JSLHR

Research Article

Intact Correction for Self-Produced Vowel Formant Variability in Individuals With Cerebellar Ataxia Regardless of Auditory Feedback Availability

Benjamin Parrell,[a] Richard B. Ivry,[b] Srikantan S. Nagarajan,[c] and John F. Houde[d]

Purpose: Individuals with cerebellar ataxia (CA) caused by cerebellar degeneration exhibit larger reactive compensatory responses to unexpected auditory feedback perturbations than neurobiologically typical speakers, suggesting they may rely more on feedback control during speech. We test this hypothesis by examining variability in unaltered speech. Previous studies of typical speakers have demonstrated a reduction in formant variability (centering) observed during the initial phase of vowel production from vowel onset to vowel midpoint. Centering is hypothesized to reflect feedback-based corrections for self-produced variability and thus may provide a behavioral assay of feedback control in unperturbed speech in the same manner as the compensatory response does for feedback perturbations.
Method: To comprehensively compare centering in individuals with CA and controls, we examine centering in two vowels (/i/ and /ɛ/) under two contexts (isolated words and connected speech). As a control, we examine speech produced both with and without noise to mask auditory feedback.
Results: Individuals with CA do not show increased centering compared to age-matched controls, regardless of vowel, context, or masking. Contrary to previous results in neurobiologically typical speakers, centering was not affected by the presence of masking noise in either group.
Conclusions: The similar magnitude of centering seen with and without masking noise questions whether centering is driven by auditory feedback. However, if centering is at least partially driven by auditory/somatosensory feedback, these results indicate that the larger compensatory response to altered auditory feedback observed in individuals with CA may not reflect typical motor control processes during normal, unaltered speech production.

There is widespread agreement that speech production relies on a combination of feedback control and feedforward (or predictive) control (Guenther, 2016; Parrell & Houde, 2019; Parrell, Lammert, et al., 2019; Parrell, Ramanarayanan, et al., 2019). Feedback control entails generating motor commands based on sensory feedback (for speech, auditory and somatosensory) and enables speakers to alter ongoing motor actions when sensory feedback does not match expectations (so-called *sensory errors*). Feedforward control, on the other hand, entails generating motor commands based on either preplanned movements (so-called *open-loop* control) or a prediction of the state of the body at any given point (*predictive* control) and enables speakers to produce movements without the need to monitor afferent sensory feedback. For a review of control mechanisms in speech, see Parrell, Lammert, et al. (2019). Feedforward control is critical for accurate, fluent speech production given that delays in detecting and responding to unexpected sensory feedback range from 20–30 ms for somatosensory errors to 150–200 ms for auditory errors (for a review, see Parrell & Houde, 2019). Notably, for the latter, the delay is longer than the duration of a majority of speech sounds in connected speech. Given these constraints, feedforward control is thought to play the dominant role in speech, with only minor contributions from the feedback control system (Guenther, 2016).

[a]Department of Communication Sciences and Disorders, University of Wisconsin–Madison
[b]Department of Psychology, University of California, Berkeley
[c]Department of Radiology, University of California, San Francisco
[d]Department of Otolaryngology, University of California, San Francisco

Correspondence to Benjamin Parrell: bparrell@wisc.edu

Editor-in-Chief: Cara E. Stepp
Editor: Adam Buchwald

Received May 18, 2020
Revision received November 2, 2020
Accepted January 19, 2021
https://doi.org/10.1044/2021_JSLHR-20-00270

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.*

To assess the feedforward and feedback control systems, the most prevalent experimental approach is to perturb the auditory feedback a speaker hears. In these studies, a participant's speech is recorded, altered, and played back through headphones in near real time. Feedback control can be probed by examining responses to feedback perturbations that are delivered on an unpredictable schedule (e.g., variation in direction of the perturbation). With unpredictable perturbations, the online compensatory response speakers make to oppose the perturbation is reflective of feedback-based control (Purcell & Munhall, 2006b; Tourville et al., 2008). In general, speakers have been shown to change their *ongoing* speech to oppose the perturbation within 150–200 ms of perturbation onset (Cai et al., 2012; Tourville et al., 2008). Although there is considerable variability in this compensatory feedback response (e.g., Daliri et al., 2020; Parrell et al., 2017; Purcell & Munhall, 2006b), the predominant observation that transient feedback perturbations cause compensatory responses in ongoing speech shows that feedback control plays some role in speech production.

Feedforward control can be probed by examining a speaker's response to consistent perturbation that is maintained across trials to allow the feedforward system to learn or *adapt* over repeated exposure to oppose the perturbation. This approach has been used to show that speakers adapt their speech to oppose perturbations to vowel formants (Houde & Jordan, 1998, 2002; Purcell & Munhall, 2006a), vocal pitch (Jones & Munhall, 2000), and the spectral center of gravity of fricatives (Shiller et al., 2009). As with compensatory feedback responses, there is substantial variability between individual speakers in the magnitude of their adaptive responses (e.g., Munhall et al., 2009; Parrell et al., 2017; Purcell & Munhall, 2006a), but the fact that speakers alter their speech to oppose the auditory perturbation and retain these changes even after the perturbation is removed provides compelling evidence that the feedforward control system relies on auditory feedback to tune and maintain accurate speech production.

In previous work, we reported that individuals with ataxia caused by cerebellar degeneration (CA) exhibit changes in both the feedforward and feedback control systems for speech. Similar to what has been observed in studies of arm movements (Donchin et al., 2012; Martin et al., 1996; Schlerf et al., 2013; Smith & Shadmehr, 2005; Tseng et al., 2007), locomotion (Morton & Bastian, 2006), and saccades (Xu-Wilson et al., 2009), individuals with ataxia showed a marked impairment in adaptation to a persistent perturbation of vowel formants (Parrell et al., 2017). At the same time, this same population shows increased online, compensatory responses to perturbations of vowel formants (Parrell et al., 2017) and pitch (Houde et al., 2019; Li et al., 2019), though these compensatory response still only partially counteract the perturbation.

These experimental results are consistent with the hypothesis that cerebellar degeneration causes impairments in the accuracy of feedforward control while sparing the feedback control system. The increased compensatory response may be indicative of increased reliance on feedback control to maintain performance accuracy in the face of impaired feedforward control, a hypothesis that has been proposed with respect to upper limb movements in individuals with ataxia (Day et al., 1998). A similar deficit in accurate feedforward control, as well as a potentially increased reliance on feedback control, has been suggested to be the cause of the speech symptoms seen in ataxic dysarthria, the motor speech disorder associated with cerebellar ataxia (CA; Hartelius et al., 2000; Spencer & Dawson, 2019; Spencer & France, 2016; Spencer & Slocomb, 2007). Computationally, the larger feedback-based response can be modeled as an increased gain on the sensory error that drives feedback-based corrections (Houde & Nagarajan, 2011; Parrell, Ramanarayanan, et al., 2019; Scott, 2004; Shadmehr & Krakauer, 2008; Todorov & Jordan, 2002; Tourville & Guenther, 2011). It is currently an unresolved question as to what causes this change in feedback gains. As suggested above, it may emerge as the system reorganizes to adapt to the impairments in feedforward control. Alternatively, the change may be caused by cerebellar degeneration directly; within the framework of the DIVA (Directions Into Velocities of Articulators) model, the cerebellum is included in the feedback control pathway (Guenther, 2016; Tourville & Guenther, 2011). The change could also reflect a combination of these mechanisms.

The hypothesis that speakers with CA rely more on feedback control than neurobiologically typical speakers is motivated by the enhanced response exhibited by these individuals in response to auditory perturbations. This work is predicated on the assumption that responses to altered auditory feedback give insight into neural mechanisms underlying feedback control during typical, unaltered speech. However, it is important to recognize that feedback responses to unexpected experimentally imposed auditory perturbations may not reflect typical speech behavior.

Establishing whether the increased feedback sensitivity seen in this population in perturbation studies extends to their normal speech is important for a variety of reasons. First, changes in feedback control have been suggested to be a partial contributor to the speech symptoms associated with ataxic dysarthria (Spencer & Dawson, 2019; Spencer & France, 2016; Spencer & Slocomb, 2007). However, it is currently unknown if feedback control during natural speech is different in this population compared to neurobiologically typical speakers. Second, understanding whether findings from studies using altered feedback reflect typical speech production mechanisms is important for our understanding of speech motor disorders more generally. While the use of experimentally imposed perturbations has been used to study a variety of disorders, including individuals who stutter (Cai et al., 2012; Daliri et al., 2018; Kim et al., 2020; Sares et al., 2018) or have Parkinson's disease (Abur et al., 2018; Huang et al., 2016, 2019; Mollaei et al., 2016, 2013), apraxia of speech (Ballard et al., 2018), or hyperfunctional voice disorders (Stepp et al., 2017), the results from these studies have not been validated in unaltered speech production. In part, the challenge with translating results from altered auditory feedback studies to unaltered

speech production is methodological: It is difficult to measure the relative contributions of feedforward and feedback control during unperturbed speech.

