

Research Article

Responses to Auditory Feedback Manipulations in Speech May Be Affected by Previous Exposure to Auditory Errors

Caroline A. Niziolek^a  and Benjamin Parrell^a 

Purpose: Speakers use auditory feedback to guide their speech output, although individuals differ in the magnitude of their compensatory response to perceived errors in feedback. Little is known about the factors that contribute to the compensatory response or how fixed or flexible they are within an individual. Here, we test whether manipulating the perceived reliability of auditory feedback modulates speakers' compensation to auditory perturbations, as predicted by optimal models of sensorimotor control.

Method: Forty participants produced monosyllabic words in two separate sessions, which differed in the auditory feedback given during an initial exposure phase. In the *veridical* session exposure phase, feedback was normal. In the *noisy* session exposure phase, small, random formant perturbations were applied, reducing reliability of auditory feedback. In each session, a subsequent test phase introduced larger unpredictable formant perturbations. We assessed whether the magnitude of within-trial compensation

for these larger perturbations differed across the two sessions.

Results: Compensatory responses to downward (though not upward) formant perturbations were larger in the veridical session than the noisy session. However, in post hoc testing, we found the magnitude of this effect is highly dependent on the choice of analysis procedures. Compensation magnitude was not predicted by other production measures, such as formant variability, and was not reliably correlated across sessions.

Conclusions: Our results, though mixed, provide tentative support that the feedback control system monitors the reliability of sensory feedback. These results must be interpreted cautiously given the potentially limited stability of auditory feedback compensation measures across analysis choices and across sessions.

Supplemental Material: <https://doi.org/10.23641/asha.14167136>

Speakers use auditory feedback to guide their speech, changing their articulation and acoustic output online to correct for apparent speech errors. Evidence for the use of auditory feedback during the online control of speech articulation comes from studies that experimentally manipulate feedback in real time, introducing intermittent discrepancies between produced and observed vowel formants (Purcell & Munhall, 2006; Tourville et al., 2008). These unpredictable discrepancies introduce errors in feedback that are

corrected through *online compensation*: adjusting articulation within a given utterance. While studies of online compensation show a consistent compensatory response at a group level, individuals differ in the magnitude of their compensatory response, with some failing to respond at all (Cai et al., 2012; Parrell et al., 2017).

Interindividual differences in compensation magnitude can be understood as population-level variation in the gains used by the feedback control system to correct for sensory errors (Guenther, 2016; Houde & Nagarajan, 2011; Parrell et al., 2019; Tourville et al., 2008). In some models of speech motor control, such as Directions Into Velocities of Articulators (Guenther, 2016; Tourville & Guenther, 2011), these feedback gains are also thought to drive changes in feed-forward control over time. A number of factors have been proposed to account for variation of these gains across individuals, including auditory and somatosensory acuity (Franken et al., 2017; Guenther, 2016; Parrell et al., 2019; Tourville et al., 2008; Villacorta et al., 2007), variation in the balance between auditory and somatosensory systems (Guenther, 2016; Katseff et al., 2012; Lametti et al., 2012; Parrell et al., 2019), production

^aDepartment of Communication Sciences and Disorders, University of Wisconsin–Madison

Correspondence to Caroline A. Niziolek: cniziolek@wisc.edu

Editor-in-Chief: Cara E. Stepp

Editor: Jason W. Bohland

Caroline A. Niziolek and Benjamin Parrell equally contributed to this work.

Received May 18, 2020

Revision received September 2, 2020

Accepted December 8, 2020

https://doi.org/10.1044/2020_JSLHR-20-00263

Publisher Note: This article is part of the Special Issue: Selected Papers From the 2020 Conference on Motor Speech—Basic Science and Clinical Innovation.

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

variability (Munhall et al., 2019), and the presence of neurological disorders such as stuttering (Cai et al., 2012), Parkinson's disease (Mollaei et al., 2016), and cerebellar ataxia (Parrell et al., 2017). These factors are all generally assumed to be relatively consistent over time, suggesting feedback gains may be a stable characteristic of the speech motor control system. This assumption is often implicit, but is reflected in attempts to establish predictors of variation in feedback gains by correlating these factors with behavioral responses to auditory perturbations (Feng et al., 2011; Martin et al., 2018; Villacorta et al., 2007).

However, little is known about the stability of the compensatory response, or the feedback gain it is thought to reflect, across time. Critically, evidence from nonspeech motor control suggests feedback gains may be more malleable than is often assumed in speech research. For example, when participants are given altered visual feedback about their hand position during a reach, the magnitude of their compensation for this perturbation is modulated by the reliability of the visual feedback. That is, compensation is largest when the feedback is most reliable—a single dot representing their hand—and smaller when it is less reliable—a cloud of dots centered at the position of their hand (Körding & Wolpert, 2004). These results show that the online compensation for sensory errors is not a fixed characteristic of the sensorimotor control system, but rather takes into account the reliability of the sensory signal, even when that reliability varies across repetitions of a movement. This behavior is not predicted by models that consider feedback gains to be tied only to individual characteristics such as sensory acuity, which are thought to be relatively stable (Bischof et al., 2002; Saito et al., 2020), but is consistent with an optimal or Bayesian model of sensory processing (Körding & Wolpert, 2006; Wei & Körding, 2009). Moreover, there is evidence from reaching tasks that people monitor their history of sensory errors over time to generate estimates of the reliability of sensory signals (Herzfeld et al., 2014). Thus, it may be possible that long-term exposure to perturbed sensory feedback could lead to a change in the perceived reliability of that feedback, which would be expected to affect measured compensation.

In speech, evidence for similar effects of sensory reliability is limited. However, there is some evidence that the control of vocal pitch is consistent with an optimal sensory system, which has higher gains for more reliable signals. When somatosensory feedback is partially blocked by the application of topical anesthetic to the vocal folds, compensation for auditory pitch perturbations increases, as the auditory signal is now relatively more reliable (Larson et al., 2008). Additionally, small pitch perturbations (100 cents) can elicit larger compensatory responses than large pitch perturbations (> 300 cents), evidence that the large perturbations may be discounted by the sensorimotor control system as externally generated (Korzyukov et al., 2017; Scheerer et al., 2013). Separately, pitch perturbations that are unpredictable in magnitude or direction lead to larger compensatory responses than predictable perturbations (Korzyukov et al., 2012; Scheerer & Jones, 2014); that is,

the nature of previously encountered auditory errors can affect the magnitude of compensation. This suggests the vocal sensorimotor system monitors a history of sensory errors, as in reaching. However, to our knowledge, no work to date has examined the effect of sensory reliability on feedback gains in the supralaryngeal motor control system.

Here, we examine how the reliability of auditory feedback signals affects feedback control during speech production. Answering this question would ideally employ a direct manipulation of feedback reliability, analogous to varying the visual representation of hand position between a single dot versus a cloud of dots (Körding & Wolpert, 2004). However, it is not clear how such a direct modulation of reliability could be implemented in the auditory system. Thus, we take an alternative approach based on the idea that the reliability of sensory inputs is estimated through a remembered history of errors (Herzfeld et al., 2014). Specifically, we test how repeated exposure to auditory errors—that is, added feedback noise in the form of small alterations to formant frequency—affects the magnitude of compensation for subsequent auditory feedback perturbations. We examine two competing hypotheses. One possibility, consistent with work in reaching, is that the long-term exposure to auditory errors created by feedback perturbations would be attributed to noise or errors in the auditory system, leading to a *decrease* in the estimated reliability of auditory feedback, a smaller weighting of auditory errors, and a decreased compensatory response to auditory perturbations. Alternatively, these errors could be attributed to motor noise, which causes deviations in speech output that are faithfully reflected by the sensory system. Given an “optimal” sensorimotor system, this would be expected to lead to an *increase* in the weighting of sensory reafferent signals, as this “reliable” sensory feedback would be informative for correcting errors in output.

In separate sessions, participants were consistently exposed to small, random perturbations of their vowel formants (*noisy* session) or received veridical auditory feedback (*veridical* session). We subsequently measured the magnitude of compensation for intermittent, unpredictable auditory perturbations. This design allows us to test for changes in compensation as auditory reliability decreases. Our results provide evidence that compensation magnitude is reduced after exposure to unreliable feedback, consistent with a downweighting of auditory feedback caused by a decrease in its reliability. Though the size of this effect is small, this is consistent with optimal accounts of sensorimotor control that predict flexible feedback gains.

