented visual stimuli. They were provided with water before and during the experiments. Participants wore a head-mounted Shure Omni Mic MX153 7 cm from the center of the mouth at an approximately 45° angle and Etymotic ER-2 insert earphones. The microphone signal was amplified +5 dB and played back to participants via the headphones with a total processing delay of < 35 ms (Kim, Wang, & Max, 2020; Weerathunge et al., 2020). Both the microphone and headphone digitized signals were recorded at 16 kHz for F_1 experiments and 44.1 kHz for f_o experiments.

For f_o perturbation experiments, the microphone signal was amplified by an RME QuadMic II Microphone Preamp and digitized by an RME Fireface sound card. Manipulations of f_o were performed by an Eventide Eclipse V4. The manipulated signal was amplified by a Behringer Xenyx Q802USB amplifier and played back to the participant via headphones, as well as being digitized via the RME Fireface sound card. Initiation of trials and Eventide Eclipse manipulations was controlled by a custom MATLAB program (MathWorks, 2016).

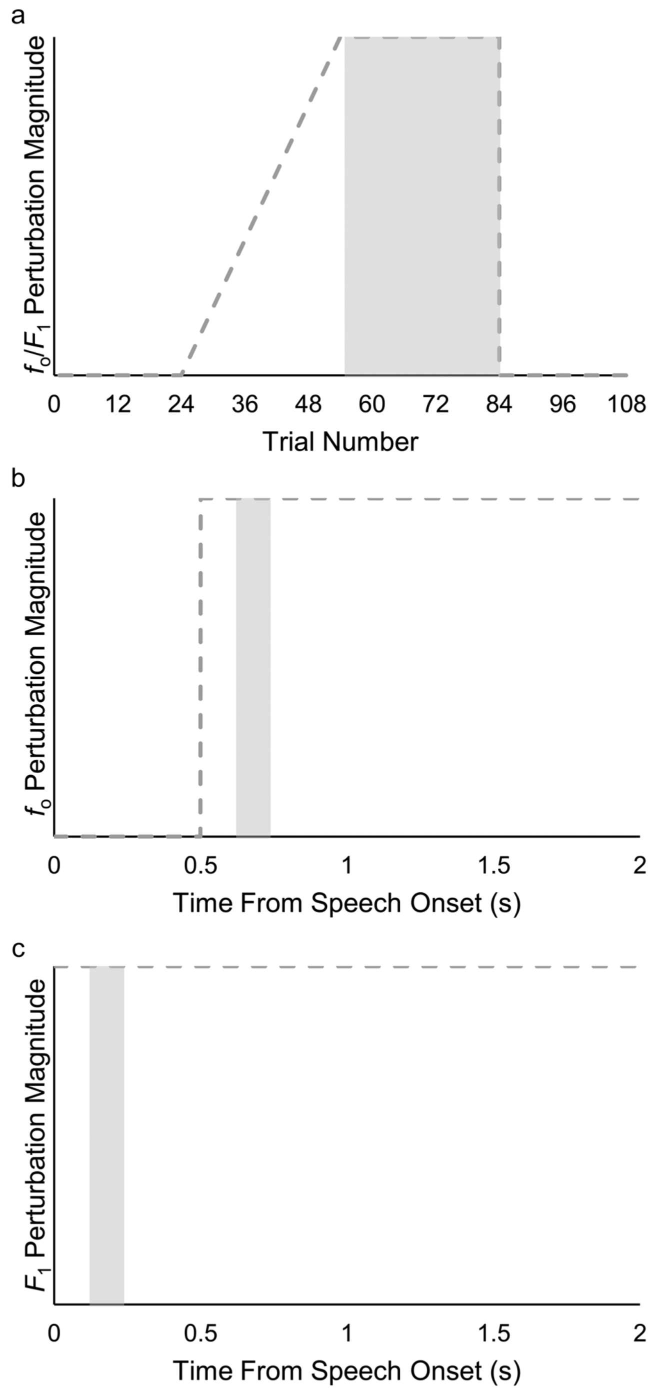
For F_1 perturbation experiments, the microphone signal was amplified and digitized by a MOTU MicroBook IIC microphone preamplifier using MicroBook II CueMix software. F_1 manipulations were performed by Audapter (Cai et al., 2008), a MATLAB software package that performs real-time transformations of acoustic parameters. The signal was then sent to headphones via a Behringer Xenyx Q802USB amplifier. Initiation of trials and Audapter manipulations were controlled by a custom MATLAB program (MathWorks, 2016).

There were three f_o perturbation conditions: adaptation shift-up, adaptation control, and reflex shift-up (10 min each; Lester-Smith et al., 2020). Adaptation conditions always occurred before the reflex condition. The order of the two adaptation conditions was counterbalanced across participants. Each condition consisted of 108 trials, during which participants produced a 2-s sustained /ɑ/. During the adaptation shift-up condition, participants received unperturbed feedback during the first 24 trials (baseline phase), gradually upward-shifted feedback of +3.4 cents per trial relative to the participant's f_o during the next 30 trials (ramp phase), 30 trials at the maximum perturbation of +100 cents (hold phase), and 24 trials of unperturbed feedback (aftereffect phase). Perturbations during the adaptation condition, when applicable, were applied for the duration of the trial. During the adaptation control condition, participants received unperturbed feedback for all 108 trials. During the reflex shift-up condition, participants produced a 2-s sustained /ɑ/ for 12 blocks of nine trials (total of 108). Each block included two perturbed trials. Perturbations consisted of a sudden +100 cent shift of f_o. Onset of perturbations was jittered between 0.5 and 1 s after the onset of phonation. After onset, perturbations remained on for the remainder of the trial.

There were two F_1 perturbation conditions: adaptation shift-up and reflex shift-up (10 min each; Lester-Smith et al., 2020). The adaptation condition always occurred before the reflex condition. Each condition consisted of 108 trials. Participants produced a prolonged word for approximately 2 s when presented with the word on the screen. During the adaptation shift-up condition, a block consisted of one presentation of each of the stimuli "bid," "tid," and "hid," in random order (36 blocks × 3 trials = 108 trials). These stimuli were chosen so that they would be perceived as a real word when F_1 was perturbed upward (toward "bed," "Ted," and "head"). Participants received unperturbed feedback during the first eight blocks (baseline phase), gradually upward-shifted F_1 feedback of

+1.03% per trial relative to F_1 produced by the participant across the next 10 blocks (ramp phase), 10 blocks at the maximum perturbation of +30% of F_1 (hold phase), and eight blocks of unperturbed feedback (aftereffect phase). During the reflex shift-up condition, a block consisted of three presentations of each of the stimuli “bid,” “tid,” and “hid,” in random order (12 blocks × 9 trials = 108 trials). Each block included two perturbed trials. Perturbations consisted of a +30% shift-up of F_1. For both the adaptation and reflex conditions, perturbations were applied for the duration of the trial, beginning with the onset of speech. Schematics for adaptation and reflex experiments are provided in Figure 1.

