JSLHR

Research Article

Auditory-Motor Perturbations of Voice Fundamental Frequency: Feedback Delay and Amplification

Hasini R. Weerathunge,[a] Defne Abur,[b] Nicole M. Enos,[a,c] Katherine M. Brown,[b] and Cara E. Stepp[a,b,d]

Purpose: Gradual and sudden perturbations of vocal fundamental frequency (f_o), also known as adaptive and reflexive f_o perturbations, are techniques to study the influence of auditory feedback on voice f_o control mechanisms. Previous vocal f_o perturbations have incorporated varied setup-specific feedback delays and amplifications. Here, we investigated the effects of feedback delays (10–100 ms) and amplifications on both adaptive and reflexive f_o perturbation paradigms, encapsulating the variability in equipment-specific delays (3–45 ms) and amplifications utilized in previous experiments.

Method: Responses to adaptive and reflexive f_o perturbations were recorded in 24 typical speakers for four delay conditions (10, 40, 70, and 100 ms) or three amplification conditions (–10, +5, and +10 dB relative to microphone) in a counterbalanced order. Repeated-measures analyses of variance were carried out on the magnitude of f_o responses to determine the effect of feedback condition.

Results: There was a statistically significant effect of the level of auditory feedback amplification on the response magnitude during adaptive f_o perturbations, driven by the difference between +10- and –10-dB amplification conditions (hold phase difference: M = 38.3 cents, SD = 51.2 cents; after-effect phase: M = 66.1 cents, SD = 84.6 cents). No other statistically significant effects of condition were found for either paradigm.

Conclusions: Experimental equipment delays below 100 ms in behavioral paradigms do not affect the results of f_o perturbation paradigms. As there is no statistically significant difference between the response magnitudes elicited by +5- and +10-dB auditory amplification conditions, this study is a confirmation that an auditory feedback amplification of +5 dB relative to microphone is sufficient to elicit robust compensatory responses for f_o perturbation paradigms.

Auditory feedback is an important modality that enables fine-tuning of vocalizations during ongoing vocal productions (Lane & Tranel, 1971; Yates, 1963). Its necessity for fluent speech is substantiated by cases of complete auditory deprivation (Ertmer et al., 2003; Périer et al., 1984). Furthermore, adults with acquired deafness experience deterioration in vocal control over time, which leads to monotonic speech and dysprosody (Kirchner & Suzuki, 1968; Penn, 1955). Altered auditory feedback experiments are a powerful technique to study the auditory-motor control mechanisms of voice fundamental frequency (f_o), the acoustic correlate of pitch (Hixon et al., 2018). These experiments systematically manipulate the frequency spectrum of the voice, such that f_o is shifted upward or downward in frequency. Speakers' responses to the experiments can be heterogeneous. Some speakers compensate for the perceived f_o changes by producing f_o changes opposing the direction of the shift to reduce the difference between the perceived and the intended vocal f_o production. Other speakers respond by following the direction of the shift. There are also speakers who do not change their f_o in response to perturbations. Most prior studies have reported a majority of compensatory responses (e.g., Burnett et al., 1998, 1997; Elman, 1981; Jones & Munhall, 2005; Larson et al., 2000; H. Liu & Larson, 2007), although a few studies have noted a more uniform distribution across response categories (Arbeiter et al., 2018; Petermann et al., 2016; Ziethe et al., 2019).

[a]Department of Biomedical Engineering, Boston University, MA
[b]Department of Speech, Language, and Hearing Sciences, Boston University, MA
[c]Department of Computer Engineering, Boston University, MA
[d]Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, MA

Correspondence to Hasini R. Weerathunge: hasiniw@bu.edu

Editor-in-Chief: Bharath Chandrasekaran
Editor: Jack J. Jiang

Received December 19, 2019
Revision received March 30, 2020
Accepted June 10, 2020
https://doi.org/10.1044/2020_JSLHR-19-00407

Disclosure: The authors have declared that no competing interests existed at the time of publication.

Shifting voice f_o has been used to elucidate the closed- and open-loop control properties of vocal motor control. The closed-loop mechanism (*feedback control*) is assessed when the perceived f_o is altered abruptly, often referred to as pitch shifts or the pitch reflexive paradigm (e.g., Elman, 1981; Bauer & Larson, 2003; Burnett et al., 1997; S. H. Chen et al., 2007; Hain et al., 2000; Larson et al., 2000). Long latencies required to process auditory feedback (Burnett et al., 1998; Hain et al., 2001; Larson et al., 2000) suggest that feedback control is probably of limited utility for rapid regulation of speech. The open-loop control mechanism (*feedforward control*) consists of a predetermined set of motor commands for an intended vocal production and is thus utilized for rapid regulation of vocal f_o. Adaptive perturbation paradigms gradually manipulate auditory feedback in a sustained and predictive manner to study the influence of auditory feedback over time and its contribution to updates to the feedforward control system (e.g., Jones & Keough, 2008; Jones & Munhall, 2000, 2002, 2005; Kawahara, 1993; Keough et al., 2013; Keough & Jones, 2009; MacDonald & Munhall, 2012).

Studies incorporating adaptive or reflexive altered auditory feedback paradigms of vocal f_o have been replicated and refined, utilizing a variety of experimental procedures and equipment for shifting f_o. The generic setup of an altered f_o perturbation experiment includes microphones to acquire speech signals from participants and headphones to deliver the altered auditory feedback to the participants in near real time. The experimental apparatus also incorporates various mechanisms to amplify the altered auditory feedback to prevent the participants from hearing their own unaltered speech. Across prior work, there is considerable variation in the experimental setup employed, mainly in terms of (a) auditory feedback delay and (b) auditory feedback amplification.

Effects of Auditory Feedback Delay

Hardware and software processing algorithms introduce apparatus-specific intrinsic delays to the acoustic signal due to the computational complexities of the acoustic perturbations and data transfer via digitization. This causes unintended delays in the auditory feedback relative to the vocalization of the speaker. In prior research, the precise auditory feedback delay of equipment has been seldom measured and reported. In the cases in which the auditory feedback delay was reported, the delays varied from 3 to 45 ms. Commercial systems for f_o shifting produce delays up to 50 ms (Heller Murray et al., 2019). Table A1 in the Appendix summarizes apparatus-specific auditory feedback delays reported across prior research. To compare participants' responses to applied auditory feedback perturbations across studies, it should be assumed that potential differences in feedback delays have no effect on the responses. However, there are suggestions from prior work that challenge this assumption.

A single study has examined the effect of delayed auditory feedback on participants' responses to f_o perturbation. Hain et al. (2001) reported that incorporating a feedback delay in the range of 50–500 ms to an f_o perturbation reflexive paradigm during sustained vowels did not lead to statistically significant variations on the f_o response magnitudes across each delay condition. It is thus far unknown whether delays smaller than 50 ms (in the range of realistic hardware or software delays) impact reflexive f_o response magnitudes.

There are no reported studies to date exploring the effects of delayed auditory feedback on f_o adaptation. However, two studies have observed effects of delayed auditory feedback on vowel formant adaptive responses. Although segmental features such as formants are expected to be controlled on a faster time scale compared to the prosodic features such as voice f_o (Guenther, 2016), these studies may provide some insight into the effects of delayed auditory feedback on the voice f_o control system. Max and Maffett (2015) gradually shifted the first and second vowel formants of the auditory feedback of participants' over 120 productions of the word "head" and observed an absence of compensatory responses when the auditory feedback was delayed by 100, 250, or 500 ms. Mitsuya et al. (2017) shifted only the first vowel formant while delaying auditory feedback systematically over the range of 0–100 ms and found a small, yet statistically significant, compensatory response even in the 100-ms delay condition, in contrast to the prior study. Mitsuya et al. also found that the amount of formant compensation had a negative linear relationship with the amount of auditory delay.

Effects of Auditory Feedback Amplification

Amplifying the perturbed auditory feedback signal may limit the impact of air- and bone-conducted feedback. The majority of the auditory perturbation studies have amplified the headphone signals by a specific gain relative to the produced speech signal in dB (Behroozmand et al., 2012; Larson et al., 2000; Max et al., 2003; Stepp et al., 2017) or maintained the headphone signals at a specific amplified sound pressure level (dB SPL; Bauer & Larson, 2003; Elman, 1981; Hain et al., 2001; Hawco & Jones, 2010; MacDonald & Munhall, 2012). Table A2 in the Appendix summarizes prior literature related to f_o perturbation paradigms, listing (a) the amplification applied between microphone and headphones, (b) masking methods utilized, and (c) sound pressure level monitoring mechanisms, if reported.