While not the primary purpose of this study, we additionally assessed whether compensation for altered feedback perturbations could be predicted by variability in production during normal, unaltered speech or by vowel centering, the reduction in variability that occurs from vowel onset to the vowel midpoint that may be driven in part by feedback-based corrections for self-produced variability (Niziolek & Kiran, 2018; Niziolek et al., 2013). The multisession design additionally allows us to examine the cross-session stability of both compensation and these other speech behaviors over time.

Method

Participants

Forty individuals participated in the current study (35 female, $M_{\text{age}} \pm SD = 21.6 \pm 4.6$ years). No participant reported any history of speech, hearing, or neurological disorders. All participants were native speakers of American English. Participants were compensated for their participation either monetarily or through extra credit in a course in the UW–Madison Communication Sciences & Disorders department. All procedures were approved by the institutional review board at UW–Madison.

Stimuli and Trial Structure

On each trial, participants produced a single, monosyllabic English target word that appeared on a computer monitor. The three target words were “bed,” “head,” and “dead.” The order of the words was pseudorandomized. The stimulus word appeared for 1.5 s, and time between trials was jittered between 0.75 and 1.5 s. Self-timed breaks were given every 30 trials throughout the experiment.

Auditory Recording and Perturbation

Participants performed the experiment in a sound-attenuated room. Participants’ speech was recorded with a head-mounted microphone (AKG C520) and played back to them over closed-back, over-the-ear headphones (Beyerdynamic DT 770) with a total latency of ~ 18 ms as measured in our labs (Kim et al., 2020). Speech was recorded at 16 kHz and digitized with a Focusrite Scarlett sound card. All speech recording and playback were done through Audapter (Cai et al., 2008; Tourville et al., 2013). On some trials, participants’ speech was altered by shifting the first and/or second vowel formants in Audapter throughout the trial. On other trials, participants received unaltered speech feedback. Speech was processed through the same pipeline and had the same feedback latency regardless of whether or not a shift was applied. Participants’ speech was played back at ~ 80 dB SPL based on a pre-experiment calibration. The actual amplitude of speech playback was dependent on the amplitude of participants’ production. Speech playback was combined with speech-shaped masking noise at ~ 60 dB SPL to mask any air- or bone-conducted auditory feedback.

Experimental Procedure

All participants completed two sessions, *veridical* and *noisy* (see Figure 1), roughly 1 week apart (mean: 9 days, range: 4–25 days). Our intention was to counterbalance the order of sessions across participants. However, due to an error in the randomization procedure, all participants completed the noisy session first.

The structure of the two sessions was the same (see Figure 1A). Each session began with a 450-trial exposure phase. In the veridical session, participants received unper-
turbed auditory feedback throughout this phase. In the

noisy session, participants received a random perturbation of the first two vowel formants (F1/F2) on each exposure trial. The direction of this perturbation was a random direction in F1/F2 space (see Figure 1B). The perturbation magnitudes were drawn from a normal distribution with a standard deviation of 10 mels, with the caveat that the 90 trials with the smallest perturbations were set to 0. That is, these 90 trials received veridical auditory feedback (see Figure 1C). These unaltered trials were included to allow measures of baseline variability and centering (see below) that could be compared across sessions.

Following the exposure phase, participants completed a 240-trial test phase. In the test phase of each session, F1 was increased by 125 mels on 1/6 of the trials (*upshift condition*: 40 trials), decreased by 125 mels on 1/6 of the trials (*downshift condition*: 40 trials), and unaltered on the remaining 2/3 of the trials (160 trials). The order of the shifts was pseudorandomized across trials, with the restriction that each shifted trial was followed by at least one unshifted trial (i.e., shifts could not occur on two consecutive trials). The perturbations were not counterbalanced with regard to stimulus word; there were no requirements that each word have the same number of shifts.

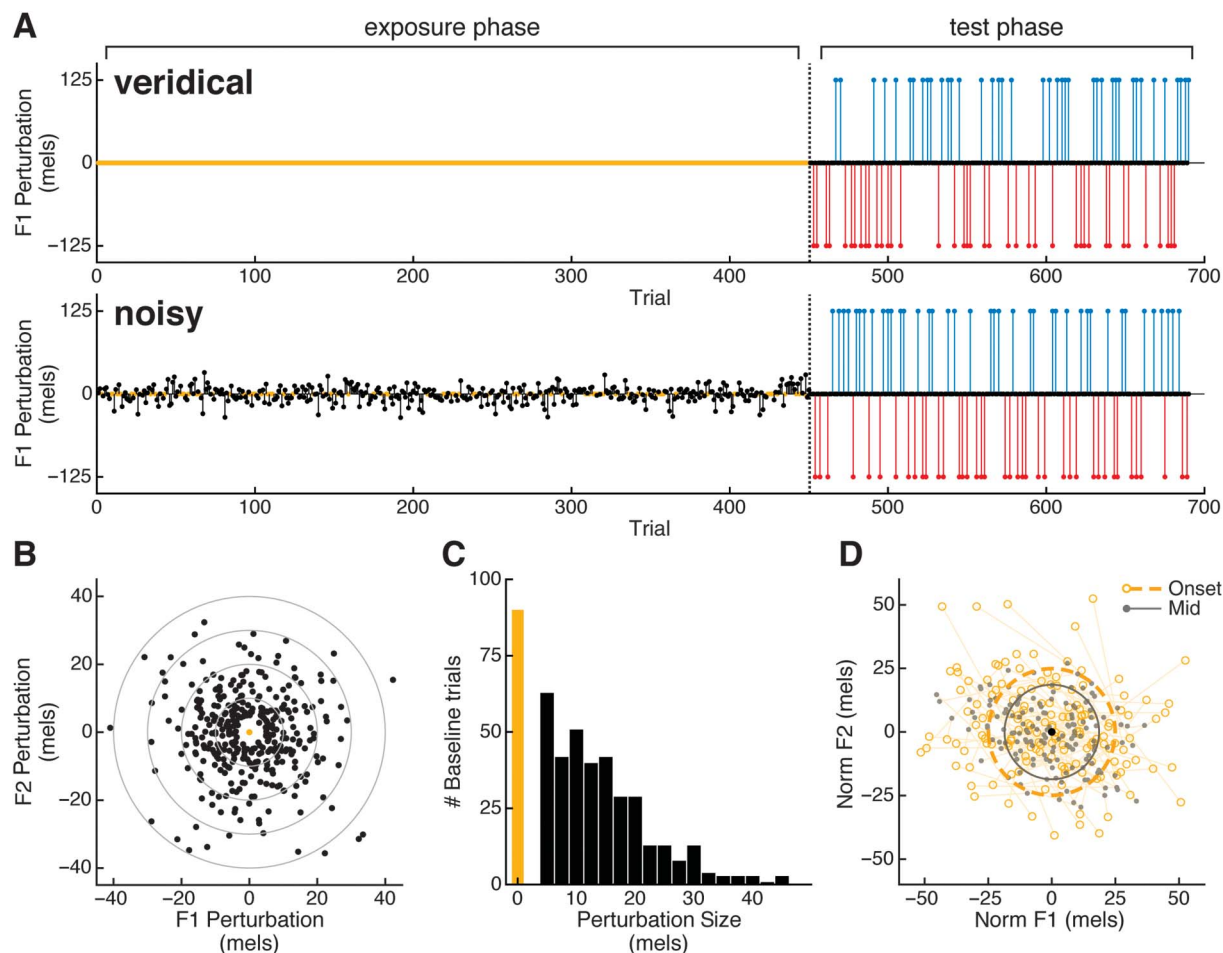
Duration Control

The latency of the compensatory response to unexpected formant perturbations is around 150 ms (Cai et al., 2012; Parrell et al., 2017; Tourville et al., 2008). To ensure there was sufficient time to observe compensatory responses, participants were trained to produce vowel durations between 250 and 500 ms. Participants were instructed to speak slightly slower than normal and were given feedback about their vowel durations after each trial. Vowel durations were estimated as the duration of the speech signal with a root-mean-square amplitude above a minimal threshold. After each trial, a visual cue appeared for 500 ms below the target word indicating whether the vowel duration fell in the target range (a green circle), was too long (a yellow circle and text reading “Speak a little faster”), or too short (a blue circle and text reading “Speak a little more slowly”). Prior to the experiment, participants completed a 10-trial training session to familiarize themselves with the visual cues and practice producing vowels with durations in the target range. Participants repeated the test until at least 8/10 trials were in the target range. The visual feedback about vowel duration was provided throughout the experiment.

