ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Practice Parameters That Influence Vocal Motor Adaptation in Altered Auditory Feedback Paradigms

Allison S. Aaron,[a] Saul A. Frankford,[a,b,c] Kaitlyn A. Siedman,[a] Edwin Maas,[d] Frank H. Guenther,[a,e] and Cara E. Stepp[a,e,f]

[a] Department of Speech, Language and Hearing Sciences, Boston University, MA [b] Department of Speech, Language, and Hearing, The University of Texas at Dallas, Richardson [c] Callier Center for Communication Disorders, The University of Texas at Dallas, Richardson [d] Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA [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 December 9, 2024
Revision received July 10, 2025
Accepted October 27, 2025

Editor-in-Chief: Jessica E. Huber

https://doi.org/10.1044/2025_JSLHR-24-00869

ABSTRACT

Purpose: The purpose of this study was to determine how the structure of vocal practice influences vocal motor adaptation.
Method: Thirty-two speakers completed four altered auditory feedback paradigms in which their fundamental frequency (f_o) was systematically shifted down by one semitone in a single session. Paradigms were compared to isolate the influence of two practice parameters—vowel and intertrial interval (ITI) duration—on vocal motor learning, as measured by vocal responses to predictable altered auditory feedback (adaptive responses). Adaptive responses were measured during and after prolonged exposure to f_o-shifted auditory feedback and compared across paradigms. We hypothesized that adaptive responses would be increased with longer vowel durations and longer ITI durations.
Results: During exposure to altered auditory feedback, adaptive responses were increased only with longer vowel durations. After exposure to the altered auditory feedback, adaptive responses were increased with both longer vowel durations and ITI durations.
Conclusions: Vocal motor adaptation is increased when speakers have more time within a vocalization to access and integrate feedback into future vocalizations. Longer ITI durations, and thus, rest times between vocalizing (i.e., distributed practice), positively influence retention.
Supplemental Material: https://doi.org/10.23641/asha.31151785

Speech scientists have attempted to better understand the complex processes of how humans monitor and regulate their speech to communicate effectively. Altered auditory feedback has been used in experiments to study the underlying mechanisms of speech production (Burnett et al., 1997; Jones & Munhall, 2000). These experiments have potential implications for understanding speech motor learning in the context of speech and voice therapy. In these studies, participants are asked to speak aloud as they are presented with their own amplified speech signal (i.e., auditory feedback) through headphones; the speech signal is manipulated and the altered auditory signal is transmitted to the speaker in near–real time via the headphones, thus altering the speaker's auditory feedback. Researchers have manipulated vowel formant frequencies to measure control within the articulatory domain (e.g., Houde & Jordan, 1998), as well as acoustic features within the voice domain. For example, the fundamental frequency (f_o) of the auditory feedback can be shifted (e.g., Burnett et al., 1997). When speakers are presented with the altered auditory feedback as they are vocalizing, there is a resulting mismatch between the speaker's intended and perceived pitch for the vocal production. In response, speakers typically compensate by opposing the shift (e.g., if the f_o is shifted down, they respond by increasing their f_o). These f_o-shift responses are used to elucidate the interacting mechanisms responsible for monitoring and regulating speech (Larson, 1998).

Several models of articulatory motor control (Houde & Nagarajan, 2011; Parrell et al., 2019; Tourville & Guenther,

---

Correspondence to Allison S. Aaron: aaaron@bu.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

2011) and vocal motor control (Weerathunge, Alzamendi, et al., 2022) have used experimental data to theorize the interacting control mechanisms for speech production. In this article, we will use the Directions into Velocities of Articulators model as a framework for interpretation of results from altered auditory feedback paradigms (Tourville & Guenther, 2011). There are two standard types of altered auditory feedback paradigms, *reflexive* and *adaptive* paradigms, which focus on two aspects of speech motor control: feedback control and feedforward control. In a reflexive paradigm, f_o shifts are applied unexpectedly, and speakers typically respond within the same production with a rapid pitch correction in the opposite direction. These rapid responses can be used to measure the control process of using feedback to correct for errors within a production (i.e., feedback control). In contrast, when f_o shifts are predictably applied across successive productions in an adaptation paradigm, the speakers *adapt* over time by integrating the feedback into their motor plans for subsequent vocalizations (i.e., refining their feedforward control), thus eliciting adaptive responses (Tourville & Guenther, 2011).

Vocal motor adaptation is the result of an internal process in which speakers update their motor plans in response to exposure to sensory errors. When adult speakers with typical voices are exposed to errors during adaptive paradigms, we can study that internal process. There is conflicting evidence of the influence of explicit factors (e.g., attention, cognitive load) on adaptation across domains (i.e., limb, articulation, and pitch control; Lametti et al., 2020; Scheerer, Tumber, & Jones, 2016; Tomassi et al., 2025; T. Wang, Li, & Ivry, 2024). However, vocal motor adaptation is thought to involve an implicit component of vocal motor learning (Krakauer et al., 2019). Given that implicit vocal motor learning is difficult to isolate, f_o adaptation paradigms offer a valuable approach to investigate this core component of changing and maintaining new vocal motor behaviors.

It is challenging to compare and interpret vocal motor learning findings across prior altered auditory feedback studies due to the significant variability in the parameters used in different paradigms throughout the literature (Miller et al., 2023). Traditionally, during adaptation paradigms, speakers produce repeated trials of vocalization across experimental runs that consist of sequential phases. Phases within a run often include a baseline phase during which speakers vocalize as they are presented with their unaltered feedback via headphones in near–real time, a hold phase during which the auditory feedback is manipulated, and an after-effect phase during which the feedback returns to its unaltered state. Vocal responses during the hold phase (i.e., when the feedback is altered) are used to measure speakers' adaptation during error-exposed practice, and vocal responses during the after-effect phase are used to measure any lasting adaptation after error-exposed practice. One parameter that varies across adaptation paradigms is how the altered auditory feedback is systematically presented to the speaker. Across prior f_o and oral articulation adaptation studies, some experiments include the use of a gradual, or ramped, shift in auditory feedback between the baseline and hold phases (Abur et al., 2021; Daliri & Dittman, 2019; Jones & Munhall, 2000; Villacorta et al., 2007), and others present the altered feedback suddenly (Acosta et al., 2023; Keough et al., 2013; Scheerer, Jacobson, & Jones, 2016). One study found that adaptive responses differ based on whether the errors are introduced gradually or suddenly (Chao & Daliri, 2023), suggesting that these varied parameters influence adaptation.

There are several other parameters that differ across f_o adaptation experiments. Reported vowel durations per trial typically range from ~2 (Behroozmand & Sangtian, 2018; Jones & Keough, 2008) to 3 s (Abur et al., 2021; Keough et al., 2013), and they have been as high as 5.4 s (Scheerer, Tumber, & Jones, 2016); however, these estimates may only reflect the experimenters' instructions rather than the actual speakers' vocalizations. Thus, true vowel durations likely vary even more widely than what is reported. Further, vowel duration estimates are not reported for every adaptation study. Similarly, the duration of rest time between trials is not always reported. However, based on the studies that do report it, rest times also vary, thus altering the distribution of trials. The rest times in adaptation runs have been as short as 1.5 s (Weerathunge et al., 2020), leading to a condensed distribution of trials (i.e., massed practice) and as long as 8 s (Abur et al., 2018), leading to an extended distribution of trials (i.e., distributed practice).

The lack of a standardized method across experiments makes it difficult to interpret experimental results. For example, if the vocalization duration per trial is too short, it is possible that speakers might not have enough time to access their feedback. Furthermore, if practice distribution is too condensed (i.e., rest time between trials is short), speakers may not have enough time between trials to update their future motor plans. Thus, due to the wide variation of these parameters in previous work, it is difficult to compare findings from differing adaptation studies and difficult to apply results to the development of clinical voice therapy protocols and guidelines. To better interpret the findings of altered auditory feedback paradigms, more work is needed to determine which variables are influencing adaptation. From a modeling perspective, we would expect that more practice with exposure to sensory error (longer vocalization durations) and thus, more time to access and integrate feedback into feedforward commands for subsequent trials would result in greater adaptation magnitudes (Tourville & Guenther, 2011). The influence of how trials are distributed among rest times (i.e., distributed

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

practice with long rest times vs. massed practice with short rest times) on adaptation is not clearly delineated in speech motor control models. A direct comparison of adaptive responses across paradigms with differing vowel durations and distribution of trials would elucidate how practice amount and distribution affect sensorimotor adaptation.