In order to compare f_o perturbation responses across different studies, it is crucial to identify whether the degree of amplification applied to the auditory feedback signals in the study apparatus has an effect on the participants' responses to all types of f_o perturbation paradigms. A single study previously examined the effects of different auditory feedback amplification conditions on reflexive f_o perturbations (Burnett et al., 1998). In this study, an f_o reflexive paradigm was carried out, utilizing three different listening conditions in terms of participants' processed voice signals' sound pressure level (65, 75, and 85 dB SPL). Variations in listening conditions in terms of sound pressure level, either alone or combined with pink masking noise (50, 60, or 70 dB SPL), had no effect on group means of response magnitude. No studies to date have examined the effect of auditory feedback amplification on f_o adaptation paradigm responses.

Current Investigation

Thus far, the effects of auditory feedback delay and amplification in f_o perturbation paradigms have not been studied comprehensively by looking at both reflexive and adaptive paradigms in the same speakers. Understanding how methodological variations in feedback delay and amplification affect f_o perturbations for both types of paradigms would have significant implications for the interpretation of results from prior work and could be utilized to identify the necessity of maintaining comparable delays and amplification in feedback for varied experimental setups.

The current study had two primary objectives. The first objective was to comprehensively investigate how delayed auditory feedback in the range of 0–100 ms affects responses for both reflexive and adaptive f_o perturbations of the auditory feedback. Experiment 1 of the current study contained reflexive and adaptive f_o perturbations for auditory feedback delay conditions of 10, 40, 70, and 100 ms. For the f_o reflexive experiment, we hypothesized that the results would closely follow the results of Hain et al. (2001), in which there were minimal effects on response magnitudes with delays in the range of 50–500 ms. With respect to f_o adaptive responses, based on the formant adaptation study by Max and Maffett (2015), we hypothesized that f_o response magnitudes would be reduced when the auditory feedback delay was increased.

The second objective of the current study was to investigate the degree to which the level of auditory feedback amplification affects responses to reflexive and adaptive f_o paradigms. Thus, Experiment 2 of the current study utilized a series of amplification levels based on the prior literature (i.e., −10-, +5-, and +10-dB amplification relative to microphone signal). Burnett et al. (1998) found no statistically significant difference between reflexive f_o response magnitudes at different auditory feedback amplifications. Thus, we hypothesized that variations in amplification level would have no effect on reflexive f_o response magnitudes. The effects of feedback amplification level have not been examined previously. Based on models of speech motor control (Guenther, 2016), changes to the feedback subsystem are integrated into the feedforward subsystem over subsequent productions. As there were no significant effects of amplification on reflexive f_o paradigm responses found in prior research, we hypothesized that the level of amplification would not show a significant effect on the adaptive f_o responses.

Method

Participants

Twenty-four young cisgender adults aged 18–26 years (12 women: M = 22 years, SD = 2 years; 12 men: M = 21 years, SD = 2 years) participated in the study. Twelve participants completed Experiment 1 (six women, six men), and the remaining 12 completed Experiment 2 (six women, six men). The participants were nonsmokers and native speakers of American English, with no history of speech, language, hearing, or neurological disorders. All participants had normal voices as determined by overall severity of dysphonia scores that were less than 34.7, as per ratings by a voice-specializing speech-language pathologist using the Consensus Auditory-Perceptual Evaluation of Voice (Kempster et al., 2009). The participants had no professional training as singers or musicians. All participants passed a hearing screening at 25 dB HL (125, 250, 500, 1000, 2000, 4000, and 8000 Hz).[1] The hearing screening was administered utilizing a Radioear 3045 DD45 circumaural audiometric headset and a Grason-Stadler GSI 18 screening audiometer. Participants provided written consent in compliance with the Boston University Institutional Review Board.

Instrumentation and Procedure

All experimental protocols were carried out in a sound-attenuated booth utilizing custom MATLAB scripts (Mathworks, 2018 Version 9.4.0.813654 [R2018a]). The participants sat in front of a computer screen that provided visual cues to alert participants to vocalize a sustained vowel /a/ for each trial. Participant voices were recorded by an omnidirectional condenser earset microphone (SHURE MX153 based on American Speech-Language-Hearing Association recommendations; R. R. Patel et al., 2018). The microphone signal was amplified via an RME Quadmic II microphone preamplifier and digitized via either a MOTU Ultralite-mk3 Hybrid or an RME Fireface UCX sound card with 32-bit resolution, both with sampling rates of 44100 Hz. The microphone signal was sent through an Eclipse V4 Harmonizer (Eclipse), which performed a full-spectrum frequency shift without formant correction. The altered speech signal was amplified via a Behringer Xenyx Q802USB earphone amplifier and presented back to the participants over Etymotic ER-2 insert earphones (Etymotic Research, Inc.) to provide near real-time auditory feedback (intrinsic hardware delay at approximately 12 ms). The resulting feedback signal was also digitized and was saved for offline analysis. The ER-2 insert earphones attenuate air-conducted sound by approximately 40 dB (Z-weighting) with proper insertion depth (Dean & Martin, 2000).

Experiment 1 (f_o Perturbations With Delayed Feedback)

For Experiment 1, 12 participants completed reflexive and adaptive f_o perturbation paradigms under four delay

[1]Pulsed pure tones (Burk & Wiley, 2004) presented for 1–2 s at 25 dB (modified version based on American National Standards Institute, 1996, 2004b).

conditions: 10, 40, 70, and 100 ms (see Figure 1 for order of condition presentation). Each trial was 3 s in duration, and the intertrial intervals were jittered from 1 to 3 s to maintain the participants' attention throughout the experiment. For each trial, participants vocalized a sustained vowel /ɑ/ at a comfortable speaking voice when a visual cue of "aaa" was displayed on the computer screen. The participants were instructed to keep the vocalization steady and consistent throughout the experiment. Before the experimental paradigm, a practice session with nine unperturbed trials was conducted to familiarize the participants to the task. In all experimental blocks in Experiment 1, the feedback amplification was maintained at +5 dB relative to the microphone intensity. The auditory feedback was delayed according to the delay condition applied for the experimental block.

f_o Reflex

The reflexive paradigm contained abrupt perturbations of f_o under the four feedback delay conditions, with the order counterbalanced across participants. Each condition consisted of a single "shift up" block of 60 trials. Within each block of vocalizations, 16 experimental trials were mixed pseudorandomly with 44 control trials, in which no f_o shift was presented. As the pseudorandomly ordered control trials wash out any adaptation effects to the reflex paradigm, no other control conditions were presented. The voicing period of 100 ms prior to f_o perturbation onset in the shifted trials was considered to be the baseline phase and utilized to normalize each of the shift up reflexive responses. In the experimental trials, an f_o shift stimuli of +100 cents[2] was presented with a random perturbation onset time in the range of 500–1,000 ms after vocalization onset. The f_o shift was maintained throughout the rest of the trial duration. The harmonizer also imposed a delay of either 10, 40, 70, or 100 ms (depending on condition) in the auditory feedback signal before vocal onset, which was maintained throughout the trial duration. The presentation of the f_o shift generated an abrupt shift in the auditory feedback f_o from the participant's true f_o. The participants did not receive any information about the differences in the four feedback conditions but were pre-instructed that there could be sudden f_o variations in their voice feedback during the experiments. Due to this specific instruction, the f_o reflex experiments were conducted after the f_o adaptation experiments.

f_o Adaptation

The adaptation paradigm contained gradual perturbations of f_o under the four feedback delay conditions (10, 40, 70, and 100 ms), with the order counterbalanced across participants. Each delay condition (except for the last condition) was presented for two blocks of trials; the first block was a "shift-up" block, and the second block was a "control" block. No control block was utilized after the final condition. Each block contained 60 trials.

The shift-up block consisted trials of four ordered stages: baseline, ramp, hold, and after-effect. The first 15 utterances, referred to as the *baseline* phase, were produced while receiving typical (unperturbed) feedback. In the following 15 trials, referred to as the *ramp* phase, the f_o in the auditory feedback increased by 6.67 cents with each successive trial, reaching a total level of 100 cents of perturbation above the participant's true f_o. For the next 15 trials, referred to as the *hold* phase, the f_o of the auditory feedback was maintained at the level of +100 cents of perturbation. In the last 15 trials, referred to as the *after-effect* phase, the f_o of auditory feedback was returned back to the participant's true f_o, similar to baseline phase. All vocal perturbations in the ramp and hold phases were maintained throughout the entire period of voicing. In the control block, participants received unperturbed f_o feedback during all 60 trials to determine natural variability in vocalization over the course of the experimental task. The amplification of the earphone signals relative to microphone signal was maintained at a set gain of +5 dB for each delay condition throughout all trials.