Figure 1. Schematics of (a) adaptation conditions and of a single perturbed reflex trial for (b) f_o and (c) F_1. Dashed lines represent the perturbation magnitude, and shaded regions represent the analysis windows. f_o = fundamental frequency; F_1 = first formant.



Data Analysis

For adaptation measures, the mean value of f_o or F_1 was calculated during the window 40–120 ms after the onset of the vowel in each trial. This window was chosen to assess contributions of the feedforward sensorimotor control system after the voice stabilized but before auditory feedback responses are expected to occur (Burnett et al., 1998). For the f_o adaptation experiment, mean f_o of each trial was extracted using an autocorrelation method via Praat (Boersma & Weenink, 2016) and converted to cents, a relative measure of f_o, by normalizing to the average of the 24 baseline trials. Trials with inadequate pitch tracking or duration were replaced with the average of the trial before and the trial after. For f_o, a total of 0.36% of all trials were replaced, averaging 0.4 trials per experimental condition (range: 0–9 trials per condition). The f_o shift-up condition was then normalized to the f_o control condition by subtracting the mean f_o in cents of the respective control trial from each shift-up trial. This normalization accounted for any change in f_o across trials that was not related to the auditory feedback perturbation. This procedure was used for f_o and not F_1 because f_o tends to drift upward over many trials (Jones & Munhall, 2000). For the F_1 adaptation experiment, F_1 was extracted using linear predictive coding and converted to a percentage relative to the mean F_1 of the 24 baseline trials (Lester-Smith et al., 2020). For F_1, no trials needed to be replaced. For both experiments, the average of the 30 trials in the hold phase was used for statistical analyses of the adaptive responses. A second measure using the mean of the last 15 trials (second half) of the hold phase was also calculated and compared to the primary measure. This was conducted as a post hoc analysis to assess whether responses were more stable in the later portion of the hold phase, for example, if adaptation has plateaued.

For reflex measures, f_o or F_1 traces were extracted for all trials. Perturbed f_o trials were normalized in cents relative to the 100-ms period preceding the start of the perturbation. Perturbed F_1 trials were normalized as a percentage relative to the mean trajectory of unperturbed trials containing the same word (Lester-Smith et al., 2020). Perturbed trials with inadequate pitch tracking or duration were removed from the analysis. For f_o, there was an average of 22.89 usable reflex trials per participant (range: 17–24 out of 24 possible trials). For F_1, only two reflex trials were removed from the entire data set, resulting in an average of 23.99 usable F_1 reflex trials per participant (range: 23–24 out of 24 possible trials). The mean normalized value of f_o or F_1 was calculated across all perturbed trials during the window 120–240 ms after the onset of the perturbation (Lester-Smith et al., 2020), consistent with timing of reflexive responses reported by Tourville et al. (2008) and Hain et al. (2000). This mean was used for statistical analyses of the reflexive responses.

Statistical Analyses

For f_o adaptation, F_1 adaptation, f_o reflex, and F_1 reflex, ICCs were used to assess the reliability of the responses at Time 1 and Time 2. ICCs and 95% confidence intervals were calculated using SPSS (Version 26, IBM, Inc.) based on single measures, absolute agreement, and two-way random effects (McGraw & Wong, 1996). Categories of poor (< .50), moderate (.50–.75), good (.76–.90), and excellent (> .90) reliability have been suggested for ICCs (Koo & Li, 2016). As a secondary analysis, two-tailed paired *t* tests were used to assess whether learning effects occurred from repeated exposure to the behavioral assays, with an α level of .05.

Prior studies of f_o adaptive responses in individuals with typical voice and speech have documented substantial variability, including both responses that oppose the perturbation and responses that follow (i.e., change in the same direction as) the perturbation (e.g., Abur, Subaciute, Kapsner-Smith, et al., 2021; Lester-Smith et al., 2020). An additional categorical analysis of f_o adaptive responses using cutoff criteria from a prior study was therefore conducted to assess its reliability. This analysis was not conducted for the other behavioral assays (f_o reflex, F_1 adaptation, F_1 reflex) because similar criterion cutoff scores were not available. The mean f_o adaptive response was categorized as typical or atypical based on 90th percentile *z*-score cutoffs taken from nonsinging control participants in the study of Abur, Subaciute, Kapsner-Smith, et al. (2021). Cohen's kappa statistic was calculated to assess agreement of this categorical score between Time 1 and Time 2.

Results

Descriptive statistics for group results for each of the four experiments (f_o adaptation, F_1 adaptation, f_o reflex, and F_1 reflex) are provided in Table 1. On average, participants opposed the f_o and F_1 auditory perturbations at both time points. The average magnitudes of f_o adaptive responses were −37.7 (SD = 70.3) and −51.3 (SD = 72.3) with a similar number of participants producing following responses at each time point. The average magnitudes of f_o reflexive responses were −8.9 (SD = 8.7) and −10.8 (SD = 10.0). The average magnitudes of F_1 adaptive responses were −9.0 (SD = 7.3) and −6.6 (SD = 5.7). The average magnitudes of F_1 reflexive responses were −1.1 (SD = 1.6) and −1.2 (SD = 1.6). Individual data are presented in Figures 2–5. As a secondary analysis, paired *t* tests were calculated for each of the four experiments to determine if systematic learning effects occurred. There was no significant difference for any of the measures at the two time points (*t* statistics and *p* values are provided in Table 1).

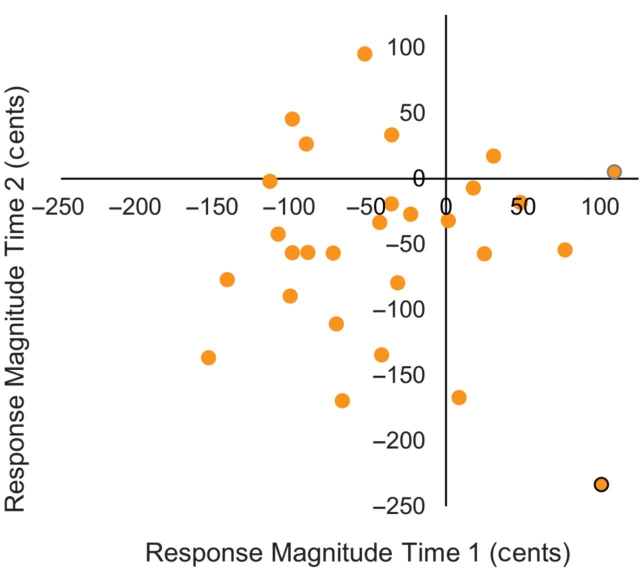
For the f_o reflex task, at Time 1, participants responded to an average of 54.7% of perturbed trials in a compensatory direction (SD = 17.8%) and 20.0% of trials in a following direction (SD = 12.0%) and did not respond to 25.3% (SD = 11.8%). Similarly, at Time 2, participants responded to an average of 55.4% of perturbed trials in a compensatory direction (SD = 18.6%) and 19.7% in a following direction (SD = 14.3%) and did not respond to 25.0% (SD = 11.8%).