Most of the evidence for the role of practice amount (i.e., the amount of time spent practicing movements) and practice distribution (i.e., how practice is distributed over time) has been derived from nonspeech motor learning studies (Baddeley & Longman, 1978; Brashers-Krug et al., 1996; C. H. Shea et al., 1990, 2000; see Maas et al., 2008, and Bislick et al., 2012, for reviews). There is evidence suggesting that practice amount impacts nonspeech motor adaptation. One study on a cursor movement task found that longer trial durations reduced error rates, suggesting that more practice (and more exposure to the error) resulted in greater adaptation (Hardwick et al., 2017). Additionally, several visuomotor studies have found that online feedback (i.e., when feedback is continuously provided throughout a trial) promotes more adaptation, as compared to posttrial feedback (i.e., when feedback is only provided after the movement trial; Barkley et al., 2014; Hinder et al., 2008; Schween et al., 2014). However, it is unclear whether this result is related to the increased practice amount with online feedback or the difference in the timing of feedback between online and posttrial feedback (T. Wang, Avraham, et al., 2024). Furthermore, there is literature showing that online feedback enhances within-session adaptation but degrades learning postsession (i.e., retention; Salmoni et al., 1984; Schmidt & Wulf, 2002). Since vocal pitch control as evaluated in adaptation paradigms may function more as a postural control system (e.g., repeatedly sustaining a single pitch) than a target-oriented movement (e.g., reaching), changes in trial duration may have distinct effects on vocal pitch control compared to target-based movements.

Practice distribution has also been found to influence sensorimotor adaptation for nonspeech motor movements (Bock et al., 2005; Brashers-Krug et al., 1996; Francis, 2005; Huang & Shadmehr, 2007; Shadmehr & Holcomb, 1997). In one sensorimotor adaptation study of limb control for a pointing task, adaptation was increased when trials were spaced out with longer rest breaks; there was a significant difference between 1-s breaks between trials (intertrial interval [ITI]), compared to ≥ 5-s breaks (Bock et al., 2005). This finding was replicated in another limb control study; a rest break of less than 5 s hampered early performance on a reaching task (Francis, 2005). These two findings were replicated by Huang and Shadmehr (2007); further, these authors posited that adaptation is enhanced after ~4 s between trials due to the fading of the error memory. Another study used a visuomotor task to compare short (5.2 s) rest times versus long (18.4 s) rest times between trials (Kim et al., 2015). Compared to the long rest time condition, the short rest time condition had faster adaptation during training but larger immediate forgetting in the retention tests, suggesting that rest breaks may influence retention more than performance.

It is unclear the extent to which sensorimotor adaptation contributes to vocal motor learning in a clinical context, which likely involves a complex intersection of both internal and external learning mechanisms (Krakauer et al., 2019). However, adaptation paradigms as models for the study of one component of vocal motor learning (i.e., related to the recalibration of movements) have the potential to inform clinical voice treatment recommendations (Weerathunge, Tomassi, & Stepp, 2022); in voice therapy, the overarching goal is to shape and change an individual's vocal motor behaviors, a process that requires implicit vocal motor learning (Tellis, 2018). Principles of motor learning for speech have been derived primarily from nonspeech motor studies (Maas et al., 2008). Nonspeech motor findings support the benefit of high overall practice amount as compared to low practice amount (Maier et al., 2019; Park & Shea, 2005; C. H. Shea & Kohl, 1991). Additionally, studies have found increased motor learning with distributed practice (when practice trials are spaced out from one another) as compared to massed practice (Lee & and Genovese, 1988; Maier et al., 2019; C. H. Shea et al., 2000). The evidence from the speech domain is limited (Bislick et al., 2012; Maas et al., 2008, 2019). In the articulatory domain, more overall practice has been shown to improve speech therapy outcomes (Allen, 2013; Edeal & Gildersleeve-Neumann, 2011; Maas et al., 2008, 2019; Namasivayam et al., 2015). Contrary to the findings in the nonspeech motor literature, Maas et al. (2019) found that massed, rather than distributed, practice was associated with increased motor learning in treatment for childhood apraxia of speech. Another study found that a more distributed practice schedule, which also had a shorter total practice amount, resulted in better speech outcomes for adults with acquired apraxia of speech (Wambaugh et al., 2018). However, this study did not control for practice amount, making it difficult to interpret which parameter might be influencing results.

The terms "practice amount" and "practice distribution" have broad definitions with several meanings. "Practice amount" can be measured as a speaker's total time directly practicing a motor task during a practice session, but also as the number of practice sessions or practice trials (Maas et al., 2008). "Practice distribution" also has several meanings: It can be defined as how a given number of trials or sessions is distributed over time. Given this ambiguity, for the current study, we focused on two practice parameter comparisons in the context of a single practice session (i.e., an altered auditory feedback experiment) with the understanding that this

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

study will only assess these specific components within the umbrella terms of practice amount and practice distribution. Vowel duration was defined as the length of vocalization within each trial, and ITI duration was defined as the time between successive voicing onsets, which scales with the amount of rest time between trials.

The purpose of this study was to determine how practice parameters influence the magnitude of vocal motor adaptation in altered auditory feedback paradigms in the context of motor learning *during* a practice session as well as retention *after* a session. We used experimental methodology from the speech motor control literature to study the influence of practice parameters on vocal motor adaptation. Participants completed four f_o adaptation conditions which varied from one another in: (a) vowel duration and (b) ITI duration (see Figure 1). We measured adaptive f_o response magnitudes during exposure to altered auditory feedback and after the altered auditory feedback was removed across the conditions. Based on both the motor control and the clinical speech motor learning literature, we hypothesized that f_o adaptive response magnitudes would be increased during and after exposure to the altered auditory feedback with longer vowel and ITI durations as compared to shorter vowel and ITI durations. The results of this project have both experimental implications for vocal motor control research and clinical implications for dosage and service delivery recommendations for voice therapy.

Method

Participants

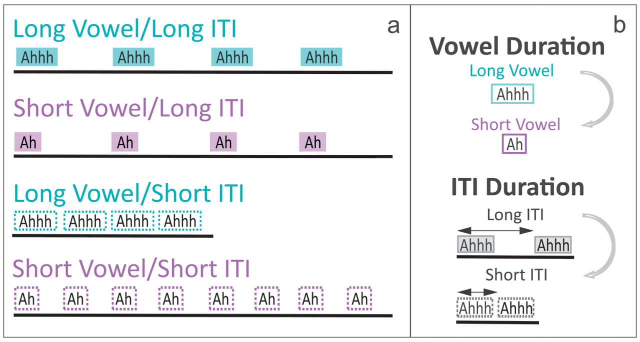
Thirty-two participants (15 cisgender women, 13 cisgender men, two transgender women, one nonbinary individual assigned female at birth, and one nonbinary individual assigned male at birth) aged 18 to 29 years (*M* = 21.8 years; *SD* = 3.2 years) participated in this study. All participants were fluent speakers of English. No individuals had a history of neurological, voice, speech, language, or hearing disorders, nor did they have a history of voice therapy or voice training. Individuals who reported more than 1 year of formal organized singing experience were excluded given the influence of singing experience on vocal motor control (Jones & Keough, 2008; W. Wang et al., 2019). Participants completed written consent in compliance with the Boston University Institutional Review Board (Protocol #2625). All included participants were required to pass a pulsed-tone hearing screening (American Speech-Language-Hearing Association, 2018). The hearing screening was conducted using over-ear headphones on a GSI 18 audiometer (Grason-Stadler). All participants had hearing thresholds at or below 25 dB HL at 250, 500, 1000, 2000, and 4000 Hz (American Speech-Language-Hearing Association, 2005).