Experiment 2 (f_o Perturbations With Varying Amplification of Auditory Feedback)

Twelve additional participants (six women and six men) completed Experiment 2. All aspects of this experiment were identical to Experiment 1 except that feedback amplification was manipulated instead of the feedback delay (see Figure 1 for order of condition presentation). The auditory feedback signal had an intrinsic delay of approximately 10 ms, which was kept constant throughout Experiment 2.

f_o Reflex

The reflexive responses to f_o were recorded under three feedback amplification conditions: −10, +5, and +10 dB relative to microphone signal, with the order counterbalanced across participants. The earphone amplifier generated the gain for each condition relative to microphone, which was maintained throughout each block.

f_o Adaptation

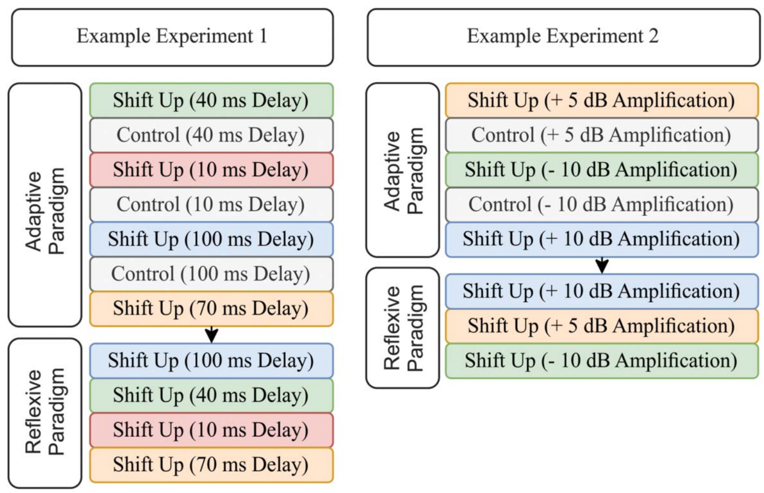
The adaptive responses to f_o were recorded under three feedback amplification conditions: −10, +5, and +10 dB, with the order counterbalanced across participants. Each amplification condition (except for the last condition) was presented for two blocks of trials, which consisted a shift-up block and a control block similar to Experiment 1.

Amplification and Calibration

In all experiments, the auditory feedback was amplified to a set gain relative to the microphone signal in order to minimize air- and bone-conducted unaltered feedback (Cornelisse et al., 1991). A 2-cc coupler (Type 4946, Bruel

[2]A cent is a logarithmic unit of measure of change in frequency (100 cents = 1 semitone).

Figure 1. Order of auditory feedback condition presentation for each paradigm in Experiments 1 and 2. The adaptive paradigm was conducted prior to the reflexive paradigm in each experiment. The order of auditory feedback delay and amplification condition presentation was counterbalanced across participants to eliminate order effects of condition. The figure illustrates an example pseudorandom ordering of condition presentation.



and Kjaer, Inc.) connected to a sound level meter (Type 2250A Handheld Analyzer with Type 4947 ½-in. Pressure Field Microphone, Bruel & Kjaer, Inc.) was utilized for calibration of auditory feedback gain levels. The earphone amplifier channel gains were adjusted such that a 1-kHz tone generated via a handheld voice recorder (Olympus LS-10 Linear PCM Recorder) positioned 7 cm from the microphone resulted in a sound pressure level gain of approximately +5 dB at the headphones for Experiment 1. For Experiment 2, the intensity was calibrated with a gain of −10, +5, or +10 dB at headphones relative to microphone.

Vocal Intensity Monitoring

Participants were asked to maintain a sound pressure level of approximately 75 dB SPL during vocalizations, resulting in an approximate insert earphone feedback level of 80 dB SPL for Experiment 1 and 65, 80, and 85 dB SPL for Experiment 2. The experimenter monitored the microphone signal using a custom sound pressure level meter.[3] The experimenter provided manual feedback (i.e., hand gestures) to the participant during the practice sessions and the first few trials of each experiment whenever the sound pressure level varied more than ± 2 dB from the target level of 75 dB SPL. Sixteen participants required feedback during approximately 10% of the experimental trials to bring the sound pressure level back within the desired bounds. All trials were utilized in data analysis regardless of investigator feedback provision.

Data and Statistical Analysis

Data analysis was carried out utilizing scripts written in MATLAB with functions to offload digitized signal f_o detection to Praat software (Boersma & Weenink, 2016; Versions 5–6.0.40). f_o traces for each trial were calculated utilizing the autocorrelation method. All acoustic signals were inspected manually for any abnormalities and corrected for f_o tracking issues. All participant responses were utilized for statistical analysis, regardless of whether they were compensatory, following, or nonresponse in nature. When f_o mistracking was present, mainly due to vocal fry, f_o traces were regenerated utilizing Praat software incorporating custom frequency bounds based on each participant's estimated vocal f_o. Auditory feedback delay between the digitized mic signal and earphone signal was calculated via a cross-correlation technique and visually inspected for errors in measurement.

Reflexive Responses

Perturbation onsets of perturbed trials of the reflexive paradigm tasks were manually annotated from the auditory feedback signal. For all perturbed trials, a region of 1,100 ms was selected, consisting of 100 ms of preshift baseline and 1,000 ms of postshift perturbed segments. The baseline region f_o trace mean was calculated and denoted as the reference frequency. Each perturbed trial f_o trace was

[3]The custom sound pressure level meter was an Arduino-based hardware connected to the computer and controlled via the custom MATLAB scripts throughout the experiments. It was equipped with an LED panel similar to a Dorrough meter to indicate when the sound pressure level at the microphone exceeded the range of 75 ± 2 dB SPL.

then converted to cents,[4] utilizing the said reference frequency. The resultant reflexive response segments were averaged per each participant to generate participant mean f_o responses per each reflexive condition. Grand-averaged mean f_o traces for each condition for all participants were also calculated by taking the mean f_o response of all participant responses. In both cases, all trials were utilized, regardless of whether they were compensatory, following, or nonresponse in nature. The *reflexive fo response magnitude* was defined as the mean f_o response of an analysis window of 120–240 ms from perturbation onset as per prior literature (Burnett & Larson, 2002; Burnett et al., 1997; Larson et al., 2001).

Adaptive Responses

For each trial, the f_o trace was selected manually starting at the vocal onset of vowel production, utilizing a custom MATLAB script to visualize the microphone signal, its frequency spectrum, and its f_o trace. The vocal motor control system generates initial feedforward motor commands and intermittent feedback responses based on adaptive error correction mechanisms of the feedback system (Guenther, 2016). According to prior studies on reflexive studies of f_o, the feedback responses are known to occur around 100–150 ms from vocal onset of the responses (Burnett & Larson, 2002; Burnett et al., 1997; Larson et al., 2001). To encapsulate only the feedforward signals and to disregard initial f_o fluctuations that occur due to initialization of voicing, mean f_o responses per trial were calculated over an analysis window of 40–120 ms from vocal onset.

The mean f_o response of the first 15 trials of the adaptation paradigm (i.e., the baseline trials) was utilized to generate the mean baseline f_o value and was denoted as the reference frequency per each participant for each adaptive condition. The mean f_o response of each trial was converted to cents,[5] utilizing the reference frequency. In order to remove the effect of the natural f_o variation over the course of 60 trials, the mean f_o response across the control block trials were subtracted from respective trials of the shift-up blocks to determine the normalized mean f_o responses. The *adaptive fo response magnitude of the hold phase* and the *adaptive fo response magnitude of the after-effect phase* were defined as the normalized mean f_o responses of each feedback condition, calculated over an analysis window of 40–120 ms from vocal onset, for trials of each respective phase. All participant responses were utilized for statistical analysis, regardless of whether they were compensatory, following, or nonresponse in nature. The normalized f_o response magnitudes of adaptive responses during the hold and after-effect phases were calculated per each subject and then consolidated per each feedback condition to get group mean f_o responses per condition.

Statistical Analysis

The distributions of reflexive and adaptive f_o response magnitudes met criteria for parametric testing. Six repeated-measures one-way analyses of variance were performed to identify differences in each f_o response magnitude (i.e., reflexive, adaptive hold phase, and adaptive after-effect phase) due to condition (i.e., auditory feedback delay or amplification condition). An alpha level of .05 or less was used for significance testing. Effect sizes were calculated using a squared partial curvilinear correlation (η_p^2). All statistical analyses were conducted in Minitab (2019) and tabulated in Table 1.

Results

Experiment 1

On average, reflexive f_o responses were compensatory (see Figure 2A), which is consistent with prior research (Hain et al., 2001). Speakers generally reduced their f_o in response to the perturbation. Mean reflexive response magnitude is shown for each delay condition for the reflexive paradigm in Figure 2B. Reflexive f_o response magnitude showed no statistically significant main effect of condition ($F = 0.16$, $p = .925$).

Adaptive f_o responses were also generally compensatory, consistent with prior research (Jones & Munhall, 2000, 2005; see Figure 3A). Speakers generally reduced their f_o during the ramp and hold phases of their shift-up responses. The mean adaptive f_o response magnitudes during the hold and after-effect phases are shown for each delay condition for the adaptation paradigm in Figures 3B and 3C, respectively. The hold phase and after-effect phase adaptive f_o response magnitudes showed no statistically significant main effects for condition (hold phase adaptive f_o response magnitude: $F = 1.07$, $p = .375$; after-effect phase adaptive f_o response magnitude: $F = 1.60$, $p = .207$).

Experiment 2

On average, reflexive f_o responses were compensatory, again consistent with prior research (Jones & Munhall, 2000, 2005; see Figure 4A). Speakers generally reduced their f_o in response to the perturbation. The mean reflexive f_o response magnitude for each amplification condition is shown in Figure 4B. The reflexive f_o response magnitude showed no statistically significant main effect of condition ($F = 0.88$, $p = .43$).

On average, adaptive f_o responses were compensatory, which is consistent with prior research (Jones & Munhall, 2000, 2005; see Figure 5A). Speakers generally reduced their f_o during the ramp and hold phases of their "shift-up" responses. Adaptive f_o response magnitude group means during the hold and after-effect phases are shown for each

[4] $f_o\ (\text{cents}) = 1200\ \log_2\left(\frac{f_o\ (\text{Hz})}{f_{oref}\ (\text{Hz})}\right)$, where f_{oref} (Hz) = the mean f_o of the 100-ms preshift baseline region of each perturbed trial of the reflexive paradigm. Cent normalization was utilized as (a) the auditory perceptual system is roughly logarithmic and (b) for comparison with previous studies that utilized cents.

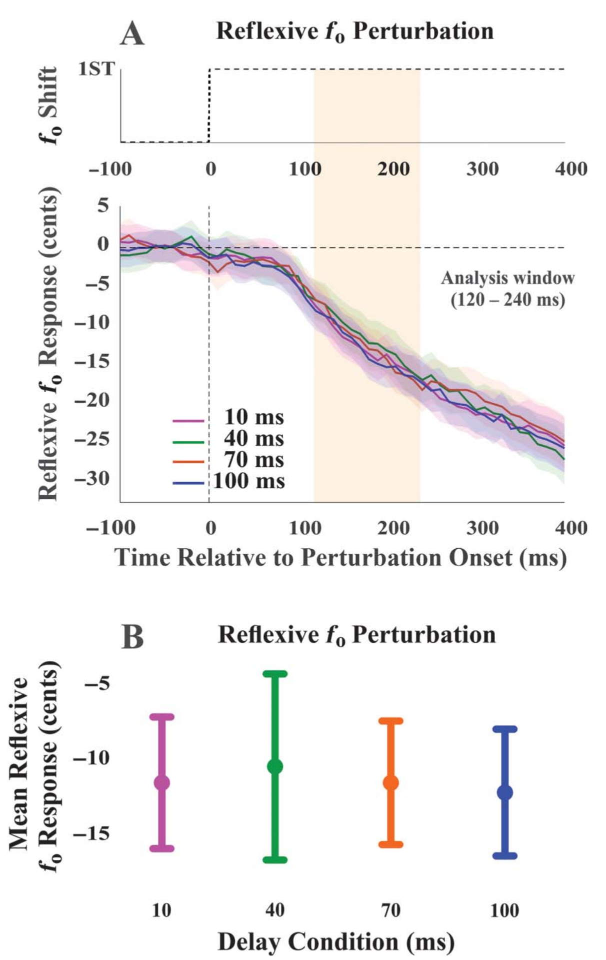
[5] $f_o\ (\text{cents}) = 1200\ \log_2\left(\frac{f_o\ (\text{Hz})}{f_{oref}\ (\text{Hz})}\right)$, where f_{oref} (Hz) = mean f_o of the first 15 baseline phase trials of the adaptation paradigm.

Table 1. Repeated-measures one-way analysis of variance on vocal fundamental frequency (f_o) response magnitudes.

Response type	Effect	*df*	η_p^2	*F*	*p*
Reflexive f_o	Delay condition	3	—	0.16	.925
Adaptive f_o (hold)	Delay condition	3	—	1.07	.375
Adaptive f_o (after-effect)	Delay condition	3	—	1.60	.207
Reflexive f_o	Amplification condition	2	—	0.88	.430
Adaptive f_o (hold)	Amplification condition	2	0.29	4.45	.024*
Adaptive f_o (after-effect)	Amplification condition	2	0.34	5.63	.011*
Post hoc: Dunnett's test				***T***	p_{adj}
Adaptive f_o (hold)	**Amplification condition**				
(–10 dB SPL) & (10 dB SPL)	(10 dB SPL control condition)			2.73	.023*
(5 dB SPL) & (10 dB SPL)	(10 dB SPL control condition)			0.31	.931
Adaptive f_o (after-effect)	**Amplification condition**				
(–10 dB SPL) & (10 dB SPL)	(10 dB SPL control condition)			3.78	.007*
(5 dB SPL) & (10 dB SPL)	(10 dB SPL control condition)			1.01	.504

*Significant at $p < .05$, $p_{adj} < .05$.

Figure 2. Experiment 1: Average reflexive responses by delay condition. (A) Fundamental frequency (f_o) perturbation and reflexive f_o responses versus time relative to perturbation onset. The shadings indicate 95% confidence intervals. (B) Mean reflexive f_o response magnitude per condition. The error bars indicate 95% confidence intervals.



amplification condition for the adaptation paradigm in Figures 5B and 5C, respectively. The hold phase and after-effect phase adaptive f_o response magnitudes both showed statistically significant main effects of condition (hold phase adaptive f_o response magnitude: $F = 4.45$, $p = .024$, $\eta_p^2 = .29$; after-effect phase adaptive f_o response magnitude: $F = 5.63$, $p = .011$, $\eta_p^2 = .34$).

Dunnett's post hoc tests were carried out, comparing the −10- and +5-dB conditions to the +10-dB condition. Results indicated that, for both the hold and after-effect phases, the response magnitudes during the −10-dB condition were statistically significantly larger than those from the +10-dB condition (hold phase adaptive f_o response magnitude: $T = 2.73$, $p_{adj} = .023$; after-effect phase adaptive f_o response magnitude: $T = 3.28$, $p_{adj} = .007$). In contrast, for both phases, there was no statistically significant difference between the response magnitudes during the +5- and +10-dB conditions (hold phase adaptive f_o response magnitude: $T = .31$, $p_{adj} = .931$; after-effect phase adaptive response magnitude: $T = 1.01$, $p_{adj} = .504$).

Discussion

The current study investigated the effects of delay and amplification in the auditory feedback presented to participants in reflexive and adaptive f_o perturbations of auditory feedback. The results indicate that auditory feedback delays in a range of 10–100 ms do not elicit any statistically significant differences in the f_o response magnitudes for either the reflexive or adaptive paradigm. Similarly, the results indicate no statistically significant differences in the response magnitudes for the reflexive f_o perturbations when the auditory feedback amplification is in the range of −10 to +10 dB relative to microphone signal. However, there was a statistically significant difference in the response magnitudes to adaptive f_o perturbations, between the −10 and +10 dB auditory amplification feedback conditions, with a large effect size.

Figure 3. Experiment 1: Average adaptive responses by delay condition. (A) Fundamental frequency (f_o) perturbation and adaptive responses f_o across three trial blocks per condition. The shadings indicate 95% confidence intervals. (B) Hold phase mean adaptive f_o response magnitude per condition. (C) After-effect phase mean adaptive f_o response magnitude per condition. The error bars indicate 95% confidence intervals.

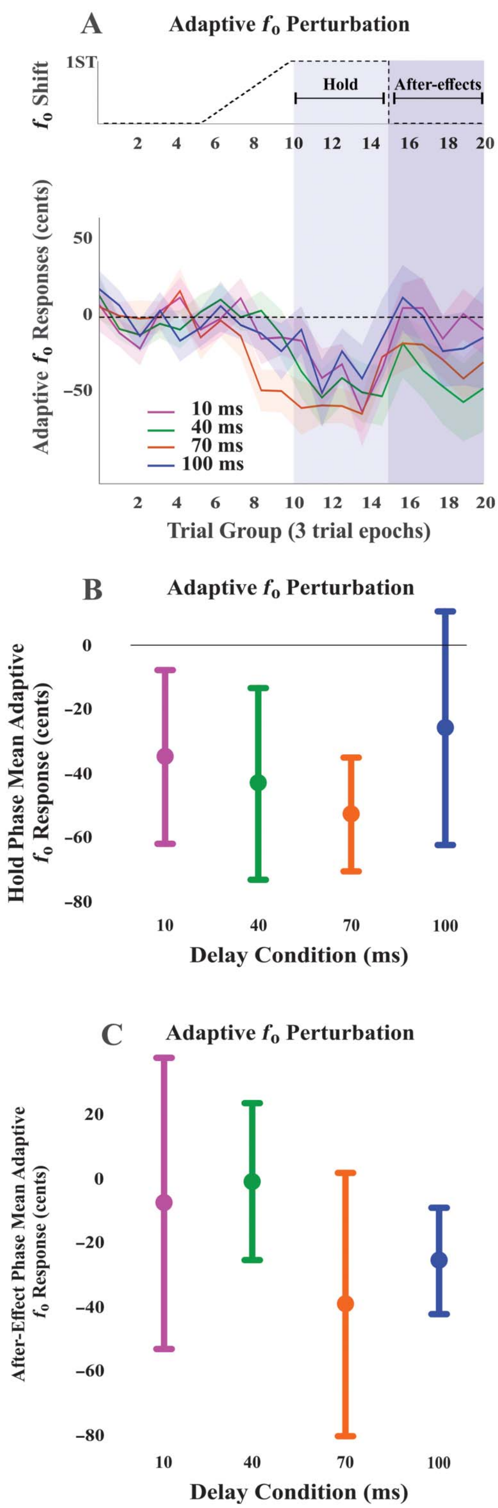
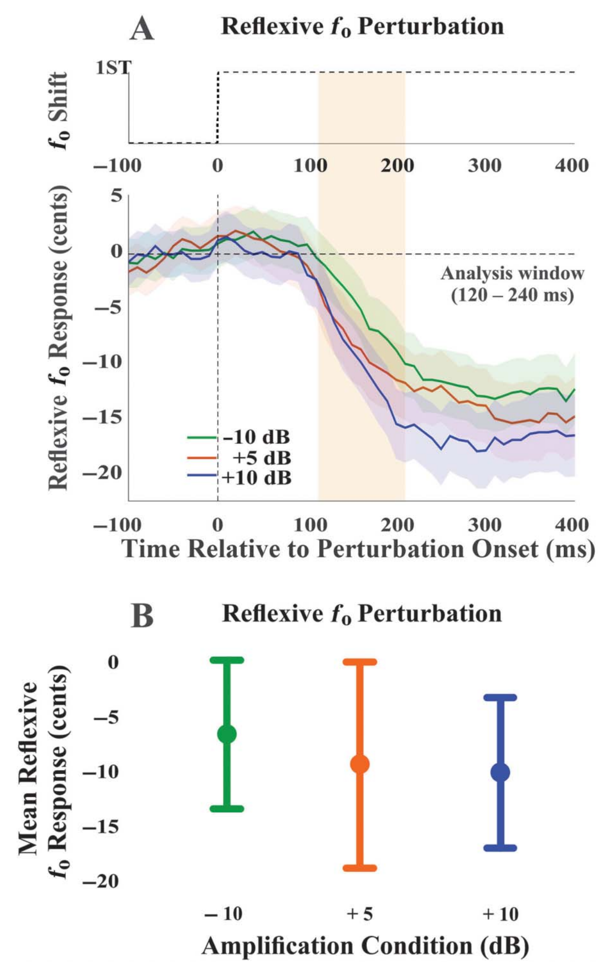


Figure 4. Experiment 2: Average reflexive responses by amplification condition. (A) Fundamental frequency (f_o) perturbation and reflexive f_o responses versus time relative to perturbation onset. The shadings indicate 95% confidence intervals. (B) Mean reflexive f_o response magnitude per condition. The error bars indicate 95% confidence intervals.

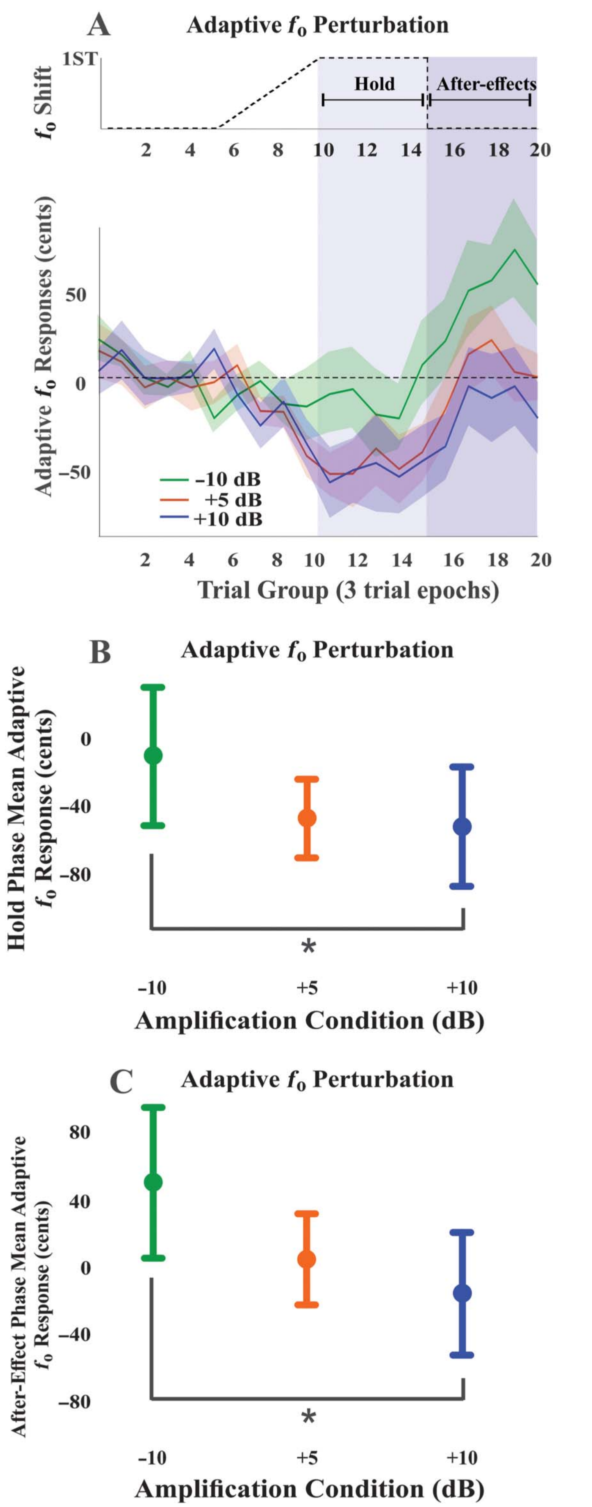


Effects of Delayed Auditory Feedback

As hypothesized, reflexive f_o response magnitudes did not vary across conditions (see Figure 2). These results are consistent with the previous f_o study by Hain et al. (2001), in which no statistically significant effects of feedback delay on the f_o response peak magnitude were found. However, the prior study reported that there was a nonsignificant increase in peak magnitude for delayed conditions compared to no delay condition. In contrast, in this study, reflexive f_o response magnitudes showed no clear trends across feedback delay conditions.

Contrary to our hypothesis, the adaptive f_o response magnitudes were not reduced when feedback delay was increased (see Figure 3). These results are in contrast to the prior research on effects of delayed auditory feedback on responses to a formant adaptation paradigm (Max & Maffett, 2015). However, vowel formants are segmental features

Figure 5. Experiment 2: Average adaptive responses by amplification condition. Fundamental frequency (f_o) perturbation and adaptive f_o responses across three trial blocks per condition. The shadings indicate 95% confidence intervals. (B) Hold phase mean adaptive f_o response magnitude per condition. (C) After-effect phase mean adaptive f_o response magnitude per condition for the adaptation paradigm. The error bars indicate 95% confidence intervals.



of speech motor control, which are thought to occur in a smaller time scale. Voice f_o is a suprasegmental feature that involves planning over longer time scales, such as over prosodic contours in a phrase. Therefore, the intrinsic delays of formant control mechanisms are expected to be smaller compared to those of the vocal f_o control mechanisms (Guenther, 2016). If the internal processing delays of the speech motor control mechanism are larger, external delays employed by the study will need to be comparably larger to affect the control mechanisms. This may explain why the feedback delay conditions applied in the current study for voice f_o did not elicit effects.

The current study can be extended to examine the effects of longer delays in auditory feedback to vocal f_o control mechanisms. Given that the prosodic subsystem is expected to have larger intrinsic delays (Guenther, 2016), there could be a reduction in f_o compensatory response magnitudes to the adaptive paradigms of f_o perturbations for delays larger than 100 ms. However, when the delays become considerably larger, the vocal f_o control system may consider auditory feedback to be unreliable and utilize other feedback modalities (i.e., somatosensory feedback) to update the feedforward control system. In that case, minimal f_o compensatory response magnitudes will be observed in the adaptive paradigm.

This study examined delays in the range that may be expected for current equipment setups. The results suggest that delays in the range of 10–100 ms do not significantly influence feedback or feedforward control mechanisms of voice f_o.