Table 1. Descriptive statistics for the magnitude of adaptive and reflexive responses and paired *t* tests.

Experiment	Time	Mean response	Min	Max	*SD*	*t*(27)	*p*	*d*
f_o adaptation	1	–37.7 cents	–154.1	109.4	70.3	0.70	.492	0.13
f_o adaptation	2	–51.3 cents	–233.5	95.2	72.3			
F_1 adaptation	1	–9.0%	–19.2	15.4	7.3	–2.04	.051	0.39
F_1 adaptation	2	–6.6%	–14.4	7.7	5.7			
f_o reflex	1	–8.9 cents	–36.4	5.3	8.7	1.85	.075	0.35
f_o reflex	2	–10.8 cents	–39.4	9.8	10.0			
F_1 reflex	1	–1.1%	–5.6	1.3	1.6	0.09	.932	0.02
F_1 reflex	2	–1.2%	–5.3	2.1	1.6			

Note. Negative values reflect opposing responses, whereas positive values reflect following responses. f_o = fundamental frequency; F_1 = first formant.

Figure 2. Individual f_o adaptive responses at Time 1 and Time 2. One participant produced a response in the atypical range at Time 1 and in the typical range at Time 2 (outlined in gray), and one produced responses in the atypical range at both times (outlined in black). f_o = fundamental frequency.



ICCs were calculated for each of the four experiments. For f_o adaptive responses, ICC(2, 1) = −.052 [−.422, .326], p = .604, consistent with poor test–retest reliability. For F_1 adaptive responses, ICC(2, 1) = .536 [.219, .753], p = .001, consistent with moderate test–retest reliability. For f_o reflexive responses, ICC(2, 1) = .833 [.663, .919], p < .001, consistent with good test–retest reliability. For F_1 reflexive responses, ICC(2, 1) = .474 [.122, .718], p = .005, consistent with poor test–retest reliability.

ICCs were also calculated for the two adaptation experiments using the mean of the last 15 trials (second half) of the hold phase. For f_o adaptive responses, ICC(2, 1) = −.115 [−.475, .270], p = .719, consistent with poor test–retest reliability. For F_1 adaptive responses, ICC(2, 1) = .525 [.206, .745], p = .001, consistent with moderate

Figure 3. Individual F_1 adaptive responses at Time 1 and Time 2. F_1 = first formant.

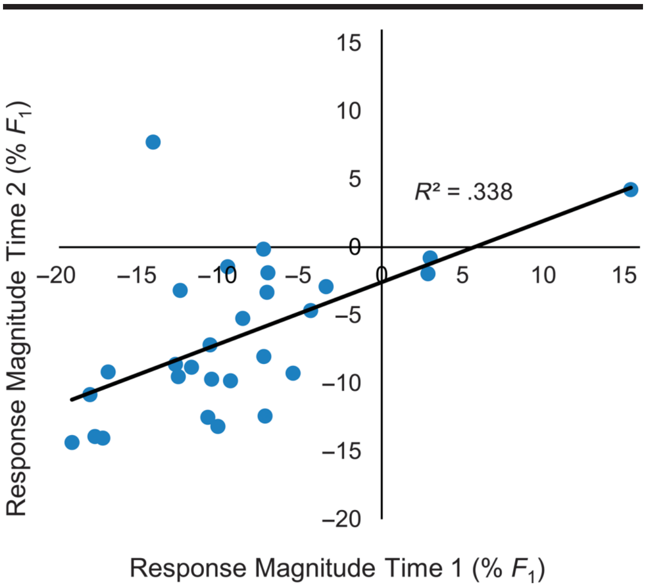


Figure 4. Individual f_o reflexive responses at Time 1 and Time 2. f_o = fundamental frequency.

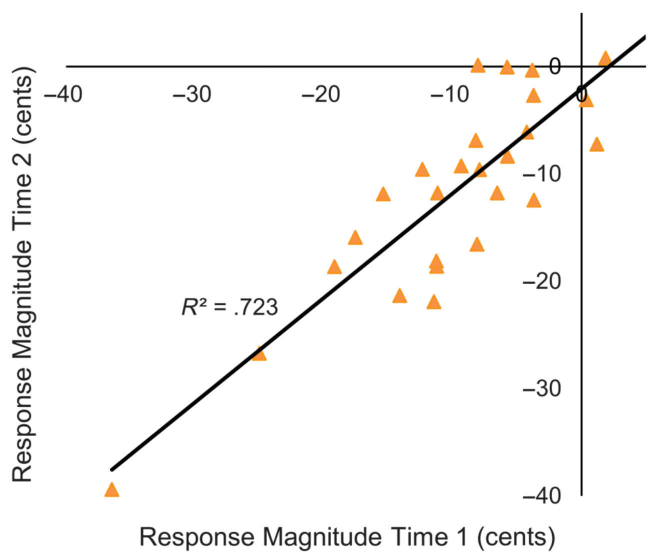
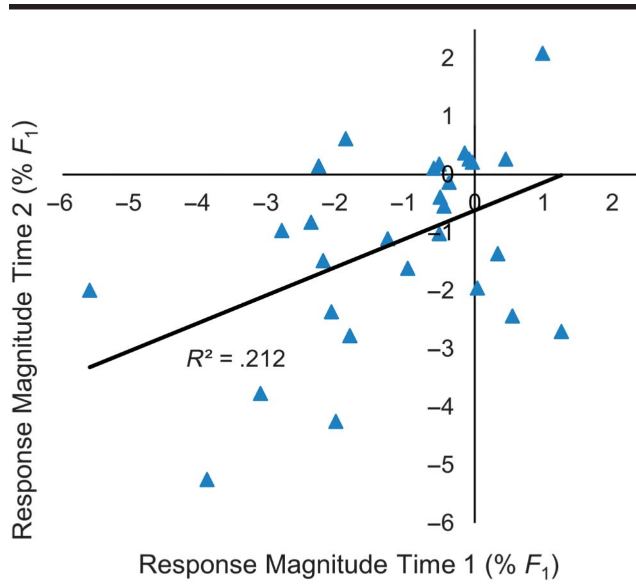


Figure 5. Individual F_1 reflexive responses at Time 1 and Time 2. F_1 = first formant.



test–retest reliability. These results were slightly less reliable than ICCs calculated using the mean of the entire hold phase.