A potential method to measure feedback control during natural speech has been established by Niziolek and colleagues (Niziolek & Kiran, 2018; Niziolek et al., 2013, 2015). They found that acoustic variability decreases from vowel onset to vowel midpoint, a phenomenon referred to as *centering*. Although the precise cause of centering has not been definitively established, two pieces of evidence suggest that it may reflect a correction for self-produced errors. First, magnetoencephalography (MEG) signals recorded from auditory cortex differentiate productions that are initially far from the median production from trials near the median (Niziolek et al., 2013). Productions near the median show a large suppression of the M100 component of the MEG signal compared to passively hearing those same productions, consistent with cancellation of the predicted sensory consequences of one's own movement (Houde et al., 2002; Kort et al., 2014). In contrast, productions far from the median show less suppression, suggesting they may be treated as sensory errors, similar to the response to (unpredicted) errors induced by altered auditory feedback (Chang et al., 2013; Kort et al., 2014). Second, auditory masking noise has been shown to reduce the magnitude of centering, suggesting at least a partial role for auditory feedback–based corrections in this phenomenon (Niziolek et al., 2015).

While the reduction in variability in centering is consistent with auditory feedback–based corrections that adjust the ongoing production to more closely approximate the vowel target, this is not the only possible explanation of the phenomenon. It is also possible that centering relies on somatosensation-based feedback correction, a feedforward controller that is able to predict variability at vowel onset, or feedforward control with smaller targets at vowel midpoint compared to vowel onset (Keating, 1990). Thus, in addition to testing for potential changes in centering in individuals with CA, we additionally include an assay of the role of auditory feedback in centering.

The goal of this study is to examine feedback control in speakers with CA during real-world, unaltered speech. We use the centering measure to test whether individuals with CA have an increased use of feedback control during normal, unperturbed speech. Based on previous studies showing increased compensatory responses to auditory feedback perturbations in this population, we hypothesize that individuals with ataxia will show increased centering compared to matched control participants. To increase the generalizability of the results, we assess centering not only in the typical case of single, isolated words (Niziolek & Kiran, 2018; Niziolek et al., 2013, 2015), but also in short multiword utterances. To evaluate the role of auditory feedback in centering, we also include a condition in which speech is produced in the presence of masking noise. Given that masking precludes online changes in production based on auditory feedback control, we further hypothesize that the ataxic and control group will perform similar in the masking condition (no centering) and differ only in the no-mask condition (larger centering in the ataxic group). The masking condition additionally allows us to test the extent to which centering is driven by auditory-based feedback control.

Method

Participants

Two groups of participants were recruited for the current experiment: individuals with clinically diagnosed ataxia due to cerebellar degeneration as confirmed by radiological assessment (n = 16, 12 women and four men, age = 55.3 ± 9.9 years) and age- and gender-matched neurologically healthy controls (n = 15, 12 women and three men, age = 56.5 ± 10.7 years). Of the 16 individuals with diagnosed cerebellar degeneration, nine were diagnosed with sporadic spinocerebellar ataxia (SCA), three with SCA2, two with SCA6, one with SCA14, and one with SCA15. Individuals with cerebellar degeneration were assessed for ataxia severity using the International Cooperative Ataxia Rating Scale (ICARS; Trouillas et al., 1997), which includes a gross measure of speech impairment. There is no standard medication used in the treatment of ataxia, and no patients reported taking any medications related to their neurological condition at the time of participation. None of the participants reported any history of hearing, speech, or neurological disorders other (Nasreddine et al., 2005) than CA. The participants also completed the Montréal Cognitive Assessment as a brief assessment of cognitive function. Two of the participants in the SCA group scored below 23 on the Montréal Cognitive Assessment exam, a score indicative of moderate cognitive impairment (Carson et al., 2017). These participants were included in the main experiment (and the pattern of results did not differ in when the analyses were performed without these two participants). Full participant information can be found in the Appendix. All protocols were approved by institutional review boards at the University of Delaware and the University of Wisconsin–Madison.

Behavioral Task

Participants sat in front of a computer screen. They were fitted with closed-back, over-the-ear headphones (Beyerdynamic DT 880). Speech was recorded through a separate head-mounted microphone (AKG C520) linked to a computer via a USB audio interface (Scarlett 2i2). Digitized speech signals were processed with MATLAB.

On each trial, a word or short phrase was displayed, and the participant was instructed to read the word(s) on the screen out loud. Each word was visible for 2.5 s, followed by an intertrial interval in which the screen was blank for 1 s. On half of the trials (randomly selected), speech-shaped noise was presented over the headphones at approximately 85 dB, with the noise commencing 0.5 s prior to the visual presentation of the stimulus word and lasting until the end of the trial. This noise served to mask

the participants' speech (see below for details on speech and noise amplitudes).

One of the known effects of speaking in the presence of loud noise is an increase in the amplitude of speech. Because this increase in amplitude could allow the speaker to hear their own voice even in the presence of 85 dB SPL masking noise, the participants were instructed to maintain a constant speech amplitude during the experiment. A visual guide (described below; see Figure 1) was provided to assist them in maintaining a relatively constant amplitude.

Amplitude Control

Prior to the experiment session itself, a target speech amplitude was determined for each participant. To set this target, participants spoke the word "Ed" 5 times, while 85-dB masking noise was played over the headphones. They were instructed to speak at a volume that was comfortable, yet quiet enough such that they did not hear themselves speaking given the presence of the masking noise. If, after the five productions, the participant indicated they could not hear themselves speaking, the target amplitude was set to the maximum amplitude of the production of "Ed" closest to the average of the five trials. If the participant reported they could hear themselves or if the experimenter deemed the productions too quiet (e.g., whispered), the process was repeated until an appropriate amplitude target could be determined.

During the experiment, a visual amplitude scale was shown to the right of the speech stimuli. The target amplitude was indicated by a horizontal line at the center of the scale. After each trial, a colored bar appeared on the scale, reflecting the maximum amplitude of the production. If the amplitude was within 25% of the target, the bar was green. If it was outside this range but within ± 50%, the bar was yellow. If the produced amplitude differed from the target by ± 50%, it was red and displayed at the top or bottom of the scale. Participants were instructed to use this information to help control their volume. The feedback bar remained visible during the intertrial interval.

Once the target amplitude was determined, a 16-trial training block was conducted to familiarize the participants with the visual amplitude scale. The block consisted of two trials of each of the four stimuli, each presented in clear and masking conditions. The training block was repeated for participants who, by the end of this block, had difficulty achieving amplitudes that fell within 25% of the target. All of the participants were able to maintain a relatively even amplitude by the second training block.

Conditions and Stimuli

The stimulus set consisted of four items (see Table 1), formed by the factorial combination of the vowel in the target word, /ɛ/ or /i/, and whether the target word was presented in isolation (one word) or in a short sentence (three words). For the three-word sentences, the target word was always the final word in the sentence. The vowels in the three-word sentences were all front vowels with an alternating high–low–high or low–high–low patterns. This pattern was chosen to maximize the amount of tongue movement between successive words.

The experimental session was composed of 480 trials, 60 repetitions of each of the four stimuli, presented in both masked and clear conditions. The order of stimuli and masking conditions was randomized for each participant.

Data Processing

For single-word utterances, onset and offset of the target vowels (/ɛ/ in *Ed* and /i/ in *eat*) were automatically labeled using a participant-specific amplitude threshold.

Figure 1. Illustration of experimental paradigm. (A) Trial structure. Participants were prompted with a stimulus word on every trial (left). Participants read those words out loud. At the end of the trial, participants were provided with visual feedback about their speech volume (right). Speech-shaped masking noise was present on 50% of trials. On these trials, the masking noise was played throughout the trial. (B) Experiment structure. First, a target speech amplitude was set for each participant at a comfortable speaking level where they could not hear their own speech under the speech-shaped masking noise used on some trials. Second, participants completed a 16-trial practice block to get used to the visual feedback about speech amplitude. This block was repeated if needed until participants could produce speech within the target region on most trials. Last, participants completed the main experiment.