Data Collection

Experimental Setup

Participants completed experimental tasks in a sound-attenuated booth at Boston University in a 2.5-hr session. They sat in front of a computer monitor and wore an omnidirectional ear-set microphone (MX153; Shure) placed approximately 45° from the midline and 7 cm from the corner of the mouth (Patel et al., 2018). During experimental tasks, participants wore Etymotic ER-2 insert earphones. The microphone signal was amplified with a preamplifier (RME Quadmic II) and digitized with a soundcard (RME Fireface UCX). The signal was transmitted through the Eventide Eclipse V4 Harmonizer. The processed signal was conditioned with another preamplifier (Behringer Xenyx Q802) and near–

Figure 1. (a) Schema of the four adaptation conditions to compare the influence of two practice parameters, vowel duration and intertrial interval (ITI) duration, on adaptation. ITI duration is defined as the time between trial onsets, which scales with amount of rest time between trials. The two long vowel conditions are in light blue and the two short vowel conditions are in purple. The long ITI duration conditions are represented with solid filled boxes, and the two short ITI duration conditions are represented with dashed unfilled boxes. (b) Comparison of vowel duration (long vs. short vowels) and ITI duration (long vs. short ITIs).



Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 1. Comparison of practice parameters (vowel duration and intertrial interval [ITI] duration) that are hypothesized to influence adaptation across four conditions.

Condition	Practice parameters		Other experimental variables			
	Vowel duration (seconds)	ITI duration (seconds)	Added inter-trial jitter (seconds)	Total speaking time (seconds)	Paradigm length (minutes)	Number of trials
Long Vowel/ Long ITI	**1.2**	**~6.00**	Range: 4.3–5.3; *M* = 4.80	~77	6	64
Short Vowel/ Long ITI	*0.6*	**~6.00**	Range: 4.9–5.9; *M* = 5.40	~38	6	64
Long Vowel/ Short ITI	**1.2**	*~3.05*	Range: 1.6–2.1; *M* = 1.85	~77	3	64
Short Vowel/ Short ITI	*0.6*	*~3.05*	Range: 2.2–2.7; *M* = 2.45	~77	6	128

Note. The comparison values for the two practice parameters are emphasized; long vowel and ITI durations are bolded and short vowel and ITI durations are italicized. Several other experimental variables (total speaking time, total length of the paradigm, and number of trials) that are less likely to influence adaptation varied in order to create the conditions.

real time feedback was presented through the earphones as participants spoke.[1] Before the microphone and earphones were placed, the experimenter completed a calibration protocol. The amplitude of the earphone output was amplified 5 dB above the microphone signal to reduce the impact of air- and bone-conducted feedback (Weerathunge et al., 2020). This calibration was completed using a sound level meter (Brüel and Kjaer Type 2250-L with a Type 4947 1/2” Pressure Field Microphone) with an insert earphone adapter (Type 4946 2CC Click-on Coupler).

Experimental Protocol

Participants completed four f_o adaptation conditions with two experimental runs per condition: shift down and control. Downward shifts were applied, rather than upward shifts, because we expected speakers’ baseline f_o productions to be close to their lower f_o boundary (pitch floor), allowing for larger response magnitudes in the opposing upward direction (Dahl et al., 2024). For each condition, participants were instructed to sustain the vowel /ɑ/ for the duration of time that the written cue, “aaa,” was presented on the screen. They were asked to keep their vocalization as steady as possible across each vowel and throughout each experimental run. Using a custom MATLAB script, the duration of the vowel production was controlled by visual cues on the screen, which differed across conditions. The added time between trials was jittered by varying amounts, depending on the condition (see Table 1). Participants completed six practice trials (no shifted feedback) prior to initiating each condition.

For all experimental runs, the trials were divided into three sequential phases: baseline, hold, and after-effect.[2] For the shift down experimental runs, during the baseline phase (25% of trials; either 16 or 32 trials depending on the condition), participants vocalized as they were presented with their unaltered feedback in near–real time. During the hold phase (37.5% of trials; either 24 or 48 trials depending on the condition), the participants’ feedback was shifted down by 100 cents (one semitone).[3] These hold phase lengths are similar to prior f_o adaptation paradigm studies with short experimental runs, and thus, short hold phases (Heller Murray & Stepp, 2020; Scheerer, Jacobson, & Jones, 2016). Further, based on pilot testing, the authors confirmed that the hold phases for all four conditions were long enough to elicit sufficient adaptive response magnitudes. During the after-effect phase (37.5% of trials; either 24 or 48 trials), the feedback returned to its unaltered state. For control runs, no f_o shift was applied in any of the phases. To control for order effects, four condition orders were generated using a balanced Latin square design (Bradley, 1958). Within each condition, half of the participants completed the shift down experimental run first, and the other half completed the control run first, resulting in a total of eight possible experimental orders. These eight orders were evenly counterbalanced across participants. Therefore, for all eight orders, the experimental runs alternated

---

[1] The processing delay from the Eclipse Eventide Harmonizer of ~11 ms (Heller Murray et al., 2019) is not expected to affect results of altered auditory feedback paradigms (Weerathunge et al., 2020).

[2] No ramp phase was used, given that prior literature has shown that speakers demonstrate adaptive responses without use of a ramp (Keough et al., 2013).

[3] The shift was applied to the full spectrum of the audio signal, which resulted in changes to both f_o and formant frequencies. However, we do not expect the small formant changes to influence results, given the relatively small magnitude of the shift and the use of a sustained /ɑ/ vowel, which was unlikely to cross a perceptual categorical boundary (Heller Murray et al., 2019).

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

between the shift down and control runs to reduce the possibility that adaptive responses from a prior condition would carry over into the following condition. Additionally, participants completed a washout task between each condition, during which they read aloud two Harvard sentences (Institute of Electrical and Electronics Engineers, 1969) as they were presented with their unaltered feedback, which provided an opportunity for participants to speak aloud within a typical communicative context for several seconds with unaltered auditory feedback.

The four conditions were designed to isolate the influence of two practice parameters: vowel duration and ITI duration. As illustrated in Figure 1, the four comparison conditions were (a) Long Vowel/Long ITI, (b) Short Vowel/Long ITI, (c) Long Vowel/Short ITI, and (d) Short Vowel/Short ITI. The distribution of mean vowel durations per experimental run across the four conditions for all participants was visualized (see Supplemental Material S1). This visualization confirms that practice variables were manipulated as intended, with mean vowel durations at ~1.2 s for the two conditions with long vowels and ~0.6 s for the two conditions with short vowels. The amount of added time between trials was jittered across varied ranges for the four conditions to avoid rhythmic cues, and the ITI duration was calculated as the duration between voicing onsets. Therefore, the mean ITI durations for the two long ITI conditions were approximately 6 s, and the mean ITI durations for the two short ITI conditions were approximately 3 s (see Table 1 for reference). Several other practice parameters (total speaking time, number of trials, and paradigm length) were varied to create the conditions and are reported in Table 1. Based on a priori hypotheses built on prior nonspeech and speech motor learning literature, these other parameters were not expected to have a strong influence on adaptive responses.[4]

Data Analysis

A trained researcher blinded to the study hypotheses used a custom MATLAB script (Version 2022a, Mathworks, Inc.) to visualize the acoustic waveform, spectrogram, and f_o trace of the microphone signal for each trial using data extracted from Praat (Boersma & Weenink, 2008). They listened to and visually inspected each trial to confirm usability. Unusable trials (e.g., speech errors, coughs, yawns) were removed, excluding 391 trials (1.6%) and leaving a final data set of 24,185 trials. The researcher used the script to manually mark the beginning and end of the utterance for each trial across both experimental runs (shift down and control) for each condition for each participant. The total vocalization time and the f_o of an early analysis window (i.e., the 40–120 ms section after vocalization onset) of each trial were then extracted from each trial. The early analysis window was chosen to replicate analysis protocols from prior adaptation studies to assess feedforward-driven f_o adaptation after the participant's voice stabilized but before auditory feedback responses were expected to occur (Abur et al., 2021; Burnett et al., 1998; Kapsner-Smith et al., 2024). If the trained researcher observed trials with tracking errors in the f_o trace, the f_o tracking settings for those trials were adjusted manually in Praat, and the f_o of the early analysis window was reanalyzed manually.