Criterion-referenced assessments are a common way to categorize responses into typical or atypical categories based on normative data. This type of approach was used by Abur, Subaciute, Kapsner-Smith, et al. (2021) to identify atypical f_o adaptive responses in a subgroup of people with vocal hyperfunction. To assess the stability of f_o adaptive responses within a typical range in the present study, participants' responses were categorized into typical versus atypical using 90th percentile z scores from data collected by Abur, Subaciute, Kapsner-Smith, et al. (2021). Cutoff scores were calculated based on nonsingers with typical voices (n = 33). f_o adaptive responses were categorized as typical if they were greater than −199.5 cents or less than 97.8 cents. Based on these cutoffs, 26 participants in the present study were categorized as typical responders at both Time 1 and Time 2, one participant was categorized as an atypical responder at both Time 1 and Time 2, and one participant was categorized as an atypical responder at Time 1 and a typical responder at Time 2. This results in 96.4% agreement with k = .65 reflecting substantial agreement.

Discussion

Test–retest reliability was assessed for common behavioral assays of feedforward and feedback auditory–motor control of voice and articulation in people with typical speech. These included adaptive learning responses and reflexive responses to predictable versus unexpected perturbation of voice (f_o) and articulation (F_1) parameters of auditory feedback. The magnitudes of F_1 adaptive and reflexive responses were similar to those reported by Lester-Smith et al. (2020), which used the same methods. Consistent with our hypothesis, test–retest reliability for the magnitude of F_1 adaptive responses was moderate, and reliability for F_1 reflexive responses was poor. However, test–retest reliability for the magnitude of f_o adaptive and reflexive responses followed an opposite pattern, with good reliability for reflexive responses and poor reliability for adaptive responses. Notably, when f_o adaptive responses were categorized into typical versus atypical scores based on the distribution of scores among control participants in a prior study (Abur, Subaciute, Kapsner-Smith, et al., 2021), there was substantial agreement between scores at the two time points.

Secondary testing revealed no significant changes from Time 1 to Time 2 suggestive of systematic learning effects in any of the auditory–motor control assays. Although results were borderline for two of the assays (F_1 adaptation and f_o reflex), effect sizes were small. This suggests that these assays may be used for repeated testing in study designs with exposure duration and spacing similar to the present study, without substantial concern for systematic changes such as learning effects due to the measurement task itself. However, reliability of these measures should be taken into account when interpreting repeated measures. Concerns regarding reliability may be particularly relevant for studies examining individual responses as well as individual changes in longitudinal and interventional designs. Such studies are crucial for characterization of causal mechanisms and for clinical applications in people with communication disorders.

The opposite patterns in reliability we observed for behavioral assays of auditory–motor control of voice versus articulation suggest differences in underlying neurophysiological mechanisms. For voice, reflexive responses to unpredictable f_o perturbations were highly reliable, whereas the magnitude of adaptive responses to predictable f_o perturbations showed no relationship between time points. It has been suggested that online feedback-based control of parameters related to voice, such as f_o and intensity, may be important for maintaining intelligibility in response to changing environmental or other conditions (Perkell et al., 1997). Vocal parameters may be controlled at multiple time scales, with a phrase component extending across numerous phoneme and word boundaries (Ladd, 2008). Fast and slow components of f_o control may be modeled separately (Fujisaki, 2004), and there is evidence for distinct neural control mechanisms (Dichter et al., 2018). The slower components of vocal parameters thus may be amenable to feedback-based control. In contrast, acoustic parameters associated with speech articulation, such as formants, change more rapidly than the time that is required for online feedback-based responses. Thus, to produce fluent and intelligible speech, a speaker must rely on an internal model for feedforward control of articulation. Theories of speech motor control include auditory feedback as crucial for both learning and maintaining such models (Parrell et al., 2019; Perkell et al., 1997; Tourville & Guenther, 2011). These views are consistent with our findings of more reliable reflexive responses to f_o perturbations than F_1, and the reverse finding for adaptive responses.

The present study adds to growing evidence for differences in sensorimotor control mechanisms underlying voice versus articulation. Studies that have compared adaptive and reflexive responses to perturbations of f_o and formants suggest that there is no significant relationship between responses to f_o and formant feedback perturbations in the same speakers (MacDonald & Munhall, 2012; Weerathunge, Voon, et al., 2022). Average compensation for unexpected perturbations of f_o (e.g., Burnett et al., 1998; S. H. Chen et al., 2007) tends to be larger as a percentage of the shift than compensation for perturbations of formants (e.g., Purcell & Munhall, 2006b; Tourville et al., 2008), including when responses are measured within the same speakers (Lester-Smith et al., 2020). The same pattern occurred in the present study. Computational models of speech motor control such as the Directions Into Velocities of Articulators model suggest that that magnitude of adaptive and reflexive responses will be related to the speaker's discrimination threshold for the perturbed parameter, with the latter closely related to the speaker's feedforward speech target (Tourville & Guenther, 2011). Indeed, some studies have found significant relationships between F_1 discrimination and F_1 reflexive responses (Lester-Smith et al., 2020) and between F_1 discrimination and F_1 adaptive responses (Nault & Munhall, 2020; Villacorta et al., 2007) in people with typical speech, although others have failed to find a relationship (Daliri et al., 2020). Specifically, individuals with better discrimination thresholds produced larger responses in both paradigms. However, studies of f_o discrimination and perturbation responses have failed to find significant relationships for reflexive or adaptive responses in people with typical speech (Abur et al., 2018; Abur, Subaciute, Kapsner-Smith, et al., 2021; Alemi et al., 2020; Lester-Smith et al., 2020). Furthermore, a study of people with congenital amusia, characterized by impaired f_o discrimination, revealed preserved reflexive responses to unpredictable f_o perturbations (Hutchins & Peretz, 2013). These findings suggest that, whereas auditory discrimination thresholds for articulatory parameters may indeed reflect characteristics of underlying control mechanisms such as the size of articulatory targets, auditory discrimination and auditory–motor control of voice appear to be independent for people with typical speech. Notably, a relationship between poor f_o discrimination and atypical f_o adaptive responses was found for some individuals with hyperfunctional voice disorders (Abur, Subaciute, Kapsner-Smith, et al., 2021), suggesting that these functions share some underlying mechanism and/or have a causal relationship when auditory–motor control of voice is atypical. Finally, additional evidence for mechanistic differences between auditory–motor control of articulation and voice comes from studies of people with PD, a neurodegenerative disease associated with hypokinetic dysarthria that impacts both articulation and voice. Studies of speakers with PD while off anti-parkinsonian medication have shown larger reflexive responses to f_o and intensity perturbations (X. Chen et al., 2013; H. Liu et al., 2012; Mollaei et al., 2016, 2019) but decreased reflexive responses to F_1 perturbations (Mollaei et al., 2016). This dissociation of impaired auditory–motor control of voice versus articulation supports mechanistic differences between the two domains.