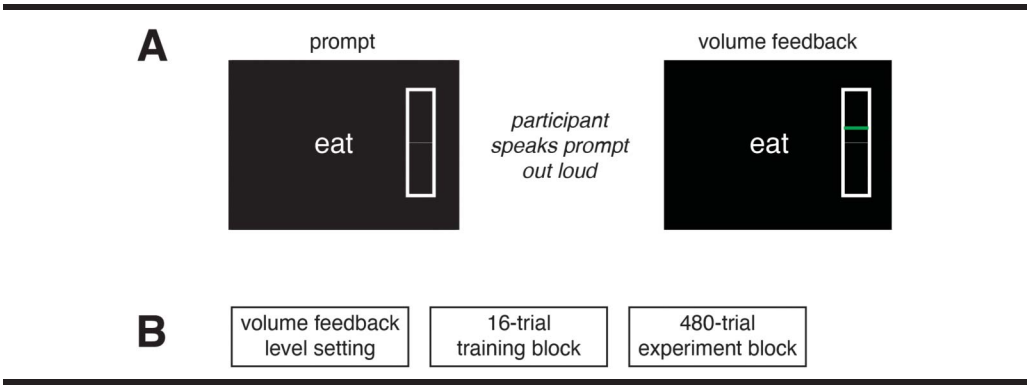


Table 1. Stimuli used in the one- and three-word conditions.

Vowel	1 word	3 words
/ɛ/	*Ed*	Dad sees *Ed.*
/i/	*eat*	He says *eat.*

Note. The target words where centering was measured appear in italics.

The labeled time points were examined using the speech waveform and spectrogram and manually adjusted when the automatic method was judged to have misidentified one or both of these points. Vowel onset was determined as the point at which periodicity in the waveform was visible, along with vowel formants in the spectrogram. Vowel offset was determined as the point where vowel formants, particularly the second formant (F2) and above, were no longer visible. Onset and offset for the target vowels in the three-word sentences were manually labeled using the same criteria, as the amplitude threshold method would detect only the first vowel (e.g., "dad" in "Dad sees Ed"). Vowel onset was identified as the first glottal pulse following the drop in high-frequency energy associated with the end of [z] in *sees* (*Dad sees Ed*) or *says* (*He says eat*). Vowel offset was identified as the point with a drop in the high-frequency energy associated with F2 (and higher formants).

The first (F1) and second (F2) vowel formants were tracked within the labeled vowel time period using Praat (Boersma & Weenink, 2019), using speaker-specific values for the number of linear predictive coding coefficients. These automatically tracked formant values were visually inspected using Wave Viewer (Niziolek & Houde, 2015). Where the vowels were judged to be incorrectly tracked, the linear predictive coding order or pre-emphasis value was adjusted. Trials with tracking errors that were not able to be corrected were discarded. These errors occurred in 8.5% of trials for the CA group and 2.4% of trials for the control group. The higher portion of errors for the CA group is likely due to voice quality changes (harsh or breathy voice) often seen in these individuals (Duffy, 2013).

Dependent Variables

The primary dependent variables were vowel duration, variability at onset, and centering. These were measured and calculated separately for each of the eight conditions (4 productions × 2 contexts, no mask/mask) on an individual basis. Duration was calculated as the time between vowel onset and vowel offset, as defined above. To calculate initial variability, the median F1 and F2 values were calculated during the first 50 ms of the vowel. Initial variability was defined as the average Euclidean distance, across trials, between each trial and the median point of all trials in that condition in the F1/F2 space. Midpoint variability was calculated the same way, using the middle 50% of the vowel around the vowel midpoint. Centering was then defined as the change between the measures of initial and midpoint variability (average initial distance − average midpoint distance; see Figure 2A). To control for the effects of coarticulation or surrounding speech segments on initial variability, centering was calculated separately for each word in each context.

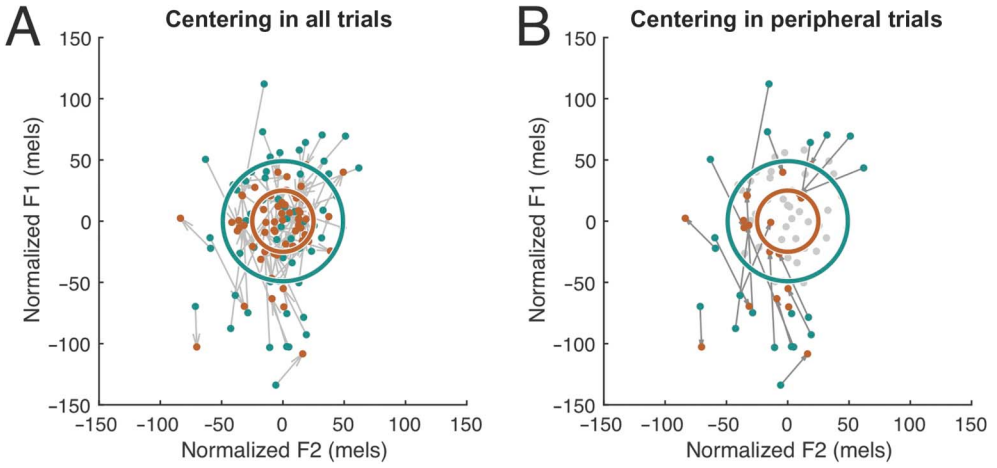
This global measure of centering captured the overall reduction in distance to the median. Motivated by the work of Niziolek and colleagues (Niziolek & Kiran, 2018; Niziolek et al., 2013, 2015), we also calculated an additional measure of centering. Using MEG, Niziolek et al. (2013) observed that the M100 response in the auditory cortex is lower for vowel productions relatively far from the median at vowel onset (peripheral trials) compared to productions near the median (center trials). Relevant to our present purposes, peripheral productions showed greater overall movement from onset to midpoint than center productions. This suggests that error correction may be greater on peripheral trials or may only occur on these trials. Given this, we include a separate measure of peripheral centering, using only peripheral trials (see Figure 2B). These were defined as the quartile of trials (approximately 15 trials) farthest from the median of the initial distance as defined above. For these trials, peripheral centering was calculated as in the global analysis (i.e., change in the distance to the median from vowel onset to vowel midpoint).

Lastly, vowel amplitude was measured to assess possible changes due to masking noise. Vowel amplitude was defined as the peak amplitude, in dB, during each target vowel. We note that absolute amplitude values cannot be compared across participants since the target amplitude for each individual was set in a subjective manner. Thus, amplitudes were normalized on an individual basis by subtracting the mean amplitude during the one-word trials with no masking noise. These normalized values were used for subsequent analysis.

Statistical Analysis

To test for hypothesized effects on centering, a linear mixed-effects model was constructed in R (R Core Team, 2013) with the *lme4* package (Bates et al., 2015). The model included the centering values as the dependent variable, with group (CA vs. control), masking condition (masking vs. clear), sentence length (one word vs. three words), and vowel (/ɛ/ vs. /i/) as fixed effects. Interaction terms between group and the other main effects were also included. Random effects of participants were included as random intercepts and slopes for main effects. Similar models were built for peripheral centering, vowel duration, and vowel amplitude. For amplitude, this model failed to converge, and a model with no random slopes was used. To test for a possible relationship between centering and initial variability (Niziolek et al., 2013), a model was constructed with centering as the dependent variable and fixed effects of initial variability and group, as well as their interaction. Again, random effects were modeled with the inclusion of random intercepts and random slopes for main effects. In order to evaluate the potential effects of speech or overall ataxia

Figure 2. Illustration of centering measurements for all trials and in peripheral trials. Example data from one participant shown. (A) Centering is defined as the reduction is distance to the vowel median in first (F1)/second (F2) formant space from vowel onset (teal) to vowel midpoint (orange). Individual trials are shown as two connected dots. Circles represent the average distance at vowel onset and midpoint. (B): The peripheral trials analysis was limited to the 25% of trials farthest from the vowel median at vowel onset (the excluded, more centralized vowels are shown in gray). Colors as in (A).



severity, additional models were constructed for each dependent measure for the CA group only. Instead of group, these models included fixed effects of both the speech subscore and overall ICARS scores. Statistical significance for all models was assessed with the *lmerTest* package (Kuznetsova et al., 2017).