Data Processing

Vowel formants were tracked with *wave_viewer* (Niziolek & Houde, 2015), an in-house software tool that provides a MATLAB GUI interface to Praat (Boersma & Weenink, 2019). Pre-emphasis values and linear predictive coding (LPC) order (the number of coefficients in the LPC model) were set for each participant individually. Vowel onset and offset were first automatically identified based on a participant-specific amplitude threshold, and formant values in

Figure 1. Experiment design. A: Schematic of perturbations applied in the veridical feedback (upper panel) and noisy feedback (lower panel) sessions. Yellow denotes exposure trials in which feedback was unaltered. B: Distribution of formant perturbations in the noisy feedback session. C: Histogram of perturbation sizes in the noisy feedback session. D: Schematic of centering analysis of unaltered trials, in which trial-wise distances to the median formants for a given word are compared at vowel onset (open yellow circles) and vowel midpoint (gray dots).



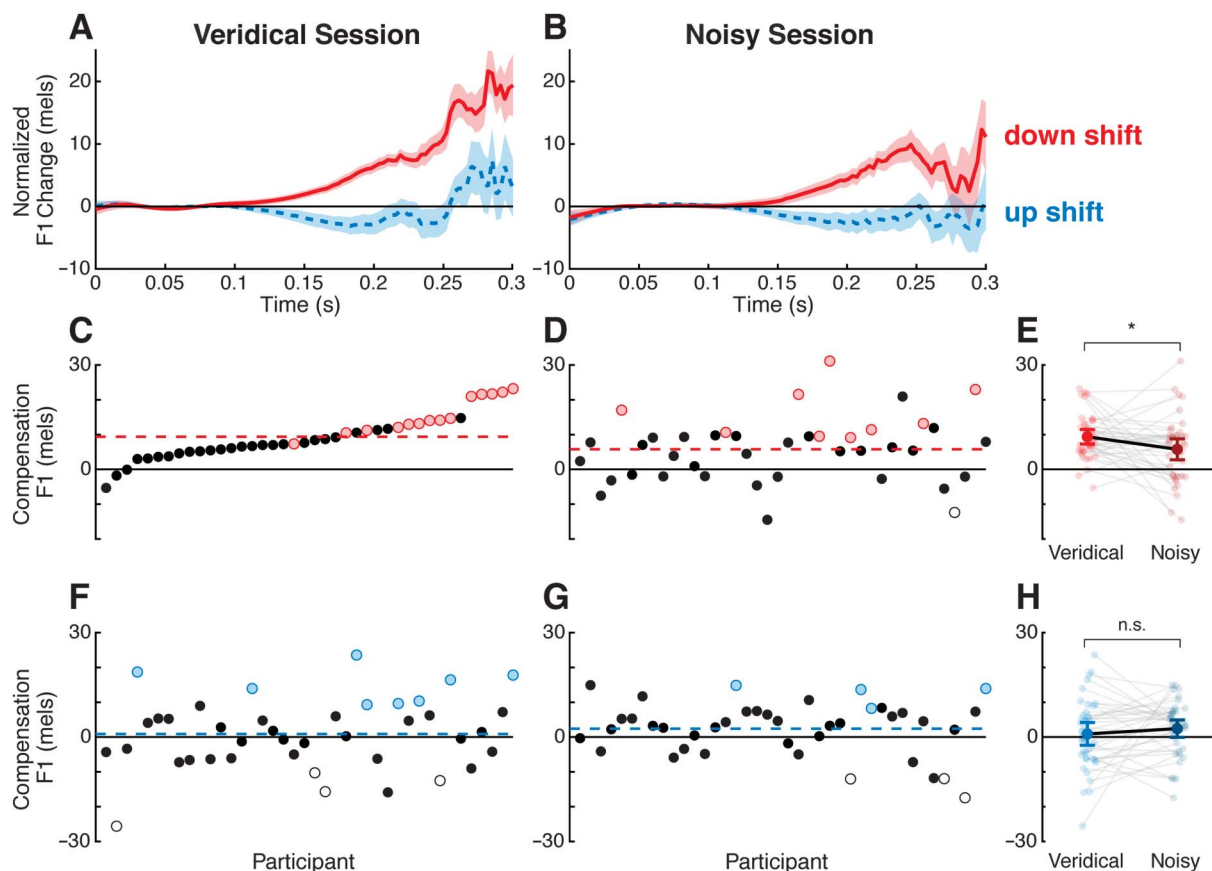
this window were tracked using the participant-specific parameters. All trials were subsequently checked for errors. Errors in vowel onset and offset were corrected by hand-labeling these times using the waveform and spectrogram. Vowel onset was identified as the point at which periodicity was visible in the waveform and formants were visible in the spectrogram. Vowel offset was identified as the point where formants, particularly F1 and F2, were no longer visible. Errors in formant tracking (such as misidentifying fundamental frequency as the first formant) were corrected by adjusting the pre-emphasis value or, if that was unsuccessful, LPC order. A limited number of trials ($M = 2.3\%$, range: 0%–19.8%) were excluded due to errors, such as the participant saying the wrong word, disfluencies, or unresolvable errors in formant tracks.

Behavioral Measurements

Our primary dependent measure was behavioral *compensation* for the feedback alterations in the test phase. To

calculate compensation for a given speaker, we first calculated the mean F1 trajectory across all unperturbed trials for each stimulus word. We then subtracted these word-specific mean trajectories from each perturbed trial, yielding F1 difference trajectories that reflected change from unperturbed trials (see Figure 2A, B). To account for any initial variability in F1 unrelated to the perturbation, each of these trajectories was normalized to its average value from 25 to 100 ms (excluding the first 25 ms of formant transitions). Such early variability is unrelated to the compensatory response, which does not begin until ~150 ms after vowel onset (Cai et al., 2012; Parrell et al., 2017; Tourville et al., 2008). Thus, this measure isolates within-trial changes in behavior in response to the auditory perturbation. Subsequently, an average difference trajectory was calculated for each shift direction by taking the mean response at each time point across all shifted trials. The magnitude of the compensatory response, for each participant, was calculated as the mean F1 value in these average difference trajectories from 150 to 300 ms

Figure 2. Compensation to F1 shifts. A, B: Change in F1 in the test phase of the veridical (A) and noisy (B) sessions. C, D, F, G: Per-participant compensation magnitude to downward (C, D) and upward (F, G) perturbations, measured as F1 change opposing the perturbation. Participants are sorted by the magnitude of their compensation to the downward perturbation in the veridical session (C); all other panels use the same participant sort order. Colored dots represent significant compensatory responses (red: downward perturbation; blue: upward perturbation); white dots represent significant following responses. Colored dashed lines show average compensation across participants. E, H: Comparison of veridical and noisy session compensation to downward (E) and upward (H) perturbations. Each pair of connected points represents a single participant.



after vowel onset. The sign of the F1 difference in response to upward perturbations was flipped such that compensation was always reflected by positive values.

During data analysis, we realized that seemingly minor changes in how compensation was measured could have surprisingly large effects on the final outcome measure. As the number of studies on online compensation for formant perturbations is relatively small, and best practices for data analysis are not well established, we additionally present the results of several unplanned analyses, including (a) evaluating compensation in a smaller time window, from 200 to 300 ms after vowel onset, and (b) calculating the mean response in the analysis window for every trial, then taking the mean across trials within each participant, rather than first calculating a single average response for each shift direction. We additionally tested how normalizing each trial to its early (25–100 ms) preresponse values results affects our results. This nonnormalized method allows easier comparison with previous work (e.g., Cai et al., 2012; Niziolek & Guenther, 2013;

Parrell et al., 2017; Purcell & Munhall, 2006; Reilly & Dougherty, 2013; Tourville et al., 2008).