The magnitude of adaptive responses to persistent f_o perturbations had poor test–retest reliability in the present study. This may reflect susceptibility of the adaptive auditory–motor learning system for voice to factors that were not controlled. People with typical speech display a wide range of f_o adaptive responses, including large compensatory responses as well as following responses, that is, changes in f_o in the same direction as the feedback perturbation (e.g., Abur, Subaciute, Kapsner-Smith, et al., 2021; Lester-Smith et al., 2020). Researchers have suggested that the f_o auditory–motor control mechanism may respond flexibly to auditory feedback errors depending on whether they are perceived as internally versus externally generated (Burnett et al., 1998) and that there may be separate compensatory versus imitative modes resulting from different underlying neural processes (Korzyukov et al., 2012; Li et al., 2013; Patel et al., 2014, 2019). However, to date, these studies have focused on reflexive responses to unpredictable f_o perturbations. When averaged across trials, these responses were highly reliable in the present study. It is possible that f_o adaptive responses may have similar flexibility and may vary depending on uncontrolled factors such as internal states of the participant during an individual testing session.

Although the magnitude of f_o adaptive responses was not reliable across testing sessions, categorical scoring of f_o adaptive responses as typical versus atypical showed substantial test–retest agreement. This scoring method was implemented by Abur, Subaciute, Kapsner-Smith, et al. (2021) to detect differences in people with and without hyperfunctional voice disorders that were masked in comparisons of group means due to the presence of atypical responses at both ends of the distribution (i.e., compensating vs. following). Because there is substantial variability in f_o adaptive responses in people with typical speech, scores that fall significantly outside of that distribution are remarkable. This scoring method has relevance for studying impaired auditory–motor control of speech in people with communication disorders. Studies with larger normative samples and/or meta-analysis of existing f_o adaptation studies are needed to establish more precise cutoff scores for typical f_o adaptive responses. Furthermore, normative data for F_1 reflexive responses could be used to assess the reliability of categorical measurement of articulatory feedback control using a similar approach.

One limitation of the present study is the difference in tasks (sustained phonation vs. words) for the f_o versus F_1 experiments. This approach was used because most current studies of f_o adaptative responses use sustained phonation, and most current studies of F_1 adaptative responses use speech or speechlike tasks. It is possible that participants' f_o feedforward targets may be less well defined for sustained phonation than their F_1 feedforward targets are for prolonged words. This could lead to more variability in f_o responses than F_1 responses. However, the opposite pattern of reliability results observed for reflexive versus adaptive responses to f_o and F_1 perturbations is not easily explained by task differences alone. Prior studies suggest differences in task may affect the magnitude of f_o reflexive responses, with larger responses observed for singing than for speech (Natke et al., 2003), and for speech than for sustained phonation (S. H. Chen et al., 2007). These studies suggest f_o feedback control is more active when the feedforward target is more defined. In addition, responses to f_o perturbations may vary depending on linguistic and musical experience, for example, singers (Zarate & Zatorre, 2008) and speakers of tonal languages (H. Liu, Wang, et al., 2010) in whom the f_o auditory–motor control system is likely more developed. Future studies should examine the reliability of f_o adaptation during different tasks and in participants with different linguistic and musical experience to see if responses are more stable depending on task and experience.

Reliability of assays of auditory–motor control of speech may be impacted by noise. F_1 and f_o production is highly variable from trial to trial, and this variability may mask the stability of feedback and feedforward control mechanisms. It is possible that reliability may be improved by increasing the number of trials or other methodological changes that decrease the noisiness of the measures. Furthermore, the sample size in this study was relatively small, with 80% power to detect medium (ICC ≥ .6) effect sizes. A larger sample size could lead to different results.

In studies of auditory–motor control of speech, numerous methodological options exist for parameters including task and perturbation type, size, and timing. For example, the use of extended words in the F_1 tasks in this study means participants' exposure to the perturbation was greater than in a natural word duration, which could have impacted the adaptive response. Outcomes such as the magnitude of responses may vary systematically with changes to task parameters, and it is conceivable that test–retest reliability could also be different depending on methodological choices. Future studies should examine the effects of such parameters on test–retest reliability.

Conclusions

Opposite patterns of reliability were observed for behavioral assays of feedforward and feedback control of voice (f_o) versus articulation (F_1). f_o reflexive responses showed good reliability and f_o adaptive responses showed poor reliability, whereas F_1 reflexive responses showed poor reliability and F_1 adaptive responses showed moderate reliability. However, a criterion-referenced categorical measurement of f_o adaptive responses as typical versus atypical showed substantial test–retest agreement. Individual responses to some behavioral assays of auditory–motor control of speech should be interpreted with caution, especially when analyzing individual performance and/or change over time. Future studies should examine the reliability of criterion-referenced measures of F_1 adaptive responses as well as f_o and F_1 reflexive responses. Furthermore, the opposite patterns of test–retest reliability observed for voice versus articulation add to growing evidence for differences in underlying neural control mechanisms.

Data Availability Statement

The data sets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grants R01 DC016270 (Cara E. Stepp), P50 DC015446 (Robert E. Hillman), T32 DC005361 (David J. Perkel), and F31 DC020359 (Mara R. Kapsner-Smith). The authors thank Chloe Thorburn, Callum Wilson, Aaron Moy, and Monique Tardif for assistance with data analysis.

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), Article e0191839. https://doi.org/10.1371/journal.pone.0191839

Abur, D., Subaciute, A., Daliri, A., Lester-Smith, R. A., Lupiani, A. A., Cilento, D., Enos N. M., Weerathunge H. R., Tardif M. C., & Stepp C. E. (2021). Feedback and feedforward auditory-motor processes for voice and articulation in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 64*(12), 4682–4694. https://doi.org/10.1044/2021_JSLHR-21-00153

Abur, D., Subaciute, A., Kapsner-Smith, M., Segina, R. K., Tracy, L., Noordzij, J. P., & Stepp, C. E. (2021). Impaired auditory-discrimination and auditory-motor integration in hyperfunctional voice disorders. *Scientific Reports, 11*(1), Article 13123. https://doi.org/10.1038/s41598-021-92250-8

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 16860. https://doi.org/10.1038/s41598-020-73932-1

American Speech-Language-Hearing Association. (2018). *Scope of practice in audiology*. https://www.asha.org/policy/sp2018-00353/

Bauer, J. J., Hubbard Seery, C., LaBonte, R., & Ruhnke, L. (2007). Voice F0 responses elicited by perturbations in pitch of auditory feedback in individuals that stutter and controls. *The Journal of the Acoustical Society of America, 121*(Suppl. 5), 3201–3201. https://doi.org/10.1121/1.4782465

Behroozmand, R., Bonilha, L., Rorden, C., Hickok, G., & Fridriksson, J. (2022). Neural correlates of impaired vocal feedback control in post-stroke aphasia. *NeuroImage, 250,* Article 118938. https://doi.org/10.1016/j.neuroimage.2022.118938.