Results

Laboratory studies with altered auditory feedback indicate that individuals with CA may exhibit increased reliance on feedback control. To examine if increased reliance is also observed in this population during normal speech production, we assessed the magnitude of centering or the degree to which vowel production becomes less variable over time. We predicted that the magnitude of centering would be larger in the CA group compared to that observed in the control group, consistent with greater corrections to self-produced variability. We expected this difference would be reduced or abolished when the speech was produced in the presence of masking noise, given the assumption that the noise would preclude auditory-based feedback control.

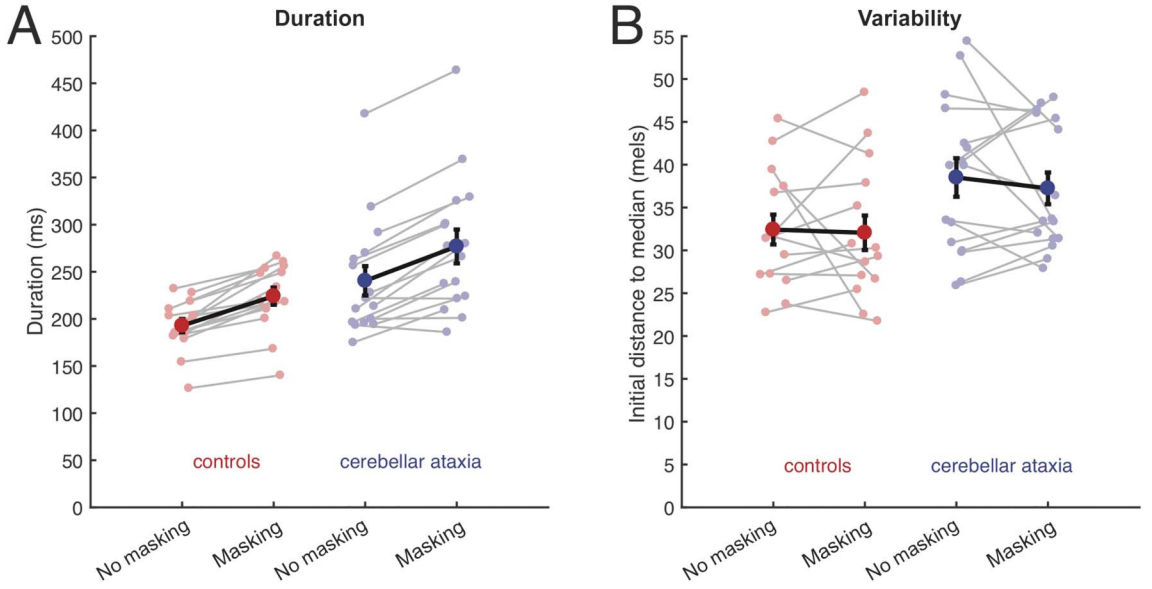
Consistent with previous reports (Hartelius et al., 2000; Kent & Kent, 2000; Kent et al., 2000, 1997), vowel duration was longer in the CA group compared to the control group (see Figure 3A; 259 ± 85 ms vs. 208 ± 40 ms), $F(1, 29) = 7.2$, $p = .01$. Also in accord with previous studies (Draegert, 1951; Hanley & Steer, 1949; Lane & Tranel, 1971; Summers et al., 1988), vowel duration was longer for both groups under masking noise compared to clear speech (252 ± 76 ms vs. 217 ± 62 ms), $F(1, 29) = 104.7$, $p < .0001$. There was no interaction between masking and group, $F(1, 29) = 1.7$, $p = .20$. The duration of the target vowel did not differ between the one- and three-word utterances, $F(1, 29) = 0.1$, $p = .83$, nor did this factor interact with group: interaction, $F(1, 29) = 3.1$, $p = .09$. Durations were shorter for /i/ than for /ɛ/ (210 ± 52 ms vs. 258 ± 81 ms), $F(1, 29) = 24.2$, $p < .0001$, and this difference was slightly larger for the CA group (64 ± 71 ms) than for the controls (31 ± 27 ms), although this difference was not significant, $F(1, 29) = 3.1$, $p = .09$.

In terms of variability, the CA group was slightly more variable than controls (see Figure 3B; mean distance to vowel median, 39 ± 12 mels vs. 33 ± 11 mels), $F(1, 29) = 5.2$, $p = .03$. There was an effect of utterance length, $F(1, 29) = 7.6$, $p = .01$, such that vowels were more variable in the three-word utterances (38 ± 13 mels) than the isolated words (33 ± 10 mels). There were no effects of masking noise, $F(1, 29) = 0.005$, $p = .95$, or vowel, $F(1, 29) = 0.003$, $p = .96$, on variability, nor any significant interactions between these factors and participant group (all $F < 1.1$, $p > .31$).

Turning to our main dependent variable of interest, centering was significantly larger than 0; two-tailed t tests: CA, $t(15) = 6.0$, $p < .0001$; controls, $t(14) = 5.1$, $p < .0002$. Critically, the amount of centering did not differ between the two groups (see Figure 4A; CA: 6.3 ± 9.7 mels; controls: 6.4 ± 9.5 mels), $F(1, 29) = 0.004$, $p = .95$. Although centering was larger in three-word utterances compared to one-word utterances, $F(1, 29) = 6.0$, $p = .02$, there was no interaction between utterance length and group, $F(1, 29) = 0.6$, $p = .43$, suggesting a similar effect of utterance length on both groups. There was no effect of vowel, $F(1, 240) = 0.6$, $p = .44$. Moreover, the similar responses in the two groups were seen across masking noise and vowels, as indicated by nonsignificant interaction effects between these terms and participant group (all $F < 1.1$, $p > .3$).

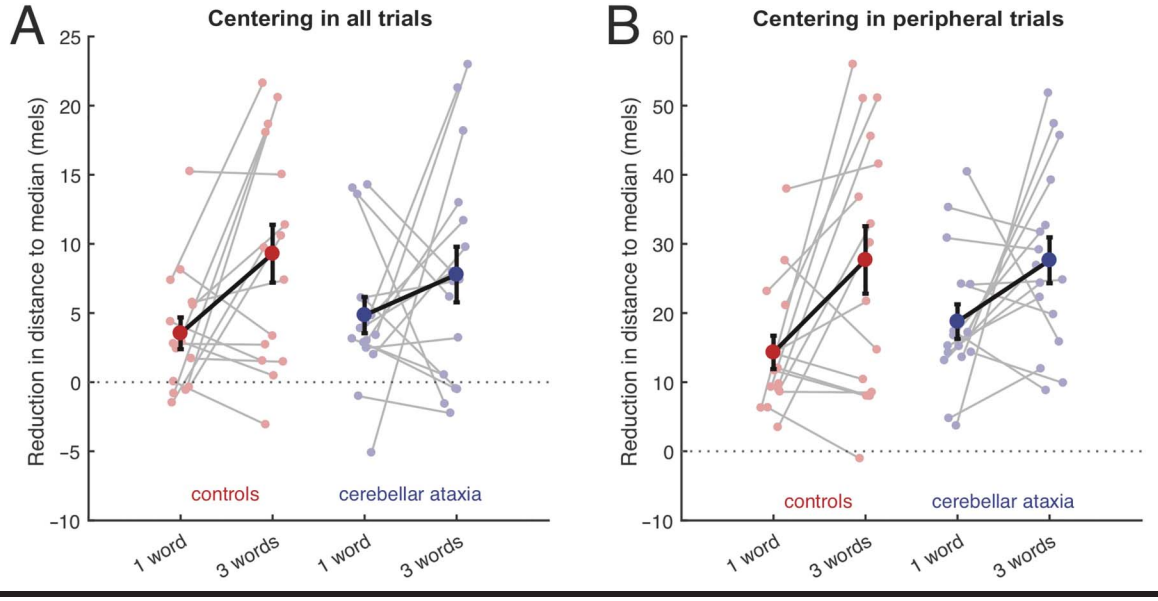
As described in the Method section, we also performed a secondary analysis restricted to the 25% of productions in which the vowel formants were initially farthest from the

Figure 3. Vowel duration and initial vowel variability. (A) Vowel duration by group. The cerebellar ataxia (CA) group is shown in blue, and controls are in red. Group means are shown in large, darker circles with standard errors; individual data are shown in small, light color circles. (B) Initial variability by group. Variability is measured as the average distance of all productions of a vowel to that vowel's median in first (F1)/second (F2) formant space. Colors and symbols as in (A). The CA group produced longer and more variable vowels. Vowel duration, but not variability, increased under masking noise for both groups.



median values (see Figure 4B). This analysis accounts for the possibility that only these "peripheral" trials may be considered to be errors by the speech sensorimotor system and drive corrective changes. The pattern of results from these peripheral trials was broadly similar to that observed for the full data set: There was no difference between the two participant groups, $F(1, 29) = 0.4$, $p = .52$, nor any interactions between participant group and the other factors,

Figure 4. Centering. (A) Centering among all productions. The cerebellar ataxia (CA) group is shown in blue, and controls are in red. Group means are shown in large, darker circles with standard errors; individual data are shown in small, light color circles. (B) Centering in peripheral trials only. Colors and symbols as in (A). There is no difference for either measure of centering between the two groups. Centering is significantly greater for the three-word utterances than in single words for both analyses.



all $F < 0.5$, $p > .49$. There was a significant effect of utterance length for the peripheral trials, such that there was more centering in three-word utterances (28 ± 16 mels) compared to single words (17 ± 10 mels), $F(1, 29) = 10.6$, $p = .003$. In summary, both the primary and secondary analyses failed to support the hypothesis that feedback control, as inferred from centering, is increased in individuals with CA.