The f_o adaptive response was constructed by converting the f_o in hertz (Hz) from the early analysis window of each trial to cents using the average of the trials in second half of the baseline phase (the late baseline phase section) as the reference frequency.[5] The late baseline phase section was used as speakers' true baseline f_o to capture their most stabilized baseline productions. The mean f_o from the control run trials was then subtracted from the corresponding trial in the shift down run for all conditions to account for natural f_o variability across extended speech production (Boominathan et al., 2010; Remacle et al., 2012; Stemple et al., 1995).

To visualize and compare adaptive responses across all four conditions, the f_o adaptive response of each trial was averaged across phase sections for each experimental run. This approach allowed for section alignment across all four conditions and observation of beginning and end points of learning in the hold and after-effect phases. The baseline phase section was divided into the early baseline phase section (first half of baseline trials) and the late baseline phase section (second half of baseline trials). The trials in the hold and after-effect phases were divided into thirds, with phase sections including early hold, mid hold, late hold, early after-effect, mid after-effect, and late after-effect sections. To visualize the temporal evolution of learning at a more fine-grained resolution, we also plotted the results of the four conditions at the block level (i.e., averaged every four trials; see Supplemental Material S2). This visualization confirms that learning reached a plateau for all conditions, including the Short Vowel/Short ITI

---

[4] Prior f_o adaptation studies have had a varying number of trials from as few as 30 trials per experimental run (Scheerer, Jacobson, et al., 2016) to as large as 480 trials per experimental run (R. Patel et al., 2011). Across studies, f_o adaptive responses have been observed plateau quickly even with a small number of trials in the hold phase (Heller Murray & Stepp, 2020; Scheerer, Jacobson, et al., 2016). Therefore, we did not expect the number of trials to influence pitch adaptation. However, we interpreted results with these varied parameters in mind.

[5] f_o (cents) = 1200 x $\log_2(f_1/f_2)$; f_1 is the f_o (Hz) of a given trial, and f_2 is the reference frequency (Hz) from the average of the late baseline phase section.

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

condition with double the number of trials compared to the other three conditions.

To quantify vocal motor adaptation, mean adaptive responses were taken from the control-normalized late hold and early after-effect phase sections. Because this study paradigm did not have a ramp phase, the late hold phase section, rather than the full hold phase, was used to measure the end point of learning. As per the literature, mean responses were included for all participants, regardless of whether they opposed the downward perturbation (with an upward f_o response), followed the perturbation (with a downward f_o response), or did not adjust to the perturbation (Miller et al., 2023). The late hold and early after-effect phase sections were used to interpret adaptation in the context of motor learning *during* a practice session (late hold phase) as well as retention immediately *after* a session (early after-effect).

Statistical Analysis

All statistical analyses were completed in *R* statistical software (RStudio Team, 2020). Significance for all statistical testing was set a priori at $\alpha = .05$. We constructed a linear mixed-effects model using the lmer package in RStudio to measure the main effects of experimental phase section (late baseline, late hold, and early after-effect), vowel duration, ITI duration, and their interactions on adaptive responses. Participant was entered as a random factor. The model specification was the following: AdaptiveResponses ~ PhaseSection * VowelDuration * ITIDuration + (1|Participant). Effect sizes were calculated as a partial squared curvilinear correlation (η_p^2) and designated as small (.01), medium (.09), or large (.25; Witte & Witte, 2017). To test whether there was a difference in f_o across phase sections in the control runs, we constructed a linear mixed-effects model using the average f_o data from the control runs as the outcome variable. Model specifications were as follows: lmer (ControlRun_Mean f_o ~ PhaseSection + (1|PtID).[6]

Post hoc pairwise comparison tests were calculated using the emmeans package in RStudio to compare the effects of vowel duration and ITI duration on adaptive responses in the late hold and early after-effect sections. To account for multiple comparisons, *p* values were adjusted using Tukey's method. Effect sizes were calculated as a Cohen's *d* and designated as small (0.2), medium (0.5), or large (0.8; Witte & Witte, 2017).

Results

On average, across all four conditions, participants adapted by opposing the 100-cent downward shift in the hold phase (see Figure 2 and Supplemental Material S2). The linear mixed-effects model revealed a statistically significant main effect of phase section ($F = 155.39$, $p < .001$) on the magnitude of adaptive responses, with a medium effect size ($\eta_p^2 = .08$). There were significant main effects of vowel duration ($F = 40.43$, $p < .001$) and ITI duration ($F = 8.83$, $p < .001$) with small effect sizes ($\eta_p^2 = .01$ and .002, respectively). There were also statistically significant interactions between phase section and vowel duration ($F = 11.92$, $p < .001$) and phase section and ITI duration ($F = 4.64$, $p < .001$) on adaptive responses with small effect sizes ($\eta_p^2 = .006$ and .003, respectively). There was no significant three-way interaction between vowel duration, ITI duration, and phase section.

In the late hold phase section, adaptive responses were significantly increased with longer compared to shorter vowel durations ($p < .001$) with a small-medium effect size ($d = 0.36$). There was no statistically significant difference between adaptive responses in the late hold phase section when comparing ITI duration. In the early after-effect phase section, adaptive responses were significantly increased with longer vowel duration ($p < .001$) and longer ITI duration ($p < .001$) relative to short vowel and ITI durations, with small effect sizes ($d = 0.24$ and 0.23, respectively; see Figure 3 and Table 2).

Discussion

The purpose of this study was to investigate the influence of two practice parameters—vowel duration and ITI duration—on adaptive responses in altered auditory feedback paradigms. We calculated adaptive responses in the late hold phase section to measure adaptation in the context of learning a task (i.e., acquisition) and in the early after-effect phase section to interpret retention and generalization of the learned vocal motor behavior. In both the late hold and early after-effect phase sections, adaptive responses were significantly increased with longer compared to short vowel durations, which was consistent with our hypothesis. Adaptive responses were also significantly increased with longer compared to short ITI durations (i.e., distributed practice) in the early after-effect phase section, which aligned with our hypothesis. However, no significant difference was observed when ITI durations varied in the late hold phase section.

Vowel Duration

Adaptive response magnitudes were increased with longer vowel durations per trial in both the late hold and

---

[6] There was a significant effect of phase section on mean f_o in the control runs, such that mean f_o was increased from baseline in the early after-effect phase section. This finding of a natural f_o drift over time aligns with prior work (Boominathan et al., 2010; Remacle et al., 2012; Stemple et al., 1995).

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 2. Results of post hoc pairwise comparisons of vowel duration and intertrial interval (ITI duration) across the phase sections of interest.

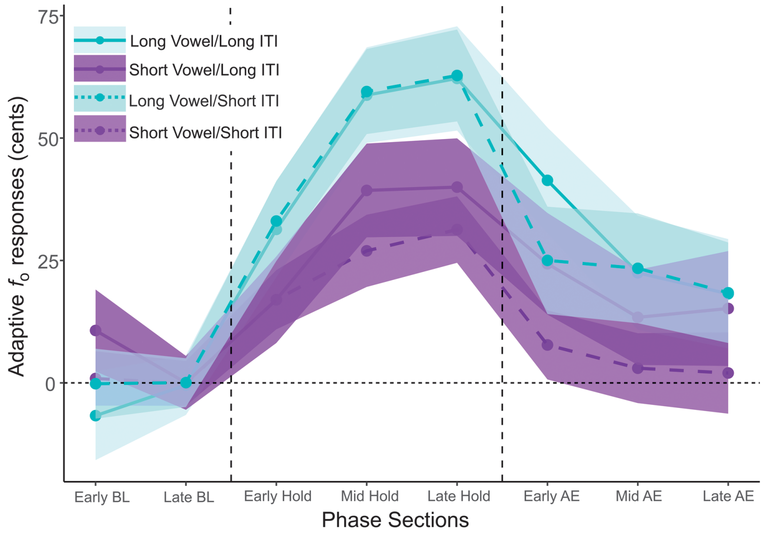
Phase section	Comparison	*p*	Cohen's *d*	Effect size	Direction
Late hold	Long vowel–short vowel	< .001*	0.36	Small-medium	Long > short
	Long ITI–short ITI	.328			
Early after-effect	Long vowel–short vowel	.027*	0.24	Small	Long > short
	Long ITI–short ITI	< .001*	0.23	Small	Long > short

*Significant at α = .05.

early after-effect phase sections. This finding suggests that vowel duration should be controlled when comparing adaptation magnitudes across differing altered auditory feedback studies. Further, the finding aligns with the nonspeech motor literature evidence of more adaptation with longer trial durations (Hardwick et al., 2017). Since vowel duration relates closely with the motor learning principle of practice amount, this finding unsurprisingly also aligns with more general nonspeech motor findings that higher practice amount (i.e., via total practice amount or treatment intensity) improves motor learning compared to low practice amounts (Park & Shea, 2005; C. H. Shea & Kohl, 1991; Yamada et al., 2019). The present study result is also consistent with findings from the speech motor literature suggesting that more overall practice improves speech therapy outcomes (Kaipa & Peterson, 2016; Maas et al., 2019).