Behroozmand, R., Karvelis, L., Liu, H., & Larson, C. R. (2009). Vocalization-induced enhancement of the auditory cortex responsiveness during voice F_0 feedback perturbation. *Clinical Neurophysiology, 120*(7), 1303–1312. https://doi.org/10.1016/j.clinph.2009.04.022

Behroozmand, R., Phillip, L., Johari, K., Bonilha, L., Rorden, C., Hickok, G., & Fridriksson, J. (2018). Sensorimotor impairment of speech auditory feedback processing in aphasia. *NeuroImage, 165,* 102–111. https://doi.org/10.1016/j.neuroimage.2017.10.014

Behroozmand, R., Shebek, R., Hansen, D. R., Oya, H., Robin, D. A., Howard, M. A., III, & Greenlee, J. D. (2015). Sensory–motor networks involved in speech production and motor control: An fMRI study. *NeuroImage, 109,* 418–428. https://doi.org/10.1016/j.neuroimage.2015.01.040

Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* (Version 5.3-6.0) [Computer program]. http://www.praat.org

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America, 103*(6), 3153–3161. https://doi.org/10.1121/1.423073

Burnett, T. A., Senner, J. E., & Larson, C. R. (1997). Voice F_0 responses to pitch-shifted auditory feedback: A preliminary study. *Journal of Voice, 11*(2), 202–211. https://doi.org/10.1016/S0892-1997(97)80079-3

Cai, S., Beal, D. S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2014). Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter. *Brain and Language, 129,* 24–29. https://doi.org/10.1016/j.bandl.2014.01.002

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLOS ONE, 7*(7), Article e41830. https://doi.org/10.1371/journal.pone.0041830

Cai, S., Boucek, M., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2008, December 8–12). *A system for online dynamic perturbation of formant frequencies and results from perturbation of the Mandarin triphthong /iau/* [Paper presentation]. 8th International Seminar on Speech Production, Strasbourg, France.

Cai, S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2011). Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *Journal of Neuroscience, 31*(45), 16483–16490. https://doi.org/10.1523/JNEUROSCI.3653-11.2011

Caudrelier, T., & Rochet-Capellan, A. (2019). Changes in speech production in response to formant perturbations: An overview of two decades of research. In S. Fuchs, J. Cleland, & A. Rochet-Capellan (Eds.), *Speech production and perception: Learning and memory,* (pp.11–75). Peter Lang.

Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proceedings of the National Academy of Sciences, 110*(7), 2653–2658. https://doi.org/10.1073/pnas.1216827110

Chen, S. H., Liu, H., Xu, Y., & Larson, C. R. (2007). Voice *F* responses to pitch-shifted voice feedback during English speech. *The Journal of the Acoustical Society of America, 121*(2), 1157–1163. https://doi.org/10.1121/1.2404624

Chen, X., Zhu, X., Wang, E. Q., Chen, L., Li, W., Chen, Z., & Liu, H. (2013). Sensorimotor control of vocal pitch production in Parkinson's disease. *Brain Research, 1527,* 99–107. https://doi.org/10.1016/j.brainres.2013.06.030

Daliri, A., Chao, S.-C., & Fitzgerald, L. C. (2020). Compensatory responses to formant perturbations proportionally decrease as perturbations increase. *Journal of Speech, Language, and Hearing Research, 63*(10), 3392–3407. https://doi.org/10.1044/2020_JSLHR-19-00422

Daliri, A., & Dittman, J. (2019). Successful auditory motor adaptation requires task-relevant auditory errors. *Journal of Neurophysiology, 122*(2), 552–562. https://doi.org/10.1152/jn.00662.2018

Daliri, A., Wieland, E. A., Cai, S., Guenther, F. H., & Chang, S. E. (2018). Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter. *Developmental Science, 21*(2), Article e12521. https://doi.org/10.1111/desc.12521

Deroche, M. L., Nguyen, D. L., & Gracco, V. L. (2017). Modulation of speech motor learning with transcranial direct current stimulation of the inferior parietal lobe. *Frontiers in Integrative Neuroscience, 11,* Article 35. https://doi.org/10.3389/fnint.2017.00035

Dichter, B. K., Breshears, J. D., Leonard, M. K., & Chang, E. F. (2018). The control of vocal pitch in human laryngeal motor cortex. *Cell, 174*(1), 21–31.e9. https://doi.org/10.1016/j.cell.2018.05.016

Flinker, A., Chang, E. F., Kirsch, H. E., Barbaro, N. M., Crone, N. E., & Knight, R. T. (2010). Single-trial speech suppression of auditory cortex activity in humans. *Journal of Neuroscience, 30*(49), 16643–16650. https://doi.org/10.1523/JNEUROSCI.1809-10.2010

Floegel, M., Fuchs, S., & Kell, C. A. (2020). Differential contributions of the two cerebral hemispheres to temporal and spectral speech feedback control. *Nature Communications, 11*(1), Article 2839. https://doi.org/10.1038/s41467-020-16743-2.

Fujisaki, H. (2004). *Information, prosody, and modeling—With emphasis on tonal features of speech* [Paper presentation]. Speech Prosody 2004, International Conference.

Greenlee, J. D., Behroozmand, R., Larson, C. R., Jackson, A. W., Chen, F., Hansen, D. R., Oya, H., Kawasaki, H., & Howard, M. A., III. (2013). Sensory-motor interactions for vocal pitch monitoring in non-primary human auditory cortex. *PLOS ONE, 8*(4), Article e60783. https://doi.org/10.1371/journal.pone.0060783

Greenlee, J. D., Jackson, A. W., Chen, F., Larson, C. R., Oya, H., Kawasaki, H., & Howard, M. A., III. (2011). Human auditory cortical activation during self-vocalization. *PLOS ONE, 6*(3), Article e14744. https://doi.org/10.1371/journal.pone.0014744

Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., & Kenney, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the

audio-vocal reflex. *Experimental Brain Research, 130*(2), 133–141. https://doi.org/10.1007/s002219900237
Hilger, A. I. (2020). *Impaired sensorimotor integration for prosodic production in ataxic dysarthria*. Northwestern University.
Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H., & Konishi, J. (1997). Cortical processing mechanism for vocalization with auditory verbal feedback. *NeuroReport, 8*(9), 2379–2382. https://doi.org/10.1097/00001756-199707070-00055
Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure Voice-Related Quality of Life (V-RQOL). *Journal of Voice, 13*(4), 557–569. https://doi.org/10.1016/S0892-1997(99)80010-1
Houde, J. F., & Chang, E. F. (2015). The cortical computations underlying feedback control in vocal production. *Current Opinion in Neurobiology, 33,* 174–181. https://doi.org/10.1016/j.conb.2015.04.006
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. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science, 279*(5354), 1213–1216. https://doi.org/10.1126/science.279.5354.1213
Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience, 14*(8), 1125–1138. https://doi.org/10.1162/089892902760807140
Hunter, E. J., & Titze, I. R. (2009). Quantifying vocal fatigue recovery: dynamic vocal recovery trajectories after a vocal loading exercise. *Annals of Otology, Rhinology & Laryngology, 118*(6), 449–460.
Hutchins, S., & Peretz, I. (2013). Vocal pitch shift in congenital amusia (pitch deafness). *Brain and Language, 125*(1), 106–117. https://doi.org/10.1016/j.bandl.2013.01.011
Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of *F*0 production: Evidence from feedback perturbation. *The Journal of the Acoustical Society of America, 108*(3), 1246–1251. https://doi.org/10.1121/1.1288414
Kim, K. S., Daliri, A., Flanagan, J. R., & Max, L. (2020). Dissociated development of speech and limb sensorimotor learning in stuttering: Speech auditory-motor learning is impaired in both children and adults who stutter. *Neuroscience, 451,* 1–21. https://doi.org/10.1016/j.neuroscience.2020.10.014
Kim, K. S., & Max, L. (2020). *Speech auditory-motor adaptation lacks an explicit component: Reduced adaptation in adults who stutter reflects limitations in implicit sensorimotor learning*. bioRxiv. https://doi.org/10.1101/2020.09.06.284638
Kim, K. S., Wang, H., & Max, L. (2020). It's about time: Minimizing hardware and software latencies in speech research with real-time auditory feedback. *Journal of Speech, Language, and Hearing Research, 63*(8), 2522–2534. https://doi.org/10.1044/2020_JSLHR-19-00419
Kiran, S., & Larson, C. R. (2001). Effect of duration of pitch-shifted feedback on vocal responses in patients with Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 44*(5), 975–987. https://doi.org/10.1044/1092-4388(2001/076)
Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012
Kort, N. S., Nagarajan, S. S., & Houde, J. F. (2014). A bilateral cortical network responds to pitch perturbations in speech feedback. *NeuroImage, 86,* 525–535. https://doi.org/10.1016/j.neuroimage.2013.09.042
Korzyukov, O., Sattler, L., Behroozmand, R., & Larson, C. R. (2012). Neuronal mechanisms of voice control are affected by implicit expectancy of externally triggered perturbations in auditory feedback. *PLOS ONE, 7*(7), Article e41216. https://doi.org/10.1371/journal.pone.0041216
Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808814
Lametti, D. R., Quek, M. Y., Prescott, C. B., Brittain, J.-S., & Watkins, K. E. (2020). The perils of learning to move while speaking: One-sided interference between speech and visuomotor adaptation. *Psychonomic Bulletin & Review, 27*(3), 544–552. https://doi.org/10.3758/s13423-020-01725-8
Lametti, D. R., Smith, H. J., Freidin, P. F., & Watkins, K. E. (2018). Cortico-cerebellar networks drive sensorimotor learning in speech. *Journal of Cognitive Neuroscience, 30*(4), 540–551. https://doi.org/10.1162/jocn_a_01216
Larson, C. R., & Robin, D. A. (2016). Sensory processing: Advances in understanding structure and function of pitch-shifted auditory feedback in voice control. *AIMS Neuroscience, 3*(1), 22–39. https://doi.org/10.3934/Neuroscience.2016.1.22
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), 3628–3642. https://doi.org/10.1044/2020_JSLHR-20-00192
Li, W., Chen, Z., Liu, P., Zhang, B., Huang, D., & Liu, H. (2013). Neurophysiological evidence of differential mechanisms involved in producing opposing and following responses to altered auditory feedback. *Clinical Neurophysiology, 124*(11), 2161–2171. https://doi.org/10.1016/j.clinph.2013.04.340
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
Liu, H., Russo, N. M., & Larson, C. R. (2010). Age-related differences in vocal responses to pitch feedback perturbations: A preliminary study. *The Journal of the Acoustical Society of America, 127*(2), 1042–1046. https://doi.org/10.1121/1.3273880
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
Liu, H., Wang, E. Q., Metman, L. V., & Larson, C. R. (2012). Vocal responses to perturbations in voice auditory feedback in individuals with Parkinson's disease. *PLOS ONE, 7*(3). https://doi.org/10.1371/journal.pone.0033629
Liu, P., Chen, Z., Jones, J. A., Huang, D., & Liu, H. (2011). Auditory feedback control of vocal pitch during sustained vocalization: A cross-sectional study of adult aging. *PLOS ONE, 6*(7), Article e22791. https://doi.org/10.1371/journal.pone.0022791.
Liu, P., Chen, Z., Larson, C. R., Huang, D., & Liu, H. (2010). Auditory feedback control of voice fundamental frequency in school children. *The Journal of the Acoustical Society of America, 128*(3), 1306–1312. https://doi.org/10.1121/1.3467773
Loucks, T., Chon, H., & Han, W. (2012). Audiovocal integration in adults who stutter. *International Journal of Language & Communication Disorders, 47*(4), 451–456. https://doi.org/10.1111/j.1460-6984.2011.00111.x

MacDonald, E. N., Johnson, E. K., Forsythe, J., Plante, P., & Munhall, K. G. (2012). Children's development of self-regulation in speech production. *Current Biology, 22*(2), 113–117. https://doi.org/10.1016/j.cub.2011.11.052

MacDonald, E. N., & Munhall, K. G. (2012). *A preliminary study of individual responses to real-time pitch and formant perturbations* [Paper presentation]. The Listening Talker.

MathWorks. (2016). *MATLAB 2016 (Version b)*.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

Merriam-Webster. (n.d.). *Reliability.* Merriam-Webster.com Dictionary.

Mollaei, F., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2016). Sensorimotor control of vocal pitch and formant frequencies in Parkinson's disease. *Brain Research, 1646,* 269–277. https://doi.org/10.1016/j.brainres.2016.06.013

Mollaei, F., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2019). The relationship between speech perceptual discrimination and speech production in Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 62*(12), 4256–4268. https://doi.org/10.1044/2019_JSLHR-S-18-0425

Natke, U., Donath, T. M., & Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *The Journal of the Acoustical Society of America, 113*(3), 1587–1593. https://doi.org/10.1121/1.1543928

Nault, D. R., & Munhall, K. G. (2020). Individual variability in auditory feedback processing: Responses to real-time formant perturbations and their relation to perceptual acuity. *The Journal of the Acoustical Society of America, 148*(6), 3709–3721. https://doi.org/10.1121/10.0002923

Niziolek, C. A., & Guenther, F. H. (2013). Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *Journal of Neuroscience, 33*(29), 12090–12098. https://doi.org/10.1523/JNEUROSCI.1008-13.2013