The degree of centering has been shown to be related, across individuals, to the amount of initial variability (Niziolek & Kiran, 2018; Niziolek et al., 2013). We observed this relationship in our data set (see Figure 5A), $F(1, 243) = 313.6$, $p < .0001$. There was no effect of group, $F(1, 173) = 0.3$, $p = .57$, nor interaction of Group × Initial Variability, $F(1, 243) = 0.48$, $p = .49$, indicating that the relationship was similar for the CA and control groups. The overall linear mixed-effects model with main effects of initial variability, group, and their interaction, had a marginal R^2 of .67 (a model fitted with only the significant predictor of variability had a marginal R^2 of .68).

We also considered how production differences between the two groups might impact our estimates of centering. First, the longer vowel durations for the CA group might distort centering effects given there is more time to implement a corrective action. However, an analysis that included duration as a factor in the model also revealed no group effect on centering, $F(1, 28) = 1.2$, $p = .29$, nor interaction of group and utterance length, masking noise, or vowel (all $F < 0.6$, $p > .44$). Second, given the positive correlation between variability and centering, and greater variability in the CA group, we derived a normalized measure of centering by dividing the basic centering score by initial variability (see Figure 5B). As with the other measures of centering, there was no group effect on this measure, $F(1, 29) = 0.2$, $p = .70$, nor did this factor interact with any other factors (all $F < 0.87$, $p > .36$).

Surprisingly, we failed to observe a centering difference between the no-mask (6.0 ± 0.8 mels) and masked (6.7 ± 0.9 mels) feedback conditions (see Figure 6A), $F(1, 29) = 0.5$, $p = .47$). This null effect is at odds with prior studies showing centering is reduced when auditory feedback is masked (Niziolek et al., 2015) and, indeed, challenges the core assumption that centering provides a measure of auditory feedback control. However, it may also indicate that our masking method was incomplete in eliminating feedback control from auditory feedback. People generally increase speech amplitude when exposed to loud masking noise (Draegert, 1951; Hanley & Steer, 1949; Lane & Tranel, 1971; Summers et al., 1988), and this effect was observed in our data: Normalized amplitude was considerably higher in trials with masking noise than in trials without masking (see Figure 6B; 6.8 ± 8.4 dB vs. −2.1 ± 2 dB), $F(1, 211) = 118.6$, $p < .0001$. Moreover, this increase under masking noise was larger in the CA group than in controls (14.0 ± 9.6 dB vs. 3.4 ± 5.2 dB), $F(1, 211) = 41.3$, $p < .0001$. Thus, despite the visual feedback, the CA participants, but not the control group, did produce louder speech under masking noise. It is not clear if this increase in amplitude would be large enough to perceive speech over the 85 dB SPL masking noise. In particular, the lack of a masking effect cannot be

Figure 5. Centering and variability. (A) Relationship between centering and initial variability. The cerebellar ataxia (CA) group is shown in blue, and controls are in red. An estimate of the regression derived from a mixed-effects linear model is shown as a solid black line with confidence intervals shown in dashed lines. Centering is highly correlated with initial variability, and there is no difference between the two groups. (B) Normalized centering as a percentage of initial variability. Colors as in (A). Group means are shown in large, darker circles with standard errors, and individual data are shown in small, light color circles. Normalized centering is similar across groups.

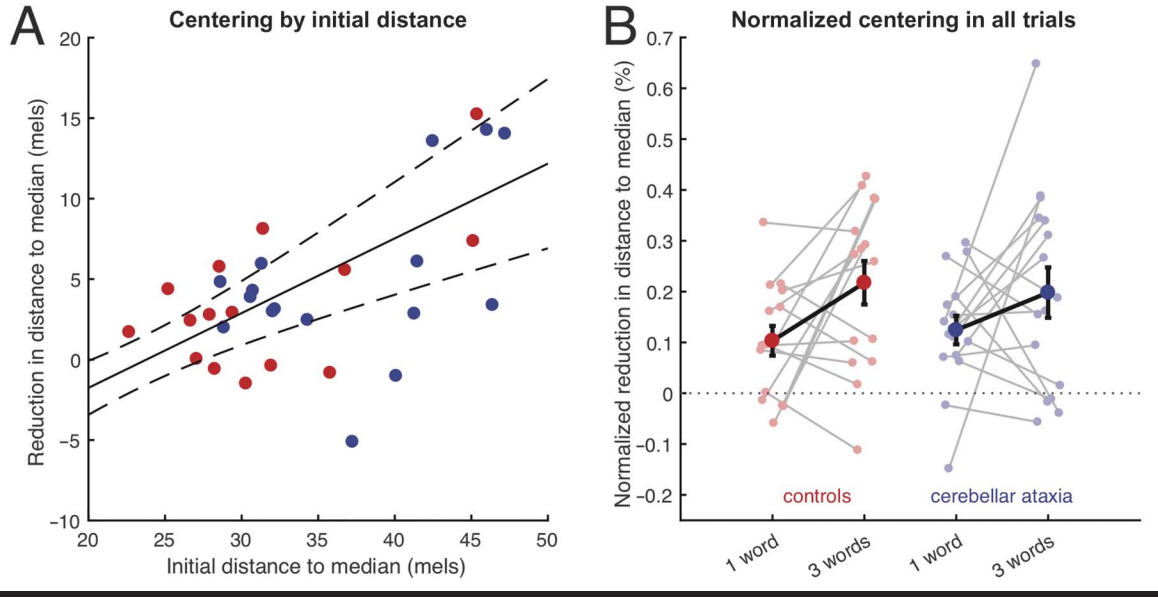
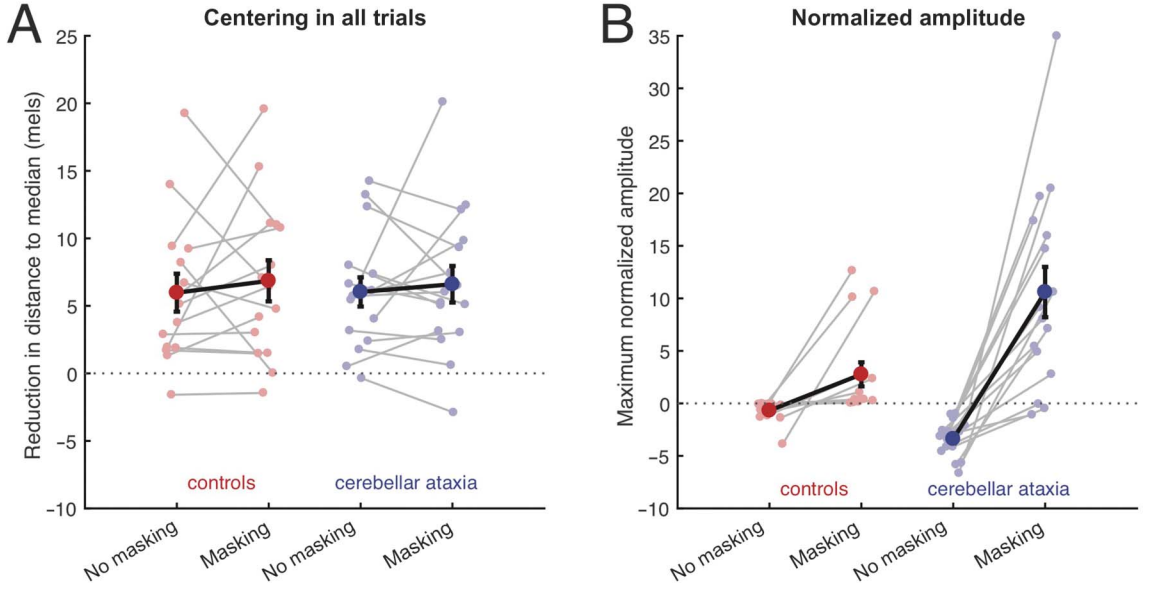


Figure 6. Effect of noise masking. (A) Centering for the cerebellar ataxia (CA) group (blue) and controls (red), as a function of masking condition. Group means are shown in large, darker circles with standard errors; individual data are shown in small, light color circles. There is no difference in the amount of centering when producing speech with and without masking noise. (B) Amplitudes are normalized to each participant's mean amplitude when producing isolated words without masking noise. Colors and symbols as in (A). Both groups produce louder speech under masking noise, and this increase was larger in the CA group.



explained by this potential increase, as we would then expect a decrease in centering under masking in the controls (who showed a minimal increase in amplitude) compared to the CA group (who produced much louder speech). This pattern was not found.