In order to test what factors may be related to inter-participant variability in compensatory behavior, we additionally included two other measures of vocal behavior taken from unperturbed trials in the exposure phase (450 trials in the veridical session and 90 trials in the noisy session): (a) variability in F1/F2 space at vowel onset and (b) *vowel centering*, defined as the reduction of variability from vowel onset to vowel midpoint (Niziolek & Kiran, 2018; Niziolek et al., 2013). Variability was defined as the average Euclidean distance of all productions of a stimulus word to the median of that distribution in F1/F2 space. Variability was measured at both vowel onset and vowel midpoint (see Figure 1D), and centering was defined as the change in variability from onset to midpoint, such that positive centering values reflect a decrease in variability between these time windows; negative values represent an increase in variability. Variability and centering were calculated

for each stimulus word separately to control for coarticulatory effects, and averages of these three measures in each session were then calculated for each participant. Only unperturbed trials were used to avoid any potential influence of altered auditory feedback on these measures.

Statistical Analysis

At the group level, compensation and centering were analyzed using linear mixed-effects models with fixed effects of session (veridical or noisy) and perturbation direction (up or down). The model included random intercepts and slopes for each fixed factor by participant. Word was not included as a factor as there were only 10–15 trials per word in each perturbation direction, not sufficient to see a reliable response. All models were constructed with the *lme4* package (Bates et al., 2015) in R (R Core Team, 2013). Statistical significance was assessed with the *lmerTest* package (Kuznetsova et al., 2017). Post hoc comparisons were conducted with the *emmeans* package (Lenth et al., 2020). The relationships between compensation and baseline variability/centering as well as the consistency of compensation, variability, and centering across sessions and the relationship between responses to upward and downward perturbations within a session were assessed with Pearson correlations.

At the individual level, compensation was considered significant if a participant's compensation magnitude exceeded a threshold found via a bootstrapping procedure. For a given participant, each perturbation direction was assessed separately by first randomly shuffling the labels of perturbed and unperturbed (baseline) trials and then computing compensation in the same manner as described above (*behavioral measurements*). This procedure was repeated 1,000 times, and the compensation threshold was defined as the 95th percentile of this null distribution. "Following" responses were considered significant if the compensation was less (more negative) than the 5th percentile of the null distribution.

Results

Compensatory responses to altered feedback trials from the test phase are shown in Figure 2 (A and B). In both sessions, trials in which participants experienced a downward perturbation to F1 tended to deviate upward from the baseline, opposing the perturbation. Trials with upward perturbations showed the same opposing trend, tending to deviate downward from the baseline, but with a much smaller response magnitude. In both sessions, the overall magnitude of the compensatory response ($M_s \pm SE$: down = 7.55 ± 1.30 mels; up = 1.66 ± 1.43 mels) was substantially smaller than in previously reported work (e.g., Cai et al., 2012; Niziolek et al., 2013; Parrell et al., 2017; Purcell & Munhall, 2006).

Overall, our method was somewhat successful in causing participants to produce vowels with durations in the target range. Average produced vowel duration was 280 ± 44 ms (standard deviation). There was no difference in duration

between the sessions, $F(1, 39) = 0.66, p = .42$. A percentage of trials (75.4%; 14,108 trials) had a duration over 250 ms, and 27.7% (5,180 trials) had a duration over 300 ms. Individually, 34 participants had a mean vowel duration over 250 ms and 10 participants had a mean duration over 300 ms. Thus, later time points reflect the contribution of fewer trials.

In the planned analysis window (150–300 ms), there was a main effect of perturbation direction, $F(1, 39) = 20.9, p < .0001$, reflecting the larger compensatory response to downward perturbations compared with upward perturbations. There was no main effect of session, $F(1, 39) = 0.7, p = .42$, but there was an interaction between session and perturbation direction, $F(1, 39) = 4.7, p = .037$. This interaction reflected a smaller compensatory response to the downward perturbation following the noisy exposure phase (5.8 ± 1.5 mels) compared to following the veridical exposure phase (9.4 ± 1.0 mels; $p = .042$; see Figure 2E). There was no significant difference in the response to the upward perturbation between noisy (2.4 ± 1.2 mels) and veridical (0.9 ± 1.6 mels) sessions ($p = .48$).

To examine changes in compensation over the course of a single session, we compared the magnitude of compensation in the first half versus the second half of each test phase, averaging across upward and downward perturbations. Compensation did not differ between the two halves of the veridical test phase (4.9 ± 1.6 vs. 4.6 ± 1.1 mels, $t(39) = 0.2, p = .83$). However, compensatory responses in the first half of the noisy test phase were significantly smaller than in the second half (2.3 ± 1.5 vs. 5.7 ± 1.1 mels, $t(39) = -2.0, p = .049$). While the effect is small, and conclusions should be tempered by the smaller number of trials in this analysis, this pattern is consistent with a reduced sensitivity to errors subsequent to the noisy exposure phase, which recovers once this noise is removed.

The bootstrapping analysis identified participants whose compensatory responses across the entire test phase significantly differed from chance. In the veridical session, 14 participants exceeded these individual compensation thresholds for the downward perturbation (see Figure 2C, red dots); eight participants exceeded these thresholds for the upward perturbation (see Figure 2F, blue dots). Participants with significant following responses are shown in white dots. In the noisy session, nine participants showed significant compensation for the downward perturbation (see Figure 2D, red dots), while four participants compensated for the upward perturbation (see Figure 2G, blue dots).

Relationship Between Compensation and Other Production Measures

In addition to our primary outcome metric of compensation for formant perturbations, we examined the baseline variability observed across multiple repetitions of a word, as well as vowel centering, the reduction in this variability from vowel onset to vowel midpoint. As described in the methods, we analyzed variability and centering in only the unperturbed trials from the exposure phase. Variability at vowel onset did not differ between the veridical (31.6 ± 6.4 mels)

and noisy (32.0 ± 5.3 mels) sessions, $t(39) = 0.4, p = .69$. Likewise, centering was not significantly different between the two sessions, $t(39) = 1.7, p = .09$, although it was significantly greater than 0 in both the veridical ($4.0 \pm .9$ mels, $t(39) = 4.7, p < .0001$) and noisy ($2.4 \pm .8$ mels, $t(39) = 2.9, p = .006$) sessions. These results show that, overall, participants do reduce their vowel formant variability from vowel onset to vowel midpoint and that the magnitude of this reduction is similar across sessions (see Figure 3).

Centering and compensation have both been hypothesized to be driven by error-corrective feedback control processes. Under this hypothesis, we would expect a significant positive correlation between the two phenomena. However, we found no evidence for this relationship (see Figure 4). In the noisy session, centering was not correlated with compensation to downward perturbations ($r = -.03, p = .83$), upward perturbations ($r = -.14, p = .40$), or to an average measure of compensation combining responses to both upward and downward F1 perturbations ($r = -.11, p = .52$). In the veridical session, centering was not correlated with the upward compensation measure ($r = -.18, p = .27$) or with average compensation ($r = -.29, p = .07$). However, we observed an unexpected negative correlation between centering and compensation to downward F1 perturbations ($r = -.32, p = .046$) in the veridical session. Taken together, these results suggest there is likely no consistent relationship between the magnitudes of centering and compensation in our study. Separately, neither initial nor midpoint variability was correlated with any measure of compensation in either session (see Appendix). The correlation between initial variability and centering was not significant in either the veridical ($r = .29, p = .07$) or noisy session ($r = .08, p = .62$), although a relationship between these measures has been reported in past work (Niziolek & Kiran, 2018).