Niziolek, C. A., & Parrell, B. (2021). Responses to auditory feedback manipulations in speech may be affected by previous exposure to auditory errors. *Journal of Speech, Language, and Hearing Research, 64*(6S), 2169–2181. https://doi.org/10.1044/2020_JSLHR-20-00263

Parkinson, A. L., Flagmeier, S. G., Manes, J. L., Larson, C. R., Rogers, B., & Robin, D. A. (2012). Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage, 61*(1), 314–322. https://doi.org/10.1016/j.neuroimage.2012.02.068

Parrell, B., Agnew, Z., Nagarajan, S., Houde, J., & Ivry, R. B. (2017). Impaired feedforward control and enhanced feedback control of speech in patients with cerebellar degeneration. *Journal of Neuroscience, 37*(38), 9249–9258. https://doi.org/10.1523/JNEUROSCI.3363-16.2017

Parrell, B., & Houde, J. (2019). Modeling the role of sensory feedback in speech motor control and learning. *Journal of Speech, Language, and Hearing Research, 62*(8S), 2963–2985. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0127

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), Article e1007321. https://doi.org/10.1371/journal.pcbi.1007321

Patel, S., Gao, L., Wang, S., Gou, C., Manes, J., Robin, D. A., & Larson, C. R. (2019). Comparison of volitional opposing and following responses across speakers with different vocal histories. *The Journal of the Acoustical Society of America, 146*(6), 4244–4254. https://doi.org/10.1121/1.5134769

Patel, S., Nishimura, C., Lodhavia, A., Korzyukov, O., Parkinson, A., Robin, D. A., & Larson, C. R. (2014). Understanding the mechanisms underlying voluntary responses to pitch-shifted auditory feedback. *The Journal of the Acoustical Society of America, 135*(5), 3036–3044. https://doi.org/10.1121/1.4870490

Perkell, J., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22*(2-3), 227–250. https://doi.org/10.1016/S0167-6393(97)00026-5

Purcell, D. W., & Munhall, K. G. (2006a). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America, 120*(2), 966–977. https://doi.org/10.1121/1.2217714

Purcell, D. W., & Munhall, K. G. (2006b). Compensation following real-time manipulation of formants in isolated vowels. *The Journal of the Acoustical Society of America, 119*(4), 2288–2297. https://doi.org/10.1121/1.2173514

Sares, A. G., Deroche, M. L., Shiller, D. M., & Gracco, V. L. (2018). Timing variability of sensorimotor integration during vocalization in individuals who stutter. *Scientific Reports, 8*(1), Article 16340. https://doi.org/10.1038/s41598-018-34517-1

Sares, A. G., Deroche, M. L. D., Ohashi, H., Shiller, D. M., & Gracco, V. L. (2020). Neural correlates of vocal pitch compensation in individuals who stutter. *Frontiers in Human Neuroscience, 14*. https://doi.org/10.3389/fnhum.2020.00018

Sato, M., & Shiller, D. M. (2018). Auditory prediction during speaking and listening. *Brain and Language, 187,* 92–103. https://doi.org/10.1016/j.bandl.2018.01.008

Scheerer, N. E, Jacobson, D. S., & Jones, J. A. (2016). Sensorimotor learning in children and adults: Exposure to frequency-altered auditory feedback during speech production. *Neuroscience, 314,* 106–115. https://doi.org/10.1016/j.neuroscience.2015.11.037

Scheerer, N. E., Jacobson, D. S., & Jones, J. A. (2020). Sensorimotor control of vocal production in early childhood. *Journal of Experimental Psychology: General, 149*(6), 1071–1077. https://doi.org/10.1037/xge0000706

Scheerer, N. E., Liu, H., & Jones, J. A. (2013). The developmental trajectory of vocal and event-related potential responses to frequency-altered auditory feedback. *European Journal of Neuroscience, 38*(8), 3189–3200. https://doi.org/10.1111/ejn.12301

Scott, T. L., Haenchen, L., Daliri, A., Chartove, J., Guenther, F. H., & Perrachione, T. K. (2020). Noninvasive neurostimulation of left ventral motor cortex enhances sensorimotor adaptation in speech production. *Brain and Language, 209,* Article 104840. https://doi.org/10.1016/j.bandl.2020.104840

Sengupta, R., Shah, S., Gore, K., Loucks, T., & Nasir, S. M. (2016). Anomaly in neural phase coherence accompanies reduced sensorimotor integration in adults who stutter. *Neuropsychologia, 93*(Pt. A), 242–250. https://doi.org/10.1016/j.neuropsychologia.2016.11.004

Shadmehr, R., & Holcomb, H. H. (1997). Neural correlates of motor memory consolidation. *Science, 277*(5327), 821–825. https://doi.org/10.1126/science.277.5327.821

Shum, M., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2011). Sensorimotor integration for speech motor learning involves the inferior parietal cortex. *European Journal of Neuroscience, 34*(11), 1817–1822. https://doi.org/10.1111/j.1460-9568.2011.07889.x

Stepp, C. E., Lester-Smith, R. A., Abur, D., Daliri, A., Pieter Noordzij, J., & Lupiani, A. A. (2017). Evidence for auditory-motor impairment in individuals with hyperfunctional voice disorders. *Journal of Speech, Language, and Hearing Research, 60*(6), 1545–1550. https://doi.org/10.1044/2017_JSLHR-S-16-0282

Tang, D.-L., McDaniel, A., & Watkins, K. E. (2021). Disruption of speech motor adaptation with repetitive transcranial magnetic stimulation of the articulatory representation in primary motor cortex. *Cortex, 145,* 115–130. https://doi.org/10.1016/j.cortex.2021.09.008

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language & 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

van Brenk, F., & Terband, H. (2020). Compensatory and adaptive responses to real-time formant shifts in adults and children. *The Journal of the Acoustical Society of America, 147*(4), 2261–2270. https://doi.org/10.1121/10.0001018

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., Abur, D., Enos, N. M., Brown, K. M., & Stepp, C. E. (2020). Auditory-motor perturbations of voice fundamental frequency: Feedback delay and amplification. *Journal of Speech, Language, and Hearing Research, 63*(9), 2846–2860. https://doi.org/10.1044/2020_JSLHR-19-00407

Weerathunge, H. R., Tomassi, N. E., & Stepp, C. E. (2022). What can altered auditory feedback paradigms tell us about vocal motor control in individuals with voice disorders? *Perspectives of the ASHA Special Interest Groups, 7*(3), 959–976. https://doi.org/10.1044/2022_PERSP-21-00195

Weerathunge, H. R., Voon, T., Tardif, M., Cilento, D., & Stepp, C. E. (2022). Auditory and somatosensory feedback mechanisms of laryngeal and articulatory speech motor control. *Experimental Brain Research, 240*(7–8), 2155–2173. https://doi.org/10.1007/s00221-022-06395-7

Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage, 40*(4), 1871–1887. https://doi.org/10.1016/j.neuroimage.2008.01.026