Dysarthria severity, as indexed by the relatively gross measure of the ICARS speech subscore, was significantly related to vowel duration, $F(1, 13) = 19.6$, $p = .0007$, such that individuals with more severe symptoms produced longer vowel durations. There was no significant relationship between either the speech severity (as indexed by the ICARS speech subscore) or general ataxia severity (full ICARS score) and any of our other dependent measures (all $p > .29$).

Discussion

Individuals with CA have been shown to have an increased compensatory feedback response for externally introduced auditory errors during speech production. Based on this work, we hypothesized that these individuals may have an increased reliance on feedback control compared to neurobiologically typical individuals. To test this hypothesis, we examined whether a similar increase in feedback sensitivity is seen in these individuals during normal, unperturbed speech production. As a measure of feedback sensitivity, we quantified the degree of centering or reduction in variability from vowel onset to vowel midpoint in a group of individuals with CA and a group of age-matched, neurobiologically typical controls. As expected, the individuals with CA produced longer and more variable vowels. However, we found no evidence that these individuals make larger corrections in response to self-produced variability than neurobiologically typical controls. This result held whether we measured centering as a reduction in overall variability, as the reduction in variability of only those trials that are initially far from the vowel median or as a normalized value based on initial variability. In addition, we found that centering was not affected by blocking auditory feedback with masking noise in either individuals with ataxia or controls.

We see three possible explanations for this pattern of results. First, our primary dependent measure of centering may not be an appropriate assay of the feedback control system in speech. Second, individuals with CA may not have any changes in feedback control for speech relative to neurobiologically typical individuals. Third, individuals with CA form a heterogeneous group, and our relatively limited sample may not have included enough speakers with altered feedback control. We address each of these possibilities in turn.

The first explanation we have for the lack of any enhancement in centering in individuals with CA may be that the centering measure itself is not assessing auditory-based feedback control. Previous work has argued that centering is at least partially driven by feedback control based on both neurological (Niziolek et al., 2013) and behavioral (Niziolek et al., 2015) data. From a neurological perspective, analysis of MEG signals indicates that productions far from a vowel's median location in F1/F2 space (the "peripheral" trials, as analyzed in this study) show less suppression in auditory cortex during speech than trials near the median location. This is similar, but smaller in

magnitude, than the so-called *release from auditory suppression* seen when auditory feedback of speech is externally perturbed (Chang et al., 2013; Houde et al., 2002; Kort et al., 2014, 2016). This suggests that peripheral trials generate sensory prediction errors, which would drive feedback-based corrections. Strikingly, participants who show a larger difference in auditory suppression between center and peripheral trials also show more centering, suggesting that this behavior may be driven by these sensory errors. Behaviorally, previous work showed that centering is modulated by the availability of sensory feedback, such that centering is reduced, though not eliminated, as the amplitude of auditory masking noise increases (Niziolek et al., 2015). These results indicate at least a partial role for auditory feedback as a driving factor in centering.

However, these results are challenged by the lack of any modulatory effect of masking noise on centering in the current study. That is, if centering is unaffected by the presence of masking noise that blocks auditory feedback, it is unlikely to be driven by auditory feedback under normal conditions. Of course, this does not prevent a role for somatosensory feedback in driving this process. Mechanical perturbations that are processed through somatosensory feedback are known to drive both corrective speech behaviors in online control (Abbs & Gracco, 1984; Fowler & Turvey, 1981; Gracco & Abbs, 1985; Kelso et al., 1984; Shaiman & Gracco, 2002) and changes to feedforward control (Lametti et al., 2012; Tremblay et al., 2003).

There are a number of differences between the current study and Niziolek et al.'s (2015) study that should be examined more carefully. Our study used a by-trial design, where masking noise was present on a random selection of trials, while Niziolek et al. used a blocked design for masking noise. This may have affected the current results in three ways. First, it seems that people are unable to prevent louder speech under unexpected masking noise, even when given explicit feedback about speech amplitude after each trial, as indicated by the louder vowels produced under masking noise in this study. The louder speech produced may have caused speech to be at least somewhat audible even in the masking condition. Second, a blocked design may cause longer term changes in behavior, as the system is able to predict whether auditory feedback will be available, which was not the case in our study. The frequent presence of noise in the current study may have led participants to rely less on auditory feedback across all trials, reducing the overall magnitude of centering but mitigating any potential modulation by the masking noise. Whatever the explanation for the unexpected (lack of an) effect of masking noise in the current study, further work is needed to isolate the source or sources of centering in speech production to determine to what extent this behavior reflects the feedback control system.

The second possibility is simply that individuals with CA have an intact feedback control system for speech motor control but do not exhibit higher feedback gains or increased reliance on this system during normal speech production. This explanation receives some support from a recent study on limb control in this population (Zimmet et al., 2020). This study used a combination of experimental results from a visuomotor tracking task and with computational modeling to show that feedback gains are equivalent in individuals with CA and matched neurobiologically typical controls. Critically, this study did not find any enhancement of feedback gains in the individuals with ataxia, consistent with the hypothesis that cerebellar atrophy predominantly affects the predictive/feedforward control system.

While this explanation of the data is both straightforward and consistent with current results from the limb control system, it may be difficult to square previous results using altered auditory feedback paradigms that show increased compensatory changes for unexpected perturbations of both vowel formants and pitch (Houde et al., 2019; Li et al., 2019; Parrell et al., 2017). We see a few possibilities here. One is that large auditory errors of the kind introduced in altered feedback studies are treated differently by the speech motor system than smaller errors that occur naturally. There is some evidence that the size of sensory errors does affect how much the system relies on or weights those errors (Wei & Körding, 2009), but it is generally thought that larger sensory errors lead to a down-weighting of sensory feedback. However, this pattern may be different in individuals with CA. This hypothesis could be tested directly.

Alternatively, the difference may stem from a distinction between sensory errors and positional or targeting errors. Auditory perturbations introduce a difference between expected and perceived sensory signals. These perturbations are, by definition, unpredictable based on the output of the speech sensorimotor system and thus result in *sensory prediction errors,* a mismatch between the predicted sensory consequences of an utterance and the resultant auditory feedback. This may differ from errors that occur during the course of typical speech, where the acoustic signal is a predictable consequence of articulator movement (even when that movement is incorrect). While the source of an auditory error does not matter in some models of speech production, such as DIVA (Guenther, 2016; Tourville & Guenther, 2011), it is relevant in other models based on state feedback control, such as the State Feedback Control and Feedback Aware Control of Tasks in Speech models proposed by our group (Houde & Nagarajan, 2011; Parrell, Ramanarayanan, et al., 2019). In state feedback control, sensory prediction errors, along with a prediction about the state of the speech system, are used to generate a weighted estimate of the speech system state. Separately, this state estimate is compared against the movement target to generate a targeting error. Thus, separate feedback gains are applied to the sensory prediction errors and to the difference between the current state of the articulators and their desired positions (Scott, 2004; Shadmehr & Krakauer, 2008; Todorov & Jordan, 2002). According to these models, an increased gain on sensory errors without any change on targeting errors would result in larger compensatory responses to perturbed feedback but would have a limited effect on real-world speech when auditory feedback is well predicted and sensory errors are close to 0. This could be the case for speech in individuals with ataxia.