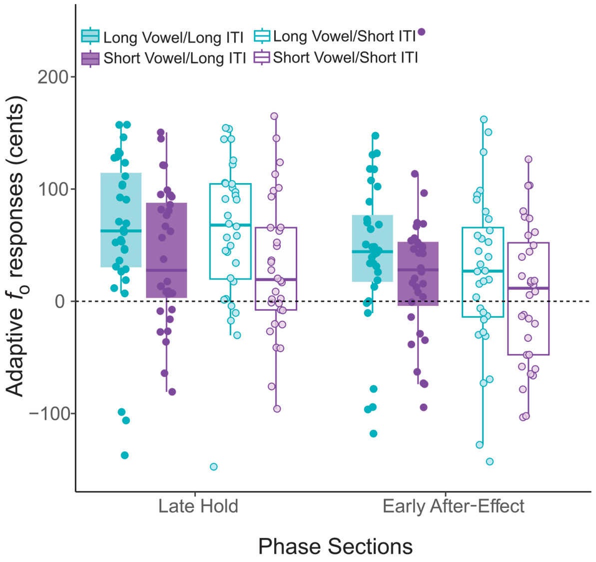
The result of increased adaptation with longer vowel durations may be driven by the difference in the amount of time per trial that participants had to access and integrate sensory feedback. This interpretation aligns with prior nonspeech adaptation studies that observed a learning benefit to continuously provided sensory feedback, compared to post-trial feedback (Barkley et al., 2014; Hinder et al., 2008; Schween et al., 2014). In speech motor control research, auditory feedback control is thought to come online after an approximately 120-ms delay (Burnett & Larson, 2002; Burnett et al., 1997; Larson et al., 2001; Tourville et al., 2008); thus, the amount of time that speakers had to notice and integrate their feedback into subsequent trials differed between the long vowel and short vowel conditions. In the two long vowel duration conditions, participants had ~1,080 ms per trial to access and integrate their feedback (1.2-s vowel duration − 120-ms delay = 1,080 ms) across the trials in the hold phase. In comparison, in the two short vowel conditions, participants only had ~480 ms per trial to access and integrate feedback (0.6-s vowel duration – 120-ms delay = 480 ms). This finding highlights the potential importance of

Figure 2. Average adaptive responses across phase sections per condition. Data points and lines connecting data points represent the means, and shaded regions represent 95% confidence intervals. The two long vowel conditions are in light blue and the two short vowel conditions are in purple. The long intertrial interval (ITI) duration conditions are represented with solid lines and the two short ITI duration conditions are represented with dashed lines. BL = baseline; AE = after-effect.



Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Figure 3. Boxplots of average adaptive responses by selected phase sections (i.e., late hold and early after-effect) per condition. Each box represents the first and third quartiles, and the central line in each box represents the median value. The two long vowel conditions are in light blue and the two short vowel conditions are in purple. The long intertrial interval (ITI) duration conditions are represented with filled boxes and the two short ITI duration conditions are represented with unfilled boxes. Mean values by participant are represented in dots and overlayed on boxplots per condition.



the amount of time during vocalization that a speaker has to access and integrate feedback into future motor commands.

ITI Duration

There was no significant difference in adaptive response magnitudes in the late hold phase section when comparing ITI duration. However, there was a significant difference in the early after-effect phase section with long ITI durations increasing adaptation. The significant finding in the early after-effect phase section suggests that researchers may also need to control for practice distribution when comparing adaptation results across differing studies. In the long ITI conditions, the time between trial onsets were longer (i.e., ~6 s) than the time between trial onsets for the short ITI duration conditions (i.e., ~3 s). This finding in the early after-effect phase section might signify that more rest time in between trials could influence retention, as rest times scale with ITI durations. In the conditions with long ITI durations, the rest time between trials was longer (i.e., the Long Vowel/Long ITI and Short Vowel/Long ITI conditions had mean added intertrial times of 4.8 s and 5.4 s, respectively), compared to the rest times between trials for the conditions with short ITI durations (i.e., the Short Vowel/Long ITI and Long Vowel/Short ITI conditions had mean added intertrial times of 1.85 and 2.45 s). This interpretation is supported by the nonspeech motor adaptation literature (Bock et al., 2005; Francis, 2005; Huang & Shadmehr, 2007; Kim et al., 2015) and general motor learning literature (Baddeley & Longman, 1978; Lee & Genovese, 1988; Maier et al., 2019; C. H. Shea et al., 2000), which suggests that distributed (rather than massed) practice increases motor learning and retention. On the other hand, the finding is in contrast with some clinical speech evidence supporting a benefit of massed practice on retention for children (Allen, 2013; Maas et al., 2019). However, one study on adults with apraxia of speech found a potential benefit to distributed practice for speech outcomes (Wambaugh et al., 2018), which aligns with the present study findings. Thus, the effect of practice distribution appears to differ between children and adult learners.

Prior adaptation work suggests that the amount of rest time between trials is specifically beneficial for retention rather than for the acquisition phase of learning during a practice session (Kim et al., 2015), which aligns with the results of this study. The distinction between acquisition and retention is an important one (e.g., Baddeley & Longman, 1978; Lee & Genovese, 1988; Salmoni et al., 1984; Schmidt & Bjork, 1992), especially in the context of long-term retention goals for clinical voice and speech therapy targets (Bislick et al., 2012). Other motor learning principles in the literature related to external feedback timing and frequency have shown a differential effect on acquisition and retention; that is, improved performance during practice does not always predict retention, and factors that enhance performance during practice are in some cases even *opposite* of those that enhance retention (Salmoni et al., 1984; Schmidt & Bjork, 1992; J. B. Shea & Morgan, 1979; see Maas et al., 2008, and Schmidt et al., 2018, for reviews). These findings motivate the differential study of vocal motor learning both during and after a practice session.

Clinical Implications

Adaptation paradigms have the potential to help us investigate practice parameters that influence implicit motor learning to guide clinical speech and voice treatment recommendations. Implicit vocal motor learning is a process that is central to voice therapy goals of shaping and changing individuals' vocal motor behaviors (Tellis, 2018). Current clinical approaches to therapy are often eclectic (Roy, 2012); clinical voice decisions are largely motivated by evidence for specific therapeutic techniques (Van Stan et al., 2015) rather than evidence for structuring practice sessions. To optimize voice therapy practice for those with voice disorders, there is a point at which overpracticing or vocalizing without adequate rest time could be detrimental and damaging to the vocal fold tissue (Ghasemzadeh et al.,

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

2024; Roy, 2012). Thus, it is important to avoid under- or overpracticing, and/or practicing suboptimally (Roy, 2012). In this study, we found that vocal motor adaptation was increased with longer vocalization time, and retention was increased with both longer vocalization time and longer rest times between trials. These findings suggest that longer vocalizations, more access to feedback, and thus, more time for speakers to integrate sensory feedback into vocal practice resulted in more implicit vocal motor learning. The findings of this study support therapy approaches that allow speakers to sufficiently practice speaking (McDonnell et al., 2018; R. R. Patel et al., 2011; Ramig et al., 1995, 2018; Sapir et al., 2011; Wenke et al., 2014) and integrating sensory feedback into their future vocal motor targets. Further, since the goal of voice therapy is to promote retention of learned vocal motor behaviors beyond the therapy session, it is crucial to consider incorporating long enough rest times between vocalizations for the speaker to consolidate sensory feedback from the prior vocalization.