Effects of Auditory Feedback Amplification

We hypothesized that neither the reflexive nor the adaptive f_o compensatory response magnitudes would be affected by the amplification level provided in the setup. There was no statistically significant differences in reflexive f_o response magnitudes with the level of feedback amplification (see Figure 4). This result is analogous to the previous study of Burnett et al. (1998), in which there were no statistically significant differences in the f_o response magnitudes across conditions. However, we did notice a nonsignificant trend in which response magnitudes decreased as the amplification level was increased. Moreover, our findings suggest that the level of amplification is a statistically significant factor in the magnitude of the responses to the f_o adaptive paradigms. There was a statistically significant decrease in hold phase and after-effect phase adaptive f_o response magnitudes for auditory feedback amplification of −10 dB compared to +10 dB (see Figure 5). This is in contrast to the two prior studies that examined headphone/earphone occlusion (Franken et al., 2019; Mitsuya & Purcell, 2016), which found no statistically significant differences in f_o response magnitudes for different levels of occlusion. Both studies utilized +10-dB amplification in headphones relative to the microphone. One possible explanation is that the gain difference between −10- and +10-dB conditions in the current study using insert earphones (higher than 30 dB attenuation; Etymotic Research, Inc.) was larger than the difference in attenuation levels of the headphones utilized in the prior study (25 dBA; Franken et al., 2019). However, there was no statistically significant difference between the response magnitudes of +10- and +5-dB conditions. Thus, we can infer that +5 dB is a sufficient amplification level to utilize in both reflexive and adaptive study setups.

Study Limitations

This study required strict control of system-specific delay perturbations in order to maintain consistent auditory feedback delays in each condition of the experiment. The Eclipse V4 Harmonizer had an intrinsic delay of 10 ms (± 5 ms), which was considered as the first auditory delay condition. Delays were increased in consistent step sizes of 30 ms for each feedback delay condition. However, due to the f_o shifting algorithms, each delay increment had a variance of ± 5 ms. While these variations are considerably smaller than the delay conditions of the study, this was a hardware-specific limitation of the study.

The sound pressure level of the participants' vocalizations was maintained around 75 dB SPL throughout the study to ensure specific levels of auditory feedback intensities were provided to participants in each auditory feedback amplification condition. The investigators were careful only to provide feedback to the participants during the practice sessions and the initial five to 10 trials of each paradigm. However, there were a few instances (16 participants and approximately 10% of total trials per participant) in which investigators were required to provide feedback to participants midway through the paradigms to ensure that sound pressure levels of participant vocalizations were within required bounds. The feedback provided in these instances may have affected the f_o responses of the participants. However, prior research has failed to identify a significant relationship between f_o and sound pressure level control of vocal productions (Burnett et al., 1997).

Presenting successive adaptation conditions could elicit fatigue or long-term adaptation in participants, which could affect the responses. However, each adaptation paradigm was only 6–7 min in duration and presented in an interleaved fashion with control conditions placed to washout the effects of long-term adaptation in the vocal control system. Moreover, each delay or amplification condition was presented in a counterbalanced order such that there were no order effects across participants.

Implications

This study is an assurance to the scientific community that potential unintended equipment-specific delays in auditory feedback of f_o perturbation paradigms do not significantly affect the comparability of results of similar studies generated via varied experimental platforms. Moreover, the current study examines sufficient bounds to the amplification of auditory feedback that can be applied such that there is minimal impact from air- and bone-conducted feedback such that reflexive and adaptive paradigms yield robust vocal responses. While higher amplification levels would ensure absolute occlusion from air- and bone-conducted auditory feedback to the participants, consideration should be given as to whether it is necessary to do so, as the participants could be at a risk of deteriorations in hearing sensitivity due to low-level exposure to high auditory feedback intensities over considerable durations of time (in the scale of hours; Liberman & Kujawa, 2017; Pienkowski, 2017). Studies have also observed that higher levels of intensity in auditory feedback and masking noise elicit a confounding Lombard effect in the participants' vocalizations (i.e., an increase in vocalization intensity as a result of increase in environmental auditory feedback intensity; Lane & Tranel, 1971; Parrell et al., 2017). According to the results of the current study, +5-dB amplification is sufficient to effectively occlude air- and bone-conducted feedback.

Conclusions

The current study systematically explored the effects of auditory feedback delay and amplification on responses to vocal f_o perturbation paradigms. Auditory feedback delays in a range of 10–100 ms did not elicit statistically significant differences in the f_o response magnitudes for either reflexive or adaptive paradigms of f_o perturbations. Similarly, there were no statistically significant differences in the response magnitudes for the reflexive f_o perturbations when the auditory feedback amplification was varied in the range of −10 to +10 dB relative to the microphone signal. In contrast, response magnitudes to the adaptation f_o perturbation paradigm were statistically significantly different between the amplification conditions of −10 and + 10 dB. This study indicates that results of similar studies of f_o perturbation paradigms are comparable across varied experimental platforms with variable feedback delays. Furthermore, results indicate that a +5-dB amplification in auditory feedback relative microphone signal is sufficient to yield robust responses to both reflexive and adaptive paradigms of f_o perturbation in typical speakers.

Acknowledgments

This research was supported by Grants P50 DC015446 (Robert E. Hillman), T32 DC013017 (Christoper A. Moore), and R01 DC016270 (Cara E. Stepp and Frank H. Guenther). We would like to acknowledge Elizabeth Heller Murray and Manuel Diaz Cádiz for assistance with equipment and software troubleshooting, Daniel Buckley and Kimberly Dahl for performing auditory perceptual ratings, Jennifer Vojtech for review of the article, and Joseph Perkell for inspiring insight at different stages of the study.

References

Abur, D., Lester-Smith, R. A., Daliri, A., Lupiani, A. A., Guenther, F. H., & Stepp, C. E. (2018). Sensorimotor adaptation of voice

fundamental frequency in Parkinson's disease. *PLOS ONE, 13*(1), e0191839. https://doi.org/10.1371/journal.pone.0191839
American National Standards Institute. (1996). *Specifications for audiometers* (ANSI S3. 6-1996).
American National Standards Institute. (2004). *Specifications for audiometers* (ANSI S3. 6-2004).
Arbeiter, M., Petermann, S., Hoppe, U., Bohr, C., Doellinger, M., & Ziethe, A. (2018). Analysis of the auditory feedback and phonation in normal voices. *Annals of Otology, Rhinology & Laryngology, 127*(2), 89–98. https://doi.org/10.1177/0003489417744567
Ballard, K. J., Halaki, M., Sowman, P., Kha, A., Daliri, A., Robin, D. A., Tourville, J. A., & Guenther, F. H. (2018). An investigation of compensation and adaptation to auditory perturbations in individuals with acquired apraxia of speech. *Frontiers in Human Neuroscience, 12,* 510. https://doi.org/10.3389/fnhum.2018.00510
Bauer, J. J., & Larson, C. R. (2003). Audio-vocal responses to repetitive pitch-shift stimulation during a sustained vocalization: Improvements in methodology for the pitch-shifting technique. *The Journal of the Acoustical Society of America, 114*(2), 1048–1054. https://doi.org/10.1121/1.1592161
Bauer, J. J., Mittal, J., Larson, C. R., & Hain, T. C. (2006). Vocal responses to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude. *The Journal of the Acoustical Society of America, 119*(4), 2363–2371. https://doi.org/10.1121/1.2173513
Behroozmand, R., Korzyukov, O., Sattler, L., & Larson, C. R. (2012). Opposing and following vocal responses to pitch-shifted auditory feedback: Evidence for different mechanisms of voice pitch control. *The Journal of the Acoustical Society of America, 132*(4), 2468–2477. https://doi.org/10.1121/1.4746984
Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience, 12*(1), 54. https://doi.org/10.1186/1471-2202-12-54
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
Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* [Computer program] (Version 6.0. 14). http://www.praat.org/
Burk, M. H., & Wiley, T. L. (2004). Continuous versus pulsed tones in audiometry. *American Journal of Audiology, 13*(1), 54–61. https://doi.org/10.1044/1059-0889(2004/008)
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., & 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., McCurdy, K. E., & Bright, J. C. (2008). Reflexive and volitional voice fundamental frequency responses to an anticipated feedback pitch error. *Experimental Brain Research, 191*(3), 341. https://doi.org/10.1007/s00221-008-1529-z
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
Chen, S. H., Liu, H., Xu, Y., & Larson, C. R. (2007). Voice F_0 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, Z., Liu, P., Jones, J. A., Huang, D., & Liu, H. (2010). Sex-related differences in vocal responses to pitch feedback perturbations during sustained vocalization. *The Journal of the Acoustical Society of America, 128*(6), EL355–EL360. https://doi.org/10.1121/1.3509124
Cornelisse, L. E., Gagné, J., & Seewald, R. C. (1991). Ear level recordings of the long-term average spectrum of speech. *Ear and Hearing, 12*(1), 47–54. https://doi.org/10.1097/00003446-199102000-00006
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), e12521. https://doi.org/10.1111/desc.12521
Dean, M. S., & Martin, F. N. (2000). Insert earphone depth and the occlusion effect. *American Journal of Audiology, 9*(2), 131–134. https://doi.org/10.1044/1059-0889(2000/011)
Demopoulos, C., Kothare, H., Mizuiri, D., Henderson-Sabes, J., Fregeau, B., Tjernagel, J., Houde, J. F., Sherr, E. H., & Nagarajan, S. S. (2018). Abnormal speech motor control in individuals with 16p11.2 deletions. *Scientific Reports, 8*(1), 1274. https://doi.org/10.1038/s41598-018-19751-x
Donath, T. M., Natke, U., & Kalveram, K. T. (2002). Effects of frequency-shifted auditory feedback on voice F_0 contours in syllables. *The Journal of the Acoustical Society of America, 111*(1), 357–366. https://doi.org/10.1121/1.1424870
Elman, J. L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *The Journal of the Acoustical Society of America, 70*(1), 45–50. https://doi.org/10.1121/1.386580
Ertmer, D. J., Strong, L. M., & Sadagopan, N. (2003). Beginning to communicate after cochlear implantation: Oral language development in a young child. *Journal of Speech, Language, and Hearing Research, 46*(2), 328–340. https://doi.org/10.1044/1092-4388(2003/026)
Franken, M. K., Hartsuiker, R. J., Johansson, P., Hall, L., Wartenberg, T., & Lind, A. (2019). Does passive sound attenuation affect responses to pitch-shifted auditory feedback? *The Journal of the Acoustical Society of America, 146*(6), 4108–4121. https://doi.org/10.1121/1.5134449
Guenther, F. H. (2016). *Neural control of speech.* MIT Press. https://doi.org/10.7551/mitpress/10471.001.0001
Guo, Z., Wu, X., Li, W., Jones, J. A., Yan, N., Sheft, S., Liu, P., & Liu, H. (2017). Top-down modulation of auditory-motor integration during speech production: The role of working memory. *The Journal of Neuroscience, 37*(43), 10323–10333. https://doi.org/10.1523/JNEUROSCI.1329-17.2017
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
Hain, T. C., Burnett, T. A., Larson, C. R., & Kiran, S. (2001). Effects of delayed auditory feedback (DAF) on the pitch-shift reflex. *The Journal of the Acoustical Society of America, 109*(5), 2146–2152. https://doi.org/10.1121/1.1366319
Hawco, C. S., & Jones, J. A. (2010). Multiple instances of vocal sensorimotor adaptation to frequency-altered feedback within a single experimental session. *The Journal of the Acoustical Society of America, 127*(1), EL13–EL18. https://doi.org/10.1121/1.3272633
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

Hixon, T. J., Weismer, G., & Hoit, J. D. (2018). *Preclinical speech science: Anatomy, physiology, acoustics, and perception*. Plural.

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

Huang, X., Fan, H., Li, J., Jones, J. A., Wang, E. Q., Chen, L., Chen, X., & Liu, H. (2019). External cueing facilitates auditory-motor integration for speech control in individuals with Parkinson's disease. *Neurobiology of Aging, 76,* 96–105. https://doi.org/10.1016/j.neurobiolaging.2018.12.020

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 F0 production: Evidence from feedback perturbation. *The Journal of the Acoustical Society of America, 108*(3), 1246–1251. https://doi.org/10.1121/1.1288414

Jones, J. A., & Munhall, K. G. (2002). The role of auditory feedback during phonation: Studies of Mandarin tone production. *Journal of Phonetics, 30*(3), 303–320. https://doi.org/10.1006/jpho.2001.0160

Jones, J. A., & Munhall, K. G. (2005). Remapping auditory-motor representations in voice production. *Current Biology, 15*(19), 1768–1772. https://doi.org/10.1016/j.cub.2005.08.063

Kawahara, H. (1993). Transformed auditory feedback: Effects of fundamental frequency perturbation. *The Journal of the Acoustical Society of America, 94*(3), 1883–1884. https://doi.org/10.1121/1.407536

Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology, 18*(2), 124–132. https://doi.org/10.1044/1058-0360(2008/08-0017)

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), 25. https://doi.org/10.1186/1471-2202-14-25

Keough, D., & Jones, J. A. (2009). The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing. *The Journal of the Acoustical Society of America, 126*(2), 837–846. https://doi.org/10.1121/1.3158600

Kirchner, J. A., & Suzuki, M. (1968). Laryngeal reflexes and voice production. *Annals of the New York Academy of Sciences, 155*(1), 98–109. https://doi.org/10.1111/j.1749-6632.1968.tb56752.x

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research, 14*(4), 677–709. https://doi.org/10.1044/jshr.1404.677

Larson, C. R., Altman, K. W., Liu, H., & Hain, T. C. (2008). Interactions between auditory and somatosensory feedback for voice F_0 control. *Experimental Brain Research, 187*(4), 613–621. https://doi.org/10.1007/s00221-008-1330-z

Larson, C. R., Burnett, T. A., Bauer, J. J., Kiran, S., & Hain, T. C. (2001). Comparison of voice F_0 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

Larson, C. R., Burnett, T. A., Kiran, S., & Hain, T. C. (2000). Effects of pitch-shift velocity on voice F_0 responses. *The Journal of the Acoustical Society of America, 107*(1), 559–564. https://doi.org/10.1121/1.428323

Larson, C. R., Sun, J., & Hain, T. C. (2007). Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F_0 and amplitude control. *The Journal of the Acoustical Society of America, 121*(5), 2862–2872. https://doi.org/10.1121/1.2715657

Liberman, M. C., & Kujawa, S. G. (2017). Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. *Hearing Research, 349,* 138–147. https://doi.org/10.1016/j.heares.2017.01.003

Liu, H., & Larson, C. R. (2007). Effects of perturbation magnitude and voice F_0 level on the pitch-shift reflex. *The Journal of the Acoustical Society of America, 122*(6), 3671–3677. https://doi.org/10.1121/1.2800254

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, 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

Liu, Y., Fan, H., Li, J., Jones, J. A., Liu, P., Zhang, B., & Liu, H. (2018). Auditory-motor control of vocal production during divided attention: Behavioral and ERP correlates. *Frontiers in Neuroscience, 12,* 113. https://doi.org/10.3389/fnins.2018.00113

MacDonald, E. N., & Munhall, K. G. (2012). A preliminary study of individual responses to real-time pitch and formant perturbations. *The Listening Talker,* 32–35.

Max, L., & Maffett, D. G. (2015). Feedback delays eliminate auditory-motor learning in speech production. *Neuroscience Letters, 591,* 25–29. https://doi.org/10.1016/j.neulet.2015.02.012

Max, L., Wallace, M. E., & Vincent, I. (2003). *Sensorimotor adaptation to auditory perturbations during speech: Acoustic and kinematic experiments*. Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain.

Minitab, L. (2019). Minitab 18 Support [Internet]. [cited 2019 May 14].

Mitsuya, T., Munhall, K. G., & Purcell, D. W. (2017). Modulation of auditory-motor learning in response to formant perturbation as a function of delayed auditory feedback. *The Journal of the Acoustical Society of America, 141*(4), 2758–2767. https://doi.org/10.1121/1.4981139

Mitsuya, T., & Purcell, D. W. (2016). Occlusion effect on compensatory formant production and voice amplitude in response to real-time perturbation. *The Journal of the Acoustical Society of America, 140*(6), 4017–4026. https://doi.org/10.1121/1.4968539

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., & Gracco, V. L. (2013). Sensorimotor adaptation of speech in Parkinson's disease. *Movement Disorders, 28*(12), 1668–1674. https://doi.org/10.1002/mds.25588

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

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

Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec, 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, 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

Penn, J. P. (1955). Voice and speech patterns of the hard of hearing. *Acta Oto-Laryngologica. Supplementum, 124,* 1–69.