Lastly, it is possible that the experimental results showing increased feedback sensitivity to auditory perturbations in individuals with ataxia are unreliable. For pitch control, this seems unlikely, given the large effects seen in multiple studies (Houde et al., 2019; Li et al., 2019). However, for the compensation to formant perturbations used to assess oromotor speech control, this may be the case. While our previous results in this paradigm did show a significant increase in this compensatory response in individuals with CA (Parrell et al., 2017), the effect was relatively modest compared to the response seen in pitch. If this effect is indeed small, we may be unlikely to detect it with the sample size in this study. Alternatively, there may be no group-level enhancement in feedback control, and our previous result did not reflect the true population. Further work in this area could sort out these possibilities.

Of course, this raises the question of why sensory gains should be higher for pitch control than oromotor control in speech. It may simply be the case that pitch control inherently relies more on auditory feedback than the oromotor system. This is consistent with the fast degradation of pitch control following the onset of postlingual deafness, but a much slower loss of fine control of the oral articulators (Lane & Webster, 1991; Svirsky et al., 1992). Thus, it may be that changes in feedback control in individuals with CA are domain specific and potentially task specific. We are currently examining this question by testing how feedback control of vocal pitch may be affected in this population in the absence of external auditory perturbations.

A final explanation for the lack of any difference in centering between groups seen in the current study is that cerebellar degeneration may have heterogeneous effects on the speech motor control system. This has been suggested in a number of studies examining the clinical symptoms seen in ataxic dysarthria, the motor speech disorder associated with cerebellar damage (Hartelius et al., 2000; Spencer & Dawson, 2019; Spencer & France, 2016; Spencer & Slocomb, 2007). Broadly speaking, it seems as though some patients produce more variable speech than normal and some less variable speech than normal, especially for temporal control. Spencer and colleagues have suggested, as one possibility, that the underlying cause of both groups of symptoms may be impairments in predictive/feedforward control that are consistently associated with cerebellar damage across motor domains. As a compensatory mechanism, some individuals may subsequently increase their reliance on sensory feedback to guide motor actions, with the effect of reducing spatial errors but consequently severely affecting the timing of speech (Spencer & France, 2016; Spencer & Slocomb, 2007). However, it is not known how prevalent these different subtypes are in general, nor how they may be associated with different subtypes of SCA. While some progress has been made in this area (Spencer & Dawson, 2019; Spencer & France, 2016), the number of patients examined in these studies is very small (< 10). Thus, given the heterogeneity in feedback control expected in this population, it is possible that our sample did not include a large portion of individuals with altered feedback control. This could be addressed in the future by testing a large sample of the population and identifying, a priori, which participants are expected to show changes in feedback control based on an analysis of dysarthria symptoms.

An additional potential limitation of the current study is that participants' hearing thresholds were not directly assessed. Thus, it is possible that individuals with CA had less sensitive hearing overall, which may have counteracted any potential increase in centering that might be driven by increased feedback usage. However, all participants were screened for any reported hearing problems, none reported any history of hearing issues, and ages were matched between individuals with ataxia and controls.

In summary, we tested whether individuals with CA show increased feedback sensitivity during unperturbed speech production by examining the magnitude of centering or reduction in vowel variability from vowel onset to vowel midpoint. We found no difference in the magnitude of centering when auditory feedback was blocked with loud masking noise, contrary to previous results. The lack of the expected decrease in centering under masking noise questions whether auditory feedback plays a prominent role in this behavior, as has previously been suggested.

Separately, we found no evidence for any difference in the magnitude of centering between individuals with CA and age-matched neurobiologically typical controls. Although our results indicate that centering is not driven by auditory feedback, it remains possible that centering is, at least partially, under feedback-based control (e.g., through somatosensation). If this is the case, the results would suggest that individuals with ataxia have an intact, but not increased, feedback control response for oromotor control in speech. This is consistent with results from limb control but unexpected given previous results from altered auditory feedback studies in speech.

The difference between feedback control in altered and unaltered speech in individuals with ataxia may suggest that speech produced under conditions of altered auditory feedback may not reflect typical speech production, at least in this population. This has implications for our understanding of the potential sources of speech symptoms observed in this population (Spencer & Slocomb, 2007) and, more broadly, highlights the importance of testing theories derived from studies using feedback alterations in more typical speech production contexts. However, some caution is warranted in drawing strong conclusions from the current results given the critical comparison here between feedback control in unaltered and perturbed speech is a post hoc comparison between experiments with different participants and, especially, given the unanticipated finding of a lack of modulation of centering magnitude by masking noise.

Acknowledgments

This work was supported by grants from the National Institute on Deafness and Other Communication Disorders (R01 DC017091 and DC176960 to J. F. H. and S. S. N.). We would

like to thank Carrie Niziolek for her assistance with the centering analysis.

References

Abbs, J. H., & Gracco, V. L. (1984). Control of complex motor gestures: Orofacial muscle responses to load perturbations of lip during speech. *Journal of Neurophysiology, 51*(4), 705–723. https://doi.org/10.1152/jn.1984.51.4.705

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

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

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

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (Version 6.0.47) [Computer software]. http://www.praat.org/

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

Carson, N., Leach, L., & Murphy, K. J. (2017). A re-examination of Montréal Cognitive Assessment (MoCA) cutoff scores. *International Journal of Geriatric Psychiatry, 33*(2), 379–388. https://doi.org/10.1002/gps.4756

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

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

Daliri, A., 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

Day, B. L., Thompson, P. D., Harding, A. E., & Marsden, C. D. (1998). Influence of vision on upper limb reaching movements in patients with cerebellar ataxia. *Brain, 121*(2), 357–372. https://doi.org/10.1093/brain/121.2.357

Donchin, O., Rabe, K., Diedrichsen, J., Lally, N., Schoch, B., Gizewski, E. R., & Timmann, D. (2012). Cerebellar regions involved in adaptation to force field and visuomotor perturbation. *Journal of Neurophysiology, 107*(1), 134–147. https://doi.org/10.1152/jn.00007.2011

Draegert, G. L. (1951). Relationships between voice variables and speech intelligibility in high level noise. *Speech Monographs, 18*(4), 272–278. https://doi.org/10.1080/03637755109375042

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). Elsevier.

Fowler, C. A., & Turvey, M. T. (1981). Immediate compensation in bite-block speech. *Phonetica, 37*(5–6), 306–326. https://doi.org/10.1159/000260000

Gracco, V. L., & Abbs, H. J. (1985). Dynamic control of the perioral system during speech: Kinematic analyses of autogenic and nonautogenic sensorimotor processes. *Journal of Neurophysiology, 54,* 418–432. https://doi.org/10.1152/jn.1985.54.2.418

Guenther, F. H. (2016). *Neural control of speech*. The MIT Press. https://doi.org/10.7551/mitpress/10471.001.0001

Hanley, T. D., & Steer, M. D. (1949). Effect of level of distracting noise upon speaking rate, duration and intensity. *Journal of Speech and Hearing Disorders, 14*(4), 363–368. https://doi.org/10.1044/jshd.1404.363

Hartelius, L., Runmarker, B., Andersen, O., & Nord, L. (2000). Temporal speech characteristics of individuals with multiple sclerosis and ataxic dysarthria: "Scanning speech" revisited. *Folia Phoniatrica et Logopaedica, 52*(5), 228–238. https://doi.org/10.1159/000021538

Houde, J. F., Gill, J. S., Agnew, Z., Kothare, H., Hickok, G., Parrell, B., Ivry, R. B., & Nagarajan, S. S. (2019). Abnormally increased vocal responses to pitch feedback perturbations in patients with cerebellar degeneration. *The Journal of the Acoustical Society of America, 145*(5), EL372–EL378. https://doi.org/10.1121/1.5100910

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science, 279*(5354), 1213–1216. https://doi.org/10.1126/science.279.5354.1213

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech, Language, and Hearing Research, 45*(2), 295–310. https://doi.org/10.1044/1092-4388(2002/023)

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

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience, 14*(8), 1125–1138. https://doi.org/10.1162/089892902760807140

Huang, X., Chen, X., Yan, N., Jones, J. A., Wang, E. Q., Chen, L., Guo, Z., Li, W., Liu, P., & Liu, H. (2016). The impact of Parkinson's disease on the cortical mechanisms that support auditory-motor integration for voice control. *Human Brain Mapping, 37*(12), 4248–4261. https://doi.org/10.1002/hbm.23306

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., & Munhall, G. K. (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

Keating, P. A. (1990). The window model of coarticulation: Articulatory evidence. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I* (pp. 451–470). Cambridge University Press. https://doi.org/10.1017/CBO9780511627736.026

Kelso, J. A. S., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C. A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance, 10*(6), 812–832. https://doi.org/10.1037/0096-1523.10.6.812

Kent, R. D., & Kent, J. F. (2000). Task-based profiles of the dysarthrias. *Folia Phoniatrica et Logopaedica, 52*(1–3), 48–53. https://doi.org/10.1159/000021512

Kent, R. D., Kent, J. F., Duffy, J. R., Thomas, J. E., Weismer, G., & Stuntebeck, S. (2000). Ataxic dysarthria. *Journal of Speech,*

Language, and Hearing Research, 43(5), 1275–1289. https://doi.org/10.1044/JSLHR.4305.1275

Kent, R. D., Kent, J. F., Rosenbek, J. C., Vorperian, H. K., & Weismer, G. (1997). A speaking task analysis of the dysarthria in cerebellar disease. *Folia Phoniatrica et Logopaedica, 49*(2), 63–82. https://doi.org/10.1159/000266440

Kim, K. S., Daliri, A., Flanagan, J. R., & Max, L. (2020). Dissociated development of speech and limb sensorimotor learning in stuttering: Speech auditory-motor learning is impaired in both children and adults who stutter. *Neuroscience, 451,* 1–21. https://doi.org/10.1016/j.neuroscience.2020.10.014

Kort, N. S., Cuesta, P., Houde, J. F., & Nagarajan, S. S. (2016). Bihemispheric network dynamics coordinating vocal feedback control. *Human Brain Mapping, 37*(4), 1474–1485. https://doi.org/10.1002/hbm.23114

Kort, N. S., Nagarajan, S. S., & Houde, J. F. (2014). A bilateral cortical network responds to pitch perturbations in speech feedback. *NeuroImage, 86,* 525–535. https://doi.org/10.1016/j.neuroimage.2013.09.042

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

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

Lane, H., & Webster, J. W. (1991). Speech deterioration in postlingually deafened adults. *The Journal of the Acoustical Society of America, 89*(2), 859–866. https://doi.org/10.1121/1.1894647

Li, W., Zhuang, J., Guo, Z., Jones, J. A., Xu, Z., & Liu, H. (2019). Cerebellar contribution to auditory feedback control of speech production: Evidence from patients with spinocerebellar ataxia. *Human Brain Mapping, 40*(16), 4748–4758. https://doi.org/10.1002/hbm.24734

Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J., & Thach, W. T. (1996). Throwing while looking through prisms. I. Focal olivocerebellar lesions impair adaptation. *Brain, 119*(4), 1183–1198. https://doi.org/10.1093/brain/119.4.1183

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

Morton, S. M., & Bastian, A. J. (2006). Cerebellar contributions to locomotor adaptations during splitbelt treadmill walking. *Journal of Neuroscience, 26*(36), 9107–9116. https://doi.org/10.1523/JNEUROSCI.2622-06.2006

Munhall, G. K., MacDonald, E. N., Byrne, S. K., & Johnsrude, I. (2009). Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *The Journal of the Acoustical Society of America, 125*(1), 384–390. https://doi.org/10.1121/1.3035829

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

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

Niziolek, C. A., Nagarajan, S. S., & Houde, J. F. (2015). The contribution of auditory feedback to corrective movements in vowel formant trajectories. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. The University of Glasgow.

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

Parrell, B., Lammert, A. C., Ciccarelli, G., & Quatieri, T. F. (2019). Current models of speech motor control: A control-theoretic overview of architectures and properties. *The Journal of the Acoustical Society of America, 145*(3), 1456–1481. https://doi.org/10.1121/1.5092807

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

Purcell, D. W., & Munhall, K. G. (2006b). Compensation following real-time manipulation of formants in isolated vowels. *The Journal of 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/

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

Schlerf, J. E., Xu, J., Klemfuss, N. M., Griffiths, T. L., & Ivry, R. B. (2013). Individuals with cerebellar degeneration show similar adaptation deficits with large and small visuomotor errors. *Journal of Neurophysiology, 109*(4), 1164–1173. https://doi.org/10.1152/jn.00654.2011

Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience, 5*(7), 532–546. https://doi.org/10.1038/nrn1427

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research, 185*(3), 359–381. https://doi.org/10.1007/s00221-008-1280-5

Shaiman, S., & Gracco, V. L. (2002). Task-specific sensorimotor interactions in speech production. *Experimental Brain Research, 146*(4), 411–418. https://doi.org/10.1007/s00221-002-1195-5

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America, 125*(2), 1103–1113. https://doi.org/10.1121/1.3058638

Smith, M. A., & Shadmehr, R. (2005). Intact ability to learn internal models of arm dynamics in Huntington's disease but not cerebellar degeneration. *Journal of Neurophysiology, 93*(5), 2809–2821. https://doi.org/10.1152/jn.00943.2004

Spencer, K. A., & Dawson, M. (2019). Dysarthria profiles in adults with hereditary ataxia. *American Journal of Speech-Language Pathology, 28*(2S), 915–924. https://doi.org/10.1044/2018_AJSLP-MSC18-18-0114

Spencer, K. A., & France, A. A. (2016). Perceptual ratings of subgroups of ataxic dysarthria. *International Journal of Language & Communication Disorders, 51*(4), 430–441. https://doi.org/10.1111/1460-6984.12219

Spencer, K. A., & Slocomb, D. L. (2007). The neural basis of ataxic dysarthria. *Cerebellum, 6*(1), 58–65. https://doi.org/10.1080/14734220601145459

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

Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *The Journal of the Acoustical Society of America, 84*(3), 917–928. https://doi.org/10.1121/1.396660

Svirsky, M. A., Lane, H., Perkell, J. S., & Wozniak, J. (1992). Effects of short-term auditory deprivation on speech production in adult cochlear implant users. *The Journal of the Acoustical Society of America, 92*(3), 1284–1300. https://doi.org/10.1121/1.403923

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience, 5*(11), 1226–1235. https://doi.org/10.1038/nn963

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

Tremblay, S., Shiller, D., & Ostry, D. (2003). Somatosensory basis of speech production. *Nature, 423,* 866–869. https://doi.org/10.1038/nature01710

Trouillas, P., Takayanagi, T., Hallett, M., Currier, R. D., Subramony, S. H., Wessel, K., Bryer, A., Diener, H. C., Massaquoi, S., Gomez, C. M., Coutinho, P., Ben Hamida, M., Campanella, G., Filla, A., Schut, L., Timann, D., Honnorat, J., Nighoghossian, N., & Manyam, B. (1997). International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. The Ataxia Neuropharmacology Committee of the World Federation of Neurology. *Journal of the Neurological Sciences, 145*(2), 205–211. https://doi.org/10.1016/S0022-510X(96)00231-6

Tseng, Y.-W. W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology, 98*(1), 54–62. https://doi.org/10.1152/jn.00266.2007

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

Xu-Wilson, M., Chen-Harris, H., Zee, D. S., & Shadmehr, R. (2009). Cerebellar contributions to adaptive control of saccades in humans. *Journal of Neuroscience, 29*(41), 12930–12939. https://doi.org/10.1523/JNEUROSCI.3115-09.2009

Zimmet, A. M., Cao, D., Bastian, A. J., & Cowan, N. J. (2020). Cerebellar patients have intact feedback control that can be leveraged to improve reaching. *ELife, 9,* e53246. https://doi.org/10.7554/eLife.53246

Appendix

Participant Information for Individuals With Cerebellar Ataxia

These include ataxia type, speech severity scores from the International Cooperative Ataxia Rating Scale (ICARS) assessment (0–8), overall ataxia severity scores from the ICARS assessment (0–100), and results of the Montréal Cognitive Assessment (MoCA).

Age	Gender	SCA type	ICARS speech	ICARS total	MoCA
69	M	Unknown sporadic	2	20	29
59	F	Unknown sporadic	5	18	22
38	F	Unknown sporadic	2	36	26
54	F	Unknown sporadic	2	38	23
44	F	Unknown sporadic	1	21	26
55	F	Unknown sporadic	2	32	28
55	F	Unknown sporadic	3	50	21
39	F	Unknown sporadic	1	21	29
55	M	Unknown sporadic	1	24	24
68	M	SCA2	3	31	27
60	F	SCA2	1	31	26
59	F	SCA2	5	43	24
64	F	SCA6	1	12	26
69	M	SCA6	6	24	26
47	F	SCA14	2	47	20
49	F	SCA15	0	21	27

Note. SCA = spinocerebellar ataxia; M = male; F = female.