Limitations and Future Directions

There are several limitations to consider in the interpretation of these study findings. We attempted to isolate the two parameters being compared (vowel duration and ITI duration). To accomplish this, additional parameters varied across the conditions (refer to Table 1), one of which is number of trials. The Short Vowel/Short ITI condition had double the number of trials compared to the three other conditions. We did not design this study to test the effects of number of trials on adaptation, given the a priori hypothesis that f_o adaptation plateaus rapidly, even with a small number of trials in the hold phase (Heller Murray & Stepp, 2020; Scheerer, Jacobson, & Jones, 2016). However, descriptively, this condition had the lowest mean adaptive response magnitudes in the hold (M = 24.7 cents) and after-effect phases (M = 3.1 cents) of the four conditions. More adaptation in the conditions with fewer trials is contrary to the prior evidence that suggests that more practice (and more repetitions) with a motor task leads to better outcomes (Kaipa & Peterson, 2016; Maas et al., 2019; Maier et al., 2019; Park & Shea, 2005; C. H. Shea & Kohl, 1991). Therefore, it is likely that, as hypothesized, vowel duration per trial and ITI duration are more important for vocal motor learning than the number of trial repetitions. Future work should test this theory further.

As compared to some prior studies of f_o adaptation paradigms (Behroozmand & Sangtian, 2018; R. Patel et al., 2011), the present study paradigm was relatively short, which potentially could have led to incomplete adaptation by the end of the hold phase. Based on the visualizations of adaptive responses in Figure 2 and Supplemental Material S2, adaptation plateaued in all four conditions, suggesting that adaptive responses from the late hold phase section may reflect the end point of learning, rather than the learning rate. Thus, we would not expect to observe significantly different results in the late hold phase section if the hold phases were extended. However, it is possible that we would have observed larger effect sizes for the differences in the early after-effect phase. Additionally, this study did not investigate whether participants differentially noticed the f_o perturbation across the conditions. It is unclear whether the manipulated parameters in the present study also influence participants' awareness of the pitch shift during vocalization in the hold phase and whether that differing awareness would have influenced the integration of sensory feedback. The influence of participant awareness of varying perturbation durations on adaptation should be considered in future studies.

A limitation in applying the results of this study to clinical contexts is that adaptation paradigms primarily measure implicit vocal motor learning rather than explicit learning. Despite the role of both implicit and explicit learning in voice therapy (Tellis, 2018), most therapy programs rely heavily on explicit clinician-directed feedback or biofeedback (Eastwood et al., 2024; Madill et al., 2020; Van Stan et al., 2017, 2019; Yiu et al., 2005) to initiate implicit learning. Research suggests that low frequency of externally provided feedback, as compared to high frequency, increases motor learning (Salmoni et al., 1984; Steinhauer & Grayhack, 2000; Winstein & Schmidt, 1990; Wulf & Schmidt, 1989). These findings might reflect the potential for a speaker to become too reliant on externally provided feedback, hindering the speaker's ability to learn from internal sensory feedback integration (the "guidance hypothesis"; Salmoni et al., 1984). Further, explicit learning is often modulated by the learner's motivation to learn, their understanding and awareness of a task, and their stimulability (Maas et al., 2008). Thus, this study did not replicate a clinical setting; therefore, it can only help us to infer how parameters might influence implicit vocal motor learning from self-produced sensory feedback. Future studies should investigate these parameters within a clinical context.

Conclusions

This study investigated the influence of two practice parameters (vowel duration and ITI duration) on implicit vocal motor learning, as measured by adaptive responses in altered auditory feedback paradigms. In the late hold phase section (during learning), adaptive responses were greater with longer vowel duration, but there was no effect of ITI duration. In the early after-effect phase section (retention), adaptive responses were greater with longer vowel duration and longer ITI duration (i.e., distributed practice). These

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

findings suggest that adaptation is increased when speakers have more exposure to errors detected by altered auditory feedback and more time within a vocalization to access and integrate that feedback into subsequent vocalizations. Further, distributed practice (i.e., longer rest times between trials) may not benefit vocal motor learning during practice, but it may positively influence learning for retention.

Data Availability Statement

The data sets generated during and/or analyzed during the current study are not publicly available due to commitments to protect participant confidentiality but are available from the senior author, Cara E. Stepp (cstepp@bu.edu), through formal mechanisms on reasonable request.

Acknowledgments

This work was supported by the National Institute on Deafness and Other Communication Disorders Grants DC015446 (awarded to Robert Hillman) and F31 DC021863 (awarded to Allison S. Aaron). The authors thank Keerthi Maddukuri, Jodi Yu, and Miracle Cariaga for their assistance in 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., 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 13123. https://doi.org/10.1038/s41598-021-92250-8

Acosta, A., Kearney, E., & Guenther, F. H. (2023, June). *Pitch and formant targets vary across days as seen in reflexive responses to auditory perturbations* [Poster presentation]. Boston Speech Motor Control Symposium, Boston, MA, United States.

Allen, M. M. (2013). Intervention efficacy and intensity for children with speech sound disorder. *Journal of Speech, Language, and Hearing Research, 56*(3), 865–877. https://doi.org/10.1044/1092-4388(2012/11-0076)

American Speech-Language-Hearing Association. (2005). *Guidelines for manual pure-tone threshold audiometry* [Guidelines]. https://www.asha.org/policy/gl2005-00014/

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

Baddeley, A. D., & Longman, D. J. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21*(8), 627–635. https://doi.org/10.1080/00140137808931764

Barkley, V., Salomonczyk, D., Cressman, E. K., & Henriques, D. Y. P. (2014). Reach adaptation and proprioceptive recalibration following terminal visual feedback of the hand. *Frontiers in Human Neuroscience, 8,* Article 705. https://doi.org/10.3389/fnhum.2014.00705

Behroozmand, R., & Sangtian, S. (2018). Neural bases of sensorimotor adaptation in the vocal motor system. *Experimental Brain Research, 236*(7), 1881–1895. https://doi.org/10.1007/s00221-018-5272-9

Bislick, L. P., Weir, P. C., Spencer, K., Kendall, D., & Yorkston, K. M. (2012). Do principles of motor learning enhance retention and transfer of speech skills? A systematic review. *Aphasiology, 26*(5), 709–728. https://doi.org/10.1080/02687038.2012.676888

Bock, O., Thomas, M., & Grigorova, V. (2005). The effect of rest breaks on human sensorimotor adaptation. *Experimental Brain Research, 163*(2), 258–260. https://doi.org/10.1007/s00221-005-2231-z

Boersma, P., & Weenink, D. (2008). *Praat: Doing phonetics by computer* [Computer software]. http://www.praat.org/

Boominathan, P., Anitha, R., Shenbagavalli, M., & Dinesh, G. (2010). Voice characteristics and recovery patterns in Indian adult males after vocal loading. *Journal of the All India Institute of Speech & Hearing, 29*(2).

Bradley, J. V. (1958). Complete counterbalancing of immediate sequential effects in a Latin square design. *Journal of the American Statistical Association, 53*(282), 525–528. https://doi.org/10.1080/01621459.1958.10501456

Brashers-Krug, T., Shadmehr, R., & Bizzi, E. (1996). Consolidation in human motor memory. *Nature, 382*(6588), 252–255. https://doi.org/10.1038/382252a0

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice f_o 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., & Larson, C. R. (2002). Early pitch-shift response is active in both steady and dynamic voice pitch control. *The Journal of the Acoustical Society of America, 112*(3), 1058–1063. https://doi.org/10.1121/1.1487844

Burnett, T. A., Senner, J. E., & Larson, C. R. (1997). Voice f_o 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

Chao, S.-C., & Daliri, A. (2023). Effects of gradual and sudden introduction of perturbations on adaptive responses to formant-shift and formant-clamp perturbations. *Journal of Speech, Language, and Hearing Research, 66*(5), 1588–1599. https://doi.org/10.1044/2023_JSLHR-21-00435

Dahl, K. L., Cádiz, M. D., Zuk, J., Guenther, F. H., & Stepp, C. E. (2024). Controlling pitch for prosody: Sensorimotor adaptation in linguistically meaningful contexts. *Journal of Speech, Language, and Hearing Research, 67*(2), 440–454. https://doi.org/10.1044/2023_JSLHR-23-00460