Unplanned Analyses

We conducted the same analyses described above with three different methods of estimating individual values for compensation. First, we reduced the analysis window from 150–300 ms to 200–300 ms after vowel onset. Second, we estimated compensation by calculating an average compensation magnitude in the analysis window for each perturbed trial, then averaging across trials. This method allowed for by-trial estimates of compensation; however, because not all trials had vowel durations as long as 300 ms, these estimates differentially weighted earlier time points more heavily in the cross-trial average. This is in contrast to our planned analysis, which calculated an average response across all perturbed trials at each time point. We conducted this by-trial analysis for both the 150- to 300- and 200- to 300-ms windows. Last, we calculated compensation without normalizing each trial to its prerespone baseline (25–100 ms). While this method is more similar to some previous studies, it is less sensitive to within-trial change. Here, we will summarize differences between these analyses. A full table of results can be found in the Appendix.

Results in the 200- to 300-ms analysis window largely replicated our findings from the 150- to 300-ms analysis

window. We found a larger compensatory response to downward compared to upward perturbations, and an interaction between direction and session, which again reflected a smaller compensatory response to the downward perturbations following the noisy exposure phase compared to the veridical exposure phase. For the by-trial analyses, the same difference between upward and downward shifts was found in the 200- to 300-ms window only. Neither analysis window for the by-trial analysis showed evidence for an interaction between session and direction. Nonnormalized compensation values showed a difference between the upward and downward shifts in both windows. Although the between-session difference in responses to the downward shift was in the same direction as in our planned analysis, this did not reach our threshold for significance in either window. This is likely due to a small offset in the prerespone (< 100 ms) window that slightly reduced the magnitude of the difference between sessions (see Supplemental Material S1, Figure S1).

Both the average response and by-trial analyses in the 150- to 300-ms window showed a significant correlation between centering and compensation in the veridical session only. This relationship was not observed in the 200- to 300-ms window for either method, and no relationship was found between centering and compensation using the nonnormalized method. No analysis method showed any significant correlation between compensation and centering in the noisy session, nor between centering and initial variability.

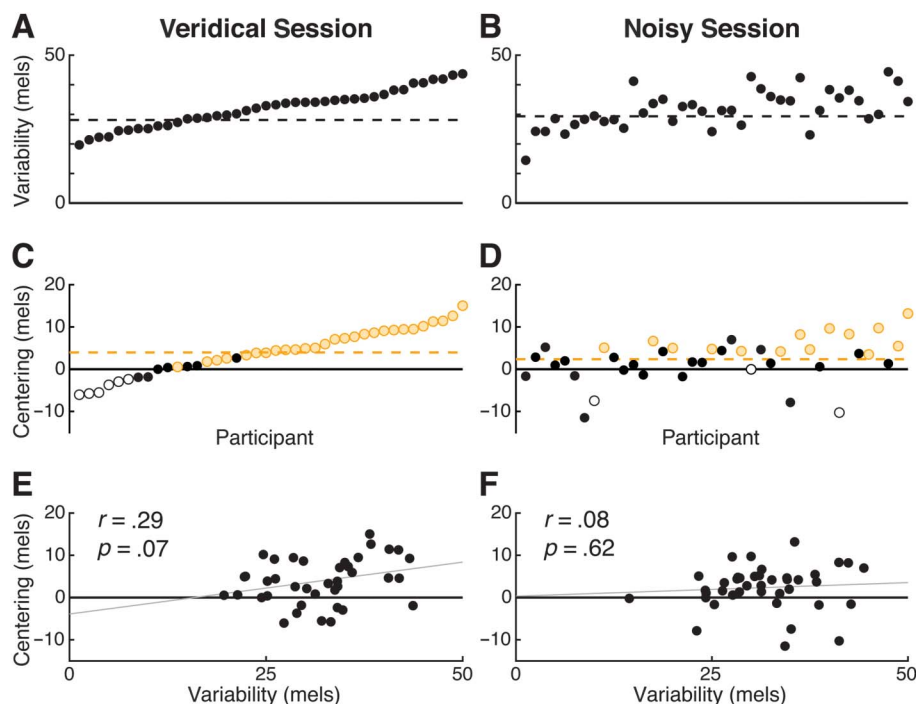
The variability observed in these results shows how seemingly minor decisions in data analysis can affect the conclusions drawn from studies on compensation for altered auditory feedback in speech. Looking across analyses, we find mixed support for our principal hypothesis that compensation magnitude would be affected by long-term exposure to altered auditory feedback. This was supported by using an average response per participant in both time windows, though not by averaging across trials nor using nonnormalized measures.

Finally, while not the original focus of the study, we additionally compared compensation magnitudes across the two sessions, veridical and noisy, as well as the relationship between compensation to upward and downward responses within each session, and found limited evidence for a significant relationship between these measures (see Supplemental Material S1, Figures S2 and S3). Formant variability and centering were more consistent across sessions, showing significant correlations between the veridical and noisy exposure phases (see Supplemental Material S1, Figures S3D and S3E).

Discussion

In this study, we examined how individuals' compensation for unexpected formant perturbations was modulated by exposure to unreliable auditory feedback. Our results suggest tentative support for our hypothesis that unreliable feedback affects sensorimotor gains. Exposure to small, quasirandom formant perturbations ($SD = 10$ mels) was

Figure 3. Variability and centering across sessions. Data are drawn from unperturbed trials in the exposure phase. A, B: Initial variability is similar between veridical (A) and noisy sessions (B). Participants are sorted by their variability in the veridical session (A), and the same order is used in (A) and (B). C, D: Centering also shows a consistent magnitude across veridical (C) and noisy sessions (D). Individuals who showed a significant change in variability are shown in yellow (reduction in variability) or white (increase in variability). Participants are sorted by the magnitude of centering in the veridical session (C), and the same order is used in (C) and (D). E, F: Initial variability and centering are not significantly correlated in either the veridical (E) or noisy sessions (F).



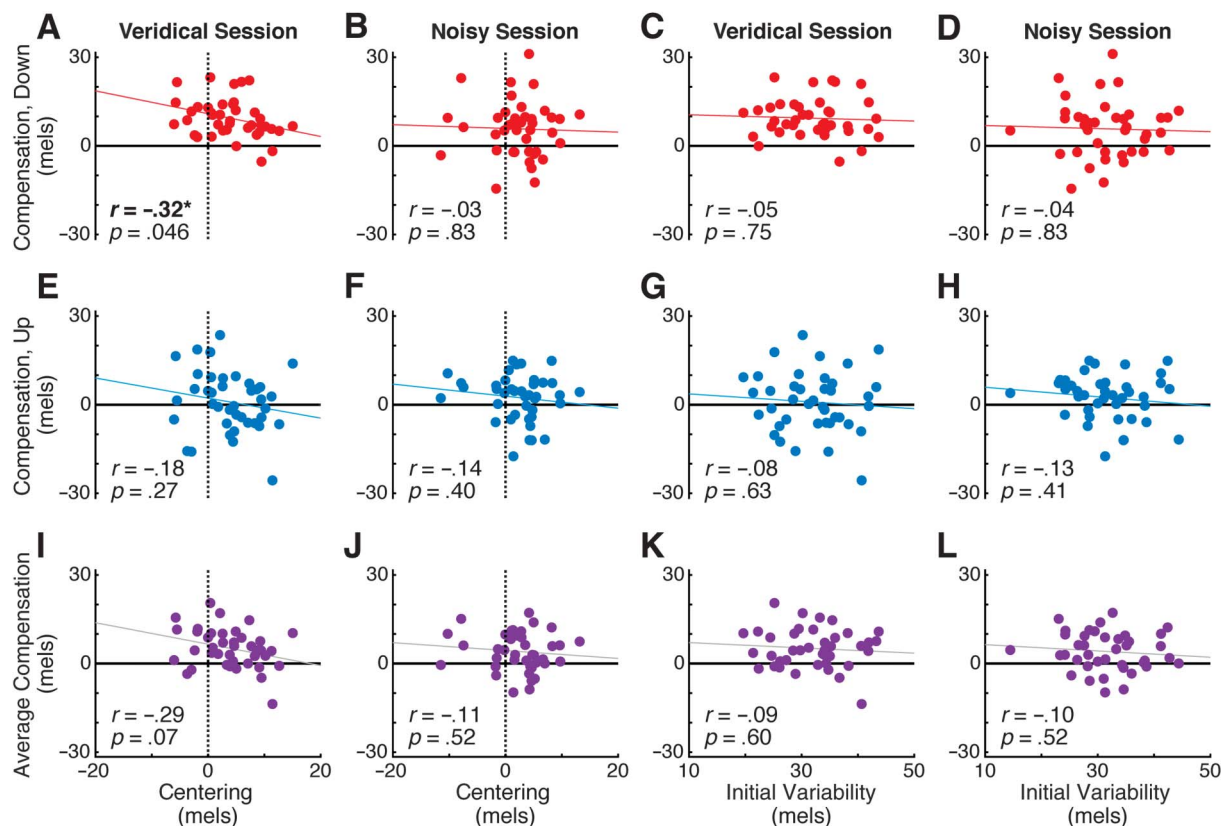
associated with a small but significant decrease in the magnitude of compensation for large downward F1 perturbations (125 mels) relative to a control session. This decrease in compensation is consistent with domain-general accounts of optimal sensory integration during sensorimotor control, which predict a downweighting of sensory feedback gains when that feedback is unreliable. Furthermore, compensation magnitude was smallest immediately after the noisy exposure phase, compatible with a decrease in error sensitivity that partially recovered over the course of the test phase.