Périer, O., Alegria, J., Buyse, M., D'alimonte, G., Gilson, D., & Serniclaes, W. (1984). Consequences of auditory deprivation in animals and humans. *Acta Oto-Laryngologica, 97*(Suppl. 412), 60–70. https://doi.org/10.3109/00016488409129739

Petermann, S., Döllinger, M., Kniesburges, S., & Ziethe, A. (2016). Analysis method for the neurological and physiological processes underlying the pitch-shift reflex. *Acta Acustica United With Acustica, 102*(2), 284–297. https://doi.org/10.3813/AAA.918944

Pienkowski, M. (2017). On the etiology of listening difficulties in noise despite clinically normal audiograms. *Ear and Hearing, 38*(2), 135–148. https://doi.org/10.1097/AUD.0000000000000388

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), 16340. https://doi.org/10.1038/s41598-018-34517-1

Scheerer, N. E., Behich, J., Liu, H., & Jones, J. (2013). ERP correlates of the magnitude of pitch errors detected in the human voice. *Neuroscience, 240,* 176–185. https://doi.org/10.1016/j.neuroscience.2013.02.054

Scheerer, N. E., Jacobson, D., & 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., & Jones, J. A. (2012). The relationship between vocal accuracy and variability to the level of compensation to altered auditory feedback. *Neuroscience Letters, 529*(2), 128–132. https://doi.org/10.1016/j.neulet.2012.09.012

Scheerer, N. E., & Jones, J. A. (2018). Detecting our own vocal errors: An event-related study of the thresholds for perceiving and compensating for vocal pitch errors. *Neuropsychologia, 114,* 158–167. https://doi.org/10.1016/j.neuropsychologia.2017.12.007

Sivasankar, M., Bauer, J. J., Babu, T., & Larson, C. R. (2005). Voice responses to changes in pitch of voice or tone auditory feedback. *The Journal of the Acoustical Society of America, 117*(2), 850–857. https://doi.org/10.1121/1.1849933

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

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin, 60*(3), 213–232. https://doi.org/10.1037/h0044155

Ziethe, A., Petermann, S., Hoppe, U., Greiner, N., Brüning, M., Bohr, C., & Döllinger, M. (2019). Control of fundamental frequency in dysphonic patients during phonation and speech. *Journal of Voice, 33*(6), 851–859. https://doi.org/10.1016/j.jvoice.2018.07.001

Appendix (p. 1 of 2)

Auditory Feedback Delays and Auditory Feedback Amplifications of Prior Research

Table A1. Apparatus-specific auditory feedback delays reported across prior research.

Study	Perturbation and study paradigm	Delay (ms)
Jones & Munhall (2000, 2002, 2005)	f_o perturbations, adaptive paradigm	3–4
Keough et al. (2013)	f_o perturbations, adaptive paradigm	10
Daliri et al. (2018), Mollaei et al. (2016, 2013)	f_o perturbations, adaptive paradigm	14
Abur et al. (2018), Stepp et al. (2017)	f_o perturbations, adaptive paradigm	45
Burnett et al. (1998)	f_o and vocal intensity perturbations, reflexive paradigm	8–20
Burnett & Larson (2002), Burnett et al. (2008)	f_o perturbations, reflexive paradigm	8–20
Hain et al. (2000), Scheerer et al. (2013, 2016)	f_o perturbations, reflexive paradigm	10
Demopoulos et al. (2018), Houde et al. (2019)	f_o perturbations, reflexive paradigm	12
Larson et al. (2007)	f_o perturbations, reflexive paradigm	14
Max et al. (2003)	f_o perturbations, reflexive paradigm	20
Sares et al. (2018)	f_o perturbations, reflexive paradigm	25
Bauer et al. (2006), Bauer & Larson (2003), Behroozmand et al. (2012, 2018), Behroozmand & Larson (2011), Burnett et al. (1998, 1997), S. H. Chen et al. (2007), Z. Chen et al. (2010), Donath et al. (2002), Elman (1981), Franken et al. (2019), Guo et al. (2017), Jones & Keough (2008), Hain et al. (2001), Hawco & Jones (2010), Huang et al. (2019), Keough & Jones (2009), Larson et al. (2008, 2001, 2000), H. Liu & Larson (2007), H. Liu, Russo, et al. (2010), P. Liu, Chen, et al. (2010), Y. Liu et al. (2018), MacDonald & Munhall (2012), Natke et al. (2003), S. Patel et al. (2014), Scheerer & Jones (2012, 2018), Sivasankar et al. (2005)		Equipment and software related feedback delays not specified

Note. f_o = fundamental frequency.

Table A2. Auditory feedback amplification methodologies utilized across prior research.

Study	Headphone amplification	+ Masking noise	Vocal intensity monitoring
f_o perturbation adaptve paradigm			
Jones & Keough (2008)	Amplification level not reported	Multitalker babble at 75 dB SPL	Loudness meter on computer screen
MacDonald & Munhall (2012)	80 dBA SPL	50 dBA SPL masking noise	Not reported
Keough et al. (2013)	Amplification level not reported	Multitalker babble at 70 dB SPL	Not reported
Hawco & Jones (2010)	+25 dB gain relative to microphone signal	Multitalker babble at 90 dB SPL	Not reported
Keough & Jones (2009)	Amplification level not reported	Multitalker babble at 80 dB SPL	Not reported
Hawco & Jones (2010)	85 dB SPL	Pink masking noise at 70 dB SPL	Loudnesss meter used to maintain subject volume at 75 dB SPL
Jones & Munhall (2000, 2002, 2005)	Amplification level not reported	Pink noise and multitalker babble played at 75 dB SPL	Reported that subject wer enot asked to maintain a specific loudness level
Behroozmand et al. (2018)	+10 dB gain relative to microphone signal	Not applied	Participants maintained steady vocalization/method not reported
Daliri et al. (2018)	Amplification applied/ level not reported	Not applied	Visual feedback presented (corresponding to whether or not participant's voice intensity was in the range of 72–88 dB SPL)
Abur et al. (2018), Stepp et al. (2017)	+5 dB gain relative to microphone signal	Not applied	Not reported

(*table continues*)

Appendix (p. 2 of 2)

Auditory Feedback Delays and Auditory Feedback Amplifications of Prior Research

Table A2.

Study	Headphone amplification	+ Masking noise	Vocal intensity monitoring
f_o perturbation reflexive paradigm			
Behroozmand et al. (2012), Behroozmand & Larson (2011), Huang et al. (2019), Y. Liu et al. (2018)	+10 dB gain relative to microphone	Not applied	Participants maintained steady vocalization/method not reported
Sares et al. (2018)	74–78 dB SPL	Pink noise applied/ 64–69 dB SPL	Not reported
Ballard et al. (2018)	+10 dB gain relative to microphone	Not applied	For vocal intensity outside 70–75 dB, an error signal was displayed after the trial (i.e., “too loud/soft”)
Arbeiter et al. (2018), Petermann et al. (2016), Ziethe et al. (2019)	75 dB SPL	Pink noise applied (upper cutoff 900 Hz)	Participant vocalization measured and kept between 75 and 85 dB SPL
Guo et al. (2017)	+10 dB gain relative to microphone	Not applied	Listeners allowed to choose preferred loudness level and maintained throughout experiment
Donath et al. (2002), Natke et al. (2003)	–5 dB gain relative to microphone signal	Low-pass filtered white noise presented at 65 dBA SPL	Participants instructed to speak with normal volume
Max et al. (2003)	+2 dB gain relative to microphone signal	Not reported	Not reported
Larson et al. (2008, 2001, 2000)	–5 dB gain relative to microphone signal	Masking noise presented/ level not reported	Participants maintained vocal loudness at 70 dB SPL aided by a Dorrough loudness monitor
Hain et al. (2000, 2001)	85 dB SPL	Pink noise played at 70 dB SPL	Participants maintained vocal loudness at 70 dB SPL aided by a Dorrough loudness monitor
Elman (1981)	85 dB SPL	Not reported	Not reported
Burnett & Larson (2002), Burnett et al. (1997)	–0.6 dB gain relative to microphone signal	Not applied	Not reported
Burnett et al. (1998)	Three different levels of amplification: 65, 75, and 85 dB SPL	Four pink noise masking levels: none, 50, 60, or 70 dB SPL	Participants maintained vocal loudness at 70 dB SPL aided by a Dorrough loudness monitor
Burnett & Larson (2002)	Not applied/intensity at microphone and headphones were equivalent	Not applied	Participants maintained vocal loudness at 80 dB SPL aided by a Dorrough loudness monitor
Sivasankar et al. (2005)	88 dB SPL	Pink noise played at 70 dB SPL	Participants maintained vocal loudness at 77 dB SPL aided by a Dorrough loudness monitor
Bauer et al. (2006)	± 1, 3, or 6 dB gain relative to microphone signal	Pink noise played at 60 dB SPL	Participants maintained vocal loudness at normal 75 dB SPL or soft amplitude level 60 dB SPL aided by a Dorrough loudness monitor
Bauer & Larson (2003)	80 dB SPL	Pink noise played at 60 dB SPL	Participants maintained vocal loudness at 70 dB SPL
H. Liu & Larson (2007)	+10 dB gain relative to microphone signal	Pink noise played at 40 dB SPL	Participants maintained vocal loudness at 70 dB SPL aided by a Dorrough loudness monitor
H. Liu, Russo, & Larson (2010)	+10 dB gain relative to microphone signal	Not applied	Not tightly controlled
Z. Chen et al. (2010)	+10 dB gain relative to microphone signal	Not reported	Not reported

Note. f_o = fundamental frequency.