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

Eastwood, C., McCabe, P., & Heard, R. (2024). Unpacking the black box of voice therapy: A clinical application and revision of the Motor Learning Classification Framework (MLCF). *International Journal of Speech-Language Pathology, 26*(1), 68–82. https://doi.org/10.1080/17549507.2022.2079723

Edeal, D. M., & Gildersleeve-Neumann, C. E. (2011). The importance of production frequency in therapy for childhood apraxia of speech. *American Journal of Speech-Language Pathology, 20*(2), 95–110. https://doi.org/10.1044/1058-0360(2011/09-0005)

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Francis, J. T. (2005). Influence of the inter-reach-interval on motor learning. *Experimental Brain Research, 167*(1), 128–131. https://doi.org/10.1007/s00221-005-0062-6

Ghasemzadeh, H., Hillman, R. E., & Mehta, D. D. (2024). Consistency of the signature of phonotraumatic vocal hyperfunction across different ambulatory voice measures. *Journal of Speech, Language, and Hearing Research, 67*(7), 1997–2020. https://doi.org/10.1044/2024_JSLHR-23-00515

Hardwick, R. M., Rajan, V. A., Bastian, A. J., Krakauer, J. W., & Celnik, P. A. (2017). Motor learning in stroke: Trained patients are not equal to untrained patients with less impairment. *Neurorehabilitation and Neural Repair, 31*(2), 178–189. https://doi.org/10.1177/1545968316675432

Heller Murray, E. S., Lupiani, A. A., Kolin, K. R., Segina, R. K., & Stepp, C. E. (2019). Pitch shifting with the commercially available eventide eclipse: Intended and unintended changes to the speech signal. *Journal of Speech, Language, and Hearing Research, 62*(7), 2270–2279. https://doi.org/10.1044/2019_JSLHR-S-18-0408

Heller Murray, E. S., & Stepp, C. E. (2020). Relationships between vocal pitch perception and production: A developmental perspective. *Scientific Reports, 10*(1), Article 3912. https://doi.org/10.1038/s41598-020-60756-2

Hinder, M. R., Tresilian, J. R., Riek, S., & Carson, R. G. (2008). The contribution of visual feedback to visuomotor adaptation: How much and when? *Brain Research, 1197*, 123–134. https://doi.org/10.1016/j.brainres.2007.12.067

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. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience, 5*, Article 82. https://doi.org/10.3389/fnhum.2011.00082

Huang, V. S., & Shadmehr, R. (2007). Evolution of motor memory during the seconds after observation of motor error. *Journal of Neurophysiology, 97*(6), 3976–3985. https://doi.org/10.1152/jn.01281.2006

Institute of Electrical and Electronics Engineers. (1969). *IEEE recommended practice for speech quality measurements (Std 297-1969)*. https://doi.org/10.1109/IEEESTD.1969.7405210

Jones, J. A., & Keough, D. (2008). Auditory-motor mapping for pitch control in singers and nonsingers. *Experimental Brain Research, 190*(3), 279–287. https://doi.org/10.1007/s00221-008-1473-y

Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of f_{o} production: Evidence from feedback perturbation. *The Journal of the Acoustical Society of America, 108*(3), 1246–1251. https://doi.org/10.1121/1.1288414

Kaipa, R., & Peterson, A. M. (2016). A systematic review of treatment intensity in speech disorders. *International Journal of Speech-Language Pathology, 18*(6), 507–520. https://doi.org/10.3109/17549507.2015.1126640

Kapsner-Smith, M. R., Abur, D., Eadie, T. L., & Stepp, C. E. (2024). Test–retest reliability of behavioral assays of feedforward and feedback auditory–motor control of voice and articulation. *Journal of Speech, Language, and Hearing Research, 67*(1), 34–48. https://doi.org/10.1044/2023_JSLHR-23-00038

Keough, D., Hawco, C., & Jones, J. A. (2013). Auditory-motor adaptation to frequency-altered auditory feedback occurs when participants ignore feedback. *BMC Neuroscience, 14*(1), Article 25. https://doi.org/10.1186/1471-2202-14-25

Kim, S., Oh, Y., & Schweighofer, N. (2015). Between-trial forgetting due to interference and time in motor adaptation. *PLOS ONE, 10*(11), Article e0142963. https://doi.org/10.1371/journal.pone.0142963

Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., & Haith, A. M. (2019). Motor learning. *Comprehensive Physiology, 9*(2), 613–663. https://doi.org/10.1002/cphy.c170043

Lametti, D. R., Quek, M. Y. M., 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

Larson, C. R. (1998). Cross-modality influences in speech motor control: The use of pitch shifting for the study of f_{o} control. *Journal of Communication Disorders, 31*(6), 489–503. https://doi.org/10.1016/S0021-9924(98)00021-5

Larson, C. R., Burnett, T. A., Bauer, J. J., Kiran, S., & Hain, T. C. (2001). Comparison of voice f_{o} responses to pitch-shift onset and offset conditions. *The Journal of the Acoustical Society of America, 110*(6), 2845–2848. https://doi.org/10.1121/1.1417527

Lee, T. D., & Genovese, E. D. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for Exercise and Sport, 59*(4), 277–287. https://doi.org/10.1080/02701367.1988.10609373

Maas, E., Gildersleeve-Neumann, C., Jakielski, K., Kovacs, N., Stoeckel, R., Vradelis, H., & Welsh, M. (2019). Bang for your buck: A single-case experimental design study of practice amount and distribution in treatment for childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 62*(9), 3160–3182. https://doi.org/10.1044/2019_JSLHR-S-18-0212

Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology, 17*(3), 277–298. https://doi.org/10.1044/1058-0360(2008/025)

Madill, C., McIlwaine, A., Russell, R., Hodges, N. J., & McCabe, P. (2020). Classifying and identifying motor learning behaviors in voice-therapy clinician–client interactions: A proposed motor learning classification framework. *Journal of Voice, 34*(5), 806.e19–806.e31. https://doi.org/10.1016/j.jvoice.2019.03.014

Maier, M., Ballester, B. R., & Verschure, P. F. (2019). Principles of neurorehabilitation after stroke based on motor learning and brain plasticity mechanisms. *Frontiers in Systems Neuroscience, 13*, Article 74. https://doi.org/10.3389/fnsys.2019.00074

McDonnell, M. N., Rischbieth, B., Schammer, T. T., Seaforth, C., Shaw, A. J., & Phillips, A. C. (2018). Lee Silverman Voice Treatment (LSVT)-BIG to improve motor function in people with Parkinson's disease: A systematic review and meta-analysis. *Clinical Rehabilitation, 32*(5), 607–618. https://doi.org/10.1177/0269215517734385

Miller, H. E., Kearney, E., Nieto-Castañón, A., Falsini, R., Abur, D., Acosta, A., Chao, S.-C., Dahl, K. L., Franken, M., Heller Murray, E. S., Mollaei, F., Niziolek, C. A., Parrell, B., Perrachione, T., Smith, D. J., Stepp, C. E., Tomassi, N., & Guenther, F. H. (2023). Do not cut off your tail: A mega-analysis of responses to auditory perturbation experiments. *Journal of Speech, Language, and Hearing Research, 66*(11), 4315–4331. https://doi.org/10.1044/2023_JSLHR-23-00315

Namasivayam, A. K., Pukonen, M., Goshulak, D., Hard, J., Rudzicz, F., Rietveld, T., Maassen, B., Kroll, R., & van Lieshout, P. (2015). Treatment intensity and childhood apraxia of speech. *International Journal of Language & Communication Disorders, 50*(4), 529–546. https://doi.org/10.1111/1460-6984.12154

Park, J.-H., & Shea, C. H. (2005). Sequence learning: Response structure and effector transfer. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 58A*(3), 387–419. https://doi.org/10.1080/02724980343000918

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

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Patel, R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Svec, J. G., & Hillman, R. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology, 27*(3), 887–905. https://doi.org/10.1044/2018_AJSLP-17-0009

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), 1051–1059. https://doi.org/10.1044/1092-4388(2010/10-0162)

Patel, R. R., Bless, D. M., & Thibeault, S. L. (2011). Boot camp: A novel intensive approach to voice therapy. *Journal of Voice, 25*(5), 562–569. https://doi.org/10.1016/j.jvoice.2010.01.010