Although compensation is known to vary widely across individuals (Cai et al., 2012; Parrell et al., 2017), many suggested predictors of compensation magnitude are taken to reflect stable individual differences. Our results suggest that the magnitude of compensation in any given instance may be less strongly influenced by stable features of the sensorimotor system, such as sensory acuity or production variability, than by more pliable aspects, such as the history of experienced sensory errors or attentional state. While converging evidence from multiple speech studies has linked auditory acuity and responses to auditory feedback errors (Martin et al., 2018; Villacorta et al., 2007), these studies were largely based on sensorimotor adaptation, not online compensation, and other similar studies have failed

to find a correlation between these measures (Feng et al., 2011). The current finding, that compensation can be modulated by externally manipulating sensory reliability, is consistent with the idea that the history of sensory errors drives changes in error sensitivity that, along with acuity, contribute to compensation differences. A similar role for error history has been invoked to explain various aspects of sensorimotor learning in reaching movements (Herzfeld et al., 2014). There is also evidence from pitch perturbation studies that previous exposure to larger perturbations may attenuate later compensatory responses (Scheerer & Jones, 2014). Separately, attention has also been shown to modulate the response of the sensorimotor system to perceived errors, though typically attention has been shown to modify trial-to-trial learning, rather than feedback control specifically (Taylor & Thoroughman, 2007). There is limited evidence that attentional load may also modulate compensatory responses to vocal pitch perturbations (Tumber et al., 2014), though these results are not consistent (Hu et al., 2015; Y. Liu et al., 2018).

Complementarily, although measures of baseline formant variability and centering were relatively consistent across sessions, reflecting stable individual differences in production, we found little evidence that these measures significantly predicted individual compensation. These results are

Figure 4. Relationship between compensation and measures of production in the exposure phase. A–D: Relationships between compensation to downward perturbations and centering (A and B) and initial variability (C and D). Data from the veridical session shown in (A) and (C); data from the noisy session shown in (B) and (D). E–H: As in (A–D), showing compensation to upward perturbations. I–L: As in (A–D), showing compensation averaged across upward and downward perturbations. Results are similar whether using initial or midpoint variability.



somewhat surprising given the hypothesized relationship between the size of a speaker's "acceptable" formant range for a given vowel, for which variability is a proxy, and the size of the corrective movement needed to keep productions within that range (Tourville & Guenther, 2011). Furthermore, both compensation and centering have been hypothesized to be driven by feedback-based corrections for auditory errors, either externally imposed (in the case of compensation) or endogenous (in the case of centering; Niziolek et al., 2013). The lack of any consistent relationship between centering and compensation in our results suggests that these are not perfectly parallel processes, and may be differently affected by exposure to auditory perturbations, which cause an auditory-somatosensory mismatch not experienced in natural speech.

However, there are a number of reasons why these results should be interpreted cautiously. First, the compensation measure itself is likely quite noisy. To date, there has been little work to establish best practices in formant perturbation experiments, and it is unknown how many trials are needed to reliably estimate the magnitude of compensation either within or across individuals. Strikingly, the standard deviation of responses within individuals (16.8–62.2 mels, when compensation is calculated separately for each trial)

is larger than the standard deviation of mean responses across individuals (6.5 mels). This suggests that it may take a very large number of trials to generate a reliable estimate for any one individual. Consistent with this idea, we found only a moderate correlation ($r = .47$, $p = .002$) between responses in the first and second half of the test phase in the veridical condition (see Supplemental Material S1). In the current study, we measured 40 responses to both upward and downward perturbations, similar to the number of trials in previous studies (Cai et al., 2012; Mollaei et al., 2016; Niziolek & Guenther, 2013; Parrell et al., 2017; Reilly & Dougherty, 2013; Tourville et al., 2008). While this number of trials may reliably elicit compensatory responses at the group level, it seems likely that this does not provide a large enough sample to give a robust estimate of compensation for a given individual, given the large intra-individual variance. Future studies should examine whether increasing the number of trials leads to more stable measurements of compensation, both within and across sessions.

Second, our results highlight the variability in outcomes caused by seemingly minor decisions in data analysis. It is not straightforward what the optimal methods are to quantify compensation, which is a highly variable response both

within and across individuals. While we have attempted to survey some of the possibilities in this article, including methods for calculating individual responses and the effects of different time windows, this should not be taken as an exhaustive assessment. In particular, our analysis window ended at 300 ms, due to the relatively short productions elicited in our paradigm. It is possible that longer analysis windows may provide measurements that are more stable to changes in analysis methods. However, the advantage of the current paradigm is that vowel durations are more similar to those found in natural speech. It is not clear how extending the duration of vowels beyond those normally produced as speech may affect the compensatory response, nor whether behavior observed in such a paradigm using extended vowel production reflects the behavior of the sensorimotor control system as it typically operates.

Limitations

Overall, we have provided evidence that compensation for external perturbations of vowel formants is reduced by exposure to unreliable auditory feedback. However, the strength of these findings is somewhat tempered by a few limiting factors. First, while we did observe a robust response to downward perturbations of F1, we observed no group-level response to upward perturbations. Subsequently, the hypothesized reduction in compensation was found only for the downward F1 perturbations. This asymmetry was unexpected given that most previous studies have shown relatively symmetric effects, although at least one study has shown different latencies, though not magnitudes, for responses to upward and downward shifts of F1 for /*ɛ*/ (Tourville et al., 2008). Some other studies have shown asymmetric responses to perturbations of F2 (Cai et al., 2011) or vocal pitch (Burnett et al., 1997; H. Liu & Larson, 2007; Sares et al., 2018). Another study has shown an asymmetric response to F1 perturbations when compensation was measured immediately following another formant perturbation experiment, though in this case, the response to downward perturbations was attenuated (Parrell et al., 2017). Thus, the asymmetry may be caused by previous exposure to auditory perturbations, though why this should be the case is not clear. Separately, the asymmetry may be due to differences in categorical vowel boundaries. Previous work has shown that compensatory responses are larger when they cause formant feedback to cross a perceptual boundary between vowels (Niziolek & Guenther, 2013). Our participants were all recruited from the Madison, Wisconsin area, and thus may produce a somewhat “raised” /*æ*/, with average F1 values similar to or even lower than those for /*ɛ*/ (Hillenbrand et al., 1995). This would create the situation where a lowered F1 would result in a categorical shift from /*ɛ*/ to /*ɪ*/, but a raised F1 may not cross into any other vowel category. Relatedly, the overall magnitude of compensation in the current study was relatively small compared to previous studies. It is possible this is related to the large number of exposure trials in the current study, which may have caused some fatigue. While we found no difference in the magnitude

of compensation between the first and second halves of the veridical test phase, participants had already produced 450 spoken trials in the exposure phase, leaving open the possibility that their response had plateaued before the beginning of the test phase. Future work should explore how compensation may change over the course of many trials.

A second limitation is the relatively short vowels produced in the current study. While we only required participants to produce vowels somewhere between 250 and 500 ms in duration, most participants produced vowels at the shorter end of that range. Thus, many vowel productions were shorter than 300 ms. This may have led to less accurate estimates of compensation in the later range of our analysis window (150–300 ms), as fewer trials contributed to the average. Separately, vowel durations shorter than 300 ms mean that we were measuring not only the steady-state portion of some vowels, but also including some of the formant transitions into the word-final /*d*/. Since F1 lowers from /*ɛ*/ to /*d*/, this may have caused a bias in the compensation values, which may also have potentially contributed to the asymmetry seen in the responses. Future work would benefit from increasing the minimum acceptable duration when examining compensation for auditory perturbations. However, this needs to be balanced with speech naturalness, as it is not clear how production of overly extended vowels may differ from more typical speech motor control.

Lastly, due to an error in our randomization code, all participants participated in the noisy session first and the veridical session second. While the effects of auditory perturbations on speech are thought to be relatively transient, if exposure to auditory perturbations has long-lasting effects on the speech motor system, this may have affected the compensation in the veridical session. However, given past studies involving multisession exposure to pitch perturbations (Behroozmand et al., 2020) or to formant adaptation paradigms (Scott et al., 2020), the likely direction of such an order effect would be a *decrease* in compensation, while our data showed a numerically *larger* compensatory response in the veridical condition. There was also no relationship between the time between sessions and change in compensation from the first to the second session ($r = .12$, $p = .47$). Nonetheless, since the effects of repeating speech compensation experiments across multiple sessions have not been established, it is possible the larger response seen in the veridical session was related to previous experience with the experiment.

Conclusions

Our results provide some support for our initial hypothesis that the sensitivity of the feedback control system in speech production is reduced by repeated exposure to auditory perturbations. This is consistent with theories of sensorimotor control that suggest sensory feedback should be weighted according to an internal estimate of its reliability. However, these results are tempered by the unexpected asymmetry between responses to upward and downward

formant perturbations, potential effects of session order, and the inconsistency of compensatory responses across analysis choices.

Acknowledgments

This work was supported by National Institutes of Health Grants R00 DC014520 (awarded to C. N.) and R01 DC017091 (awarded to B. P.).

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Behroozmand, R., Johari, K., Bridwell, K., Hayden, C., Fahey, D., & den Ouden, D.-B. (2020). Modulation of vocal pitch control through high-definition transcranial direct current stimulation of the left ventral motor cortex. *Experimental Brain Research*, 238(6), 1525–1535. <https://doi.org/10.1007/s00221-020-05832-9>
- Bischof, J., Gratzka, V., Strehlow, U., Haffner, J., Parzer, P., & Resch, F. (2002). Reliabilität, Trainierbarkeit und Stabilität auditiv diskriminativer Leistungen bei zwei computergestützten Mess- und Trainingsverfahren [Reliability, trainability and stability of auditory discrimination performance in 2 computer-assisted assessment and training methods]. *Zeitschrift Für Kinder-Und Jugendpsychiatrie Und Psychotherapie*, 30(4), 261–270. <https://doi.org/10.1024/1422-4917.30.4.261>
- Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer (Version 6.0.47) [Computer software]*. <http://www.praat.org/>
- 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](https://doi.org/10.1016/S0892-1997(97)80079-3)
- Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLOS ONE*, 7(7), e41830. <https://doi.org/10.1371/journal.pone.0041830>
- Cai, S., Boucek, M., Ghosh, S., Guenther, F. H., & Perkell, J. (2008). A system for online dynamic perturbation of formant trajectories and results from perturbations of the Mandarin Triphthong /iaul/. *Proceedings of the 8th International Seminar on Speech Production*, 65–68.
- Cai, S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2011). Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing. *Journal of Neuroscience*, 31(45), 16483–16490. <https://doi.org/10.1523/JNEUROSCI.3653-11.2011>
- Feng, Y., Gracco, V. L., & Max, L. (2011). Integration of auditory and somatosensory error signals in the neural control of speech movements. *Journal of Neurophysiology*, 106(2), 667–679. <https://doi.org/10.1152/jn.00638.2010>
- Franken, M. K., Acheson, D. J., McQueen, J. M., Eisner, F., & Hagoort, P. (2017). Individual variability as a window on production-perception interactions in speech motor control. *The Journal of the Acoustical Society of America*, 142(4), 2007–2018. <https://doi.org/10.1121/1.5006899>
- Guenther, F. H. (2016). *Neural control of speech*. The MIT Press. <https://doi.org/10.7551/mitpress/10471.001.0001>
- Herzfeld, D. J., Vaswani, P. A., Marko, M. K., & Shadmehr, R. (2014). A memory of errors in sensorimotor learning. *Science*, 345(6202), 1349–1353. <https://doi.org/10.1126/science.1253138>
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, 97(5), 3099–3111. <https://doi.org/10.1121/1.411872>
- Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience*, 5, 82. <https://doi.org/10.3389/fnhum.2011.00082>
- Hu, H., Liu, Y., Guo, Z., Li, W., Liu, P., Chen, S., & Liu, H. (2015). Attention modulates cortical processing of pitch feedback errors in voice control. *Scientific Reports*, 5(1), 7812. <https://doi.org/10.1038/srep07812>
- Katseff, S., Houde, J. F., & Johnson, K. (2012). Partial compensation for altered auditory feedback: A tradeoff with somatosensory feedback. *Language and Speech*, 55(2), 295–308. <https://doi.org/10.1177/0023830911417802>
- Kim, K., Wang, H., & Max, L. (2020). It's about time: Minimizing hardware and software latencies in speech research with real-time auditory feedback. *Journal of Speech, Language, and Hearing Research*, 63(8), 2522–2534. https://doi.org/10.1044/2020_JSLHR-19-00419
- Korzyukov, O., Bronder, A., Lee, Y., Patel, S., & Larson, C. R. (2017). Bioelectrical brain effects of one's own voice identification in pitch of voice auditory feedback. *Neuropsychologia*, 101, 106–114. <https://doi.org/10.1016/j.neuropsychologia.2017.04.035>
- Korzyukov, O., Sattler, L., Behroozmand, R., & Larson, C. R. (2012). Neuronal mechanisms of voice control are affected by implicit expectancy of externally triggered perturbations in auditory feedback. *PLOS ONE*, 7(7), e41216. <https://doi.org/10.1371/journal.pone.0041216>
- Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971), 244–247. <https://doi.org/10.1038/nature02169>
- Körding, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, 10(7), 319–326. <https://doi.org/10.1016/j.tics.2006.05.003>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13). <https://doi.org/10.18637/jss.v082.i13>
- Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of Neuroscience*, 32(27), 9351–9358. <https://doi.org/10.1523/JNEUROSCI.0404-12.2012>
- 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>
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). *emmeans: Estimated marginal means, aka least-squares means*. <https://cran.r-project.org/package=emmeans>
- 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, 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 Human Neuroscience*, 12, 113. <https://doi.org/10.3389/fnhum.2018.00113>
- Martin, C. D., Niziolek, C. A., Duñabeitia, J. A., Perez, A., Hernandez, D., Carreiras, M., & Houde, J. F. (2018). Online adaptation to altered auditory feedback is predicted by auditory acuity and not by domain-general executive control resources. *Frontiers in Human Neuroscience*, 12. <https://doi.org/10.3389/fnhum.2018.00091>

- Mollaie, 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>
- Munhall, K. G., Mitsuya, T., Normann, R. E., Nault, D. R., Haque, M. K., & Purcell, D. W. (2019, June 21). *Auditory feedback variability in vowel production* [Paper presentation]. Boston Speech Motor Control Symposium, Boston, MA, United States.
- Niziolek, C. A., & Guenther, F. H. (2013). Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *The Journal of Neuroscience*, 33(29), 12090–12098. <https://doi.org/10.1523/JNEUROSCI.1008-13.2013>
- Niziolek, C. A., & Houde, J. (2015). *Wave_Viewer: First release*. <https://doi.org/10.5281/ZENODO.13839>
- Niziolek, C. A., & Kiran, S. (2018). Assessing speech correction abilities with acoustic analyses: Evidence of preserved online correction in persons with aphasia. *International Journal of Speech-Language Pathology*, 20(6), 659–668. <https://doi.org/10.1080/17549507.2018.1498920>
- Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2013). What does motor efference copy represent? Evidence from speech production. *Journal of Neuroscience*, 33(41), 16110–16116. <https://doi.org/10.1523/JNEUROSCI.2137-13.2013>
- Parrell, B., Agnew, Z., Nagarajan, S., Houde, J. F., & Ivry, R. B. (2017). Impaired feedforward control and enhanced feedback control of speech in patients with cerebellar degeneration. *Journal of Neuroscience*, 37(38), 9249–9258. <https://doi.org/10.1523/JNEUROSCI.3363-16.2017>
- Parrell, B., Ramanarayanan, V., Nagarajan, S., & Houde, J. (2019). The FACTS model of speech motor control: Fusing state estimation and task-based control. *PLOS Computational Biology*, 15(9), e1007321. <https://doi.org/10.1371/journal.pcbi.1007321>
- Purcell, D. W., & Munhall, K. G. (2006). Compensation following real-time manipulation of formants in isolated vowels. *Journal of Acoustical Society of America*, 119(4), 2288–2297. <https://doi.org/10.1121/1.2173514>
- R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.R-project.org/>
- Reilly, K. J., & Dougherty, K. E. (2013). The role of vowel perceptual cues in compensatory responses to perturbations of speech auditory feedback. *Journal of the Acoustical Society of America*, 134(2), 1314–1323. <https://doi.org/10.1121/1.4812763>
- Saito, K., Sun, H., & Tierney, A. (2020). Brief report: Test–retest reliability of explicit auditory processing measures. *BioRxiv*. 2020.06.12.149484. <https://doi.org/10.1101/2020.06.12.149484>
- Sares, A. G., Deroche, M. L. D., 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. A. (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., & Jones, J. A. (2014). The predictability of frequency-altered auditory feedback changes the weighting of feedback and feedforward input for speech motor control. *European Journal of Neuroscience*, 40(12), 3793–3806. <https://doi.org/10.1111/ejn.12734>
- Scott, T. L., Haenchen, L., Daliri, A., Chartove, J., Guenther, F. H., & Perrachione, T. K. (2020). Noninvasive neurostimulation of left ventral motor cortex enhances sensorimotor adaptation in speech production. *Brain and Language*, 209, 104840. <https://doi.org/10.1016/j.bandl.2020.104840>
- Taylor, J. A., & Thoroughman, K. A. (2007). Divided attention impairs human motor adaptation but not feedback control. *Journal of Neurophysiology*, 98(1), 317–326. <https://doi.org/10.1152/jn.01070.2006>
- Tourville, J. A., Cai, S., & Guenther, F. (2013). Exploring auditory-motor interactions in normal and disordered speech. *Proceedings of Meetings on Acoustics*, 19(1), 060180. <https://doi.org/10.1121/1.4800684>
- 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>
- Tumber, A. K., Scheerer, N. E., & Jones, J. A. (2014). Attentional demands influence vocal compensations to pitch errors heard in auditory feedback. *PLOS ONE*, 9(10), e109968. <https://doi.org/10.1371/journal.pone.0109968>
- 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 Acoustical Society of America*, 122(4), 2306–2319. <https://doi.org/10.1121/1.2773966>
- Wei, K., & Körding, K. (2009). Relevance of error: What drives motor adaptation? *Journal of Neurophysiology*, 101(2), 655–664. <https://doi.org/10.1152/jn.90545.2008>

Appendix

Magnitude of Compensation Across Measurement Methods

	150–300 ms			200–300 ms		
	Average response	By trial [‡]	Nonnormalized (avg. response)	Average response [‡]	By trial [‡]	Nonnormalized (avg. response)
Compensation magnitude:						
Model	compensation~ direction × session + (1 + direction + session participant)	compensation~ direction × session + (1 + session participant)	compensation~ direction × session + (1 + direction + session participant)	compensation~ direction × session + (1 + direction participant)	compensation~ direction × session + (1 participant)	compensation~ direction × session + (1 + direction + session participant)
Direction	$F(1, 39) = 20.9$, $p < .001$	$F(1, 78) = 1.7$, $p = .19$	$F(1, 39) = 12.5$, $p = .001$	$F(1, 39) = 23.1$, $p < .0001$	$F(1, 117) = 14.6$, $p < .001$	$F(1, 39) = 14.9$, $p < .001$
Session	$F(1, 39) = .7$, $p = .42$	$F(1, 39) = .6$, $p = .45$	$F(1, 39) = .07$, $p = .80$	$F(1, 78) = .6$, $p = .48$	$F(1, 117) = .3$, $p = .56$	$F(1, 39) = .05$, $p = .83$
Direction × Session	$F(1, 39) = 4.6$, $p = .037$	$F(1, 78) = 1.8$, $p = .18$	$F(1, 39) = 2.2$, $p = .15$	$F(1, 78) = 6.0$, $p = .017$	$F(1, 117) = 1.8$, $p = .18$	$F(1, 39) = 3.6$, $p = .06$
Correlation between compensation and centering						
Veridical average	$r = -.29$, $p = .07$	$r = -.30$, $p = .06$	$r = -.18$, $p = .26$	$r = -.26$, $p = .10$	$r = -.25$, $p = .12$	$r = -.15$, $p = .35$
Veridical down	$r = -.32$, $p = .046$	$r = -.33$, $p = .04$	$r = -.13$, $p = .42$	$r = -.27$, $p = .10$	$r = -.24$, $p = .13$	$r = -.10$, $p = .54$
Veridical up	$r = -.18$, $p = .27$	$r = -.20$, $p = .22$	$r = -.18$, $p = .28$	$r = -.16$, $p = .33$	$r = -.16$, $p = .31$	$r = -.14$, $p = .37$
Noisy average	$r = -.11$, $p = .52$	$r = -.12$, $p = .47$	$r = -.08$, $p = .64$	$r = -.08$, $p = .60$	$r = -.09$, $p = .56$	$r = -.07$, $p = .69$
Noisy down	$r = -.03$, $p = .83$	$r = -.05$, $p = .76$	$r = -.01$, $p = .94$	$r = .00$, $p = .99$	$r = -.03$, $p = .86$	$r = .01$, $p = .95$
Noisy up	$r = -.14$, $p = .40$	$r = -.12$, $p = .48$	$r = -.12$, $p = .45$	$r = -.15$, $p = .38$	$r = -.09$, $p = .58$	$r = -.11$, $p = .49$
Correlation between compensation and initial variability						
Veridical average	$r = -.09$, $p = .60$	$r = -.02$, $p = .88$	$r = -.04$, $p = .83$	$r = -.12$, $p = .47$	$r = -.06$, $p = .71$	$r = -.09$, $p = .59$
Veridical down	$r = -.05$, $p = .75$	$r = -.11$, $p = .49$	$r = -.08$, $p = .62$	$r = -.04$, $p = .81$	$r = -.10$, $p = .56$	$r = -.06$, $p = .70$
Veridical up	$r = -.08$, $p = .63$	$r = .03$, $p = .84$	$r = .00$, $p = .99$	$r = -.12$, $p = .45$	$r = -.03$, $p = .88$	$r = -.08$, $p = .64$
Noisy average	$r = -.10$, $p = .52$	$r = -.18$, $p = .26$	$r = -.12$, $p = .45$	$r = -.09$, $p = .58$	$r = -.21$, $p = .20$	$r = -.07$, $p = .67$
Noisy down	$r = -.04$, $p = .83$	$r = -.10$, $p = .55$	$r = .07$, $p = .69$	$r = -.04$, $p = .83$	$r = -.13$, $p = .43$	$r = .08$, $p = .62$
Noisy up	$r = -.13$, $p = .41$	$r = -.17$, $p = .31$	$r = -.24$, $p = .14$	$r = -.11$, $p = .50$	$r = -.16$, $p = .32$	$r = -.20$, $p = .23$

Note. For the analyses of compensation magnitude, models with random slopes for each factor failed to converge or provided singular fits for some models. Reduced models were used in those cases, which increases the degrees of freedom of the F statistic. The final models used can be found in the table below. Cases where models with full random slopes failed to converge are indicated with a ‡. The estimated degrees of freedom for these models are larger than for the full models.