Ramig, L. O., Countryman, S., Thompson, L. L., & Horii, Y. (1995). Comparison of two forms of intensive speech treatment for Parkinson disease. *Journal of Speech and Hearing Research, 38*(6), 1232–1251. https://doi.org/10.1044/jshr.3806.1232

Ramig, L. O., Halpern, A., Spielman, J., Fox, C., & Freeman, K. (2018). Speech treatment in Parkinson's disease: Randomized controlled trial (RCT). *Movement Disorders, 33*(11), 1777–1791. https://doi.org/10.1002/mds.27460

Remacle, A., Finck, C., Roche, A., & Morsomme, D. (2012). Vocal impact of a prolonged reading task at two intensity levels: Objective measurements and subjective self-ratings. *Journal of Voice, 26*(4), e177–e186. https://doi.org/10.1016/j.jvoice.2011.07.016

Roy, N. (2012). Optimal dose–response relationships in voice therapy. *International Journal of Speech-Language Pathology, 14*(5), 419–423. https://doi.org/10.3109/17549507.2012.686119

RStudio Team. (2020). *RStudio: Integrated Development for R* [Computer software]. RStudio, PBC.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin, 95*(3), 355–386. https://doi.org/10.1037/0033-2909.95.3.355

Sapir, S., Ramig, L. O., & Fox, C. M. (2011). Intensive voice treatment in Parkinson's disease: Lee Silverman Voice Treatment. *Expert Review of Neurotherapeutics, 11*(6), 815–830. https://doi.org/10.1586/ern.11.43

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., Tumber, A. K., & Jones, J. A. (2016). Attentional demands modulate sensorimotor learning induced by persistent exposure to changes in auditory feedback. *Journal of Neurophysiology, 115*(2), 826–832. https://doi.org/10.1152/jn.00799.2015

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–217. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x

Schmidt, R. A., Lee, T. D., Winstein, C., Wulf, G., & Zelaznik, H. N. (2018). *Motor control and learning: A behavioral emphasis*. Human Kinetics.

Schmidt, R. A., & Wulf, G. (2002). Continuous concurrent feedback degrades skill learning: Implications for training and simulation. In F. Jentsch, & M. Curtis (Eds.), *Simulation in aviation training* (pp. 375–392). Routledge.

Schween, R., Taube, W., Gollhofer, A., & Leukel, C. (2014). Online and post-trial feedback differentially affect implicit adaptation to a visuomotor rotation. *Experimental Brain Research, 232*(9), 3007–3013. https://doi.org/10.1007/s00221-014-3992-z

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

Shea, C. H., Kohl, R., & Indermill, C. (1990). Contextual interference: Contributions of practice. *Acta Psychologica, 73*(2), 145–157. https://doi.org/10.1016/0001-6918(90)90076-R

Shea, C. H., & Kohl, R. M. (1991). Composition of practice: Influence on the retention of motor skills. *Research Quarterly for Exercise and Sport, 62*(2), 187–195. https://doi.org/10.1080/02701367.1991.10608709

Shea, C. H., Lai, Q., Black, C., & Park, J.-H. (2000). Spacing practice sessions across days benefits the learning of motor skills. *Human Movement Science, 19*(5), 737–760. https://doi.org/10.1016/S0167-9457(00)00021-X

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*(2), 179–187. https://doi.org/10.1037/0278-7393.5.2.179

Steinhauer, K., & Grayhack, J. P. (2000). The role of knowledge of results in performance and learning of a voice motor task. *Journal of Voice, 14*(2), 137–145. https://doi.org/10.1016/S0892-1997(00)80020-X

Stemple, J. C., Stanley, J., & Lee, L. (1995). Objective measures of voice production in normal subjects following prolonged voice use. *Journal of Voice, 9*(2), 127–133. https://doi.org/10.1016/S0892-1997(05)80245-0

Tellis, C. (2018). New concepts in motor learning and training related to voice rehabilitation. *Perspectives of the ASHA Special Interest Groups, 3*(3), 56–67. https://doi.org/10.1044/persp3.SIG3.56

Tomassi, N. E., Turashvili, D. M., Williams, A., Walsh, B., Stephen, E. P., & Stepp, C. E. (2025). Investigating cognitive load and autonomic arousal during voice production and vocal auditory-motor adaptation. *Journal of Speech, Language, and Hearing Research, 68*(4), 1634–1653. https://doi.org/10.1044/2024_JSLHR-24-00601

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

Van Stan, J. H., Dijkers, M. P., Whyte, J., Hart, T., Turkstra, L. S., Zanca, J. M., & Chen, C. (2019). The rehabilitation treatment specification system: Implications for improvements in research design, reporting, replication, and synthesis. *Archives of Physical Medicine and Rehabilitation, 100*(1), 146–155. https://doi.org/10.1016/j.apmr.2018.09.112

Van Stan, J. H., Mehta, D. D., Petit, R. J., Sternad, D., Muise, J., Burns, J. A., & Hillman, R. E. (2017). Integration of motor learning principles into real-time ambulatory voice biofeedback and example implementation via a clinical case study with vocal fold nodules. *American Journal of Speech-Language Pathology, 26*(1), 1–10. https://doi.org/10.1044/2016_AJSLP-15-0187

Van Stan, J. H., Roy, N., Awan, S., Stemple, J., & Hillman, R. E. (2015). A taxonomy of voice therapy. *American Journal of Speech-Language Pathology, 24*(2), 101–125. https://doi.org/10.1044/2015_AJSLP-14-0030

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

Wambaugh, J. L., Wright, S., Boss, E., Mauszycki, S. C., DeLong, C., Hula, W., & Doyle, P. J. (2018). Effects of treatment intensity

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

on outcomes in acquired apraxia of speech. *American Journal of Speech-Language Pathology, 27*(1S), 306–322. https://doi.org/10.1044/2017_AJSLP-16-0188

Wang, T., Avraham, G., Tsay, J. S., Thummala, T., & Ivry, R. B. (2024). Advanced feedback enhances sensorimotor adaptation. *Current Biology, 34*(5), 1076–1085.e5. https://doi.org/10.1016/j.cub.2024.01.073

Wang, T., Li, J., & Ivry, R. B. (2024). *Attention defines the context for implicit sensorimotor adaptation*. bioRxiv. https://doi.org/10.1101/2024.09.03.611108

Wang, W., Wei, L., Chen, N., Jones, J. A., Gong, G., & Liu, H. (2019). Decreased gray-matter volume in insular cortex as a correlate of singers' enhanced sensorimotor control of vocal production. *Frontiers in Neuroscience, 13,* Article 815. https://doi.org/10.3389/fnins.2019.00815

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., 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), Article e1010159. https://doi.org/10.1371/journal.pcbi.1010159

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

Wenke, R. J., Stabler, P., Walton, C., Coman, L., Lawrie, M., O'Neill, J., Theodoros, D., & Cardell, E. (2014). Is more intensive better? Client and service provider outcomes for intensive versus standard therapy schedules for functional voice disorders. *Journal of Voice, 28*(5), 652.e31–652.e43. https://doi.org/10.1016/j.jvoice.2014.02.005

Winstein, C. J., & Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(4), 677–691. https://doi.org/10.1037/0278-7393.16.4.677

Witte, R. S., & Witte, J. S. (2017). *Statistics*. Wiley.

Wulf, G., & Schmidt, R. A. (1989). The learning of generalized motor programs: Reducing the relative frequency of knowledge of results enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(4), 748–757. https://doi.org/10.1037/0278-7393.15.4.748

Yamada, C., Itaguchi, Y., & Fukuzawa, K. (2019). Effects of the amount of practice and time interval between practice sessions on the retention of internal models. *PLOS ONE, 14*(4), Article e0215331. https://doi.org/10.1371/journal.pone.0215331

Yiu, E. M.-L., Verdolini, K., & Chow, L. P. Y. (2005). Electromyographic study of motor learning for a voice production task. *Journal of Speech, Language, and Hearing Research, 48*(6), 1254–1268. https://doi.org/10.1044/1092-4388(2005/087a)

Downloaded from: https://pubs.asha.org Boston University on 02/12/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions