1 Relevance of Auditory Errors Decreases When Errors Are Introduced Suddenly

2
3 Sara-Ching Chao and Ayoub Daliri

4 College of Health Solutions, Arizona State University, Tempe, AZ, United States

5 Author Note

6 Correspondence should be addressed to Ayoub Daliri, Ph.D., College of Health Solutions,
7 Arizona State University, 975 S. Myrtle Ave., Coor Hall 3456, Tempe, AZ, 85287, USA. E-mail:
8 Ayoub.daliri@asu.edu. Phone: (480) 727-8397. Fax: (480) 965-8516.

9 This work was supported by a grant from the National Institutes of Health awarded to A. Daliri
10 (R21 DC017563).

11 Conflict of Interest: There are no relevant conflicts of interest.

Abstract

Purpose: When the speech motor system encounters errors, it generates adaptive responses to compensate for the errors. We previously showed that adaptive responses to task-irrelevant errors are significantly smaller than responses to task-relevant errors when errors are introduced gradually. The current study aimed to examine responses to task-irrelevant and task-relevant errors when errors are introduced suddenly.

Method: We used an adaptation paradigm in which participants experienced task-relevant errors (induced by formant-shift perturbations) and task-irrelevant errors (induced by formant-clamp perturbations). For one group of participants ($N = 30$), we applied the perturbations gradually. The second group of participants ($N = 30$) received the perturbations suddenly. We designed the perturbations based on participant-specific vowel configurations such that a participant's first and second formants of /ε/ were perturbed toward their /æ/. To estimate adaptive responses, we measured formant changes (within 0–100 ms of the vowel onset) in response to the formant perturbations.

Results: We found that (1) the difference between adaptive responses to formant-shift and formant-clamp perturbations was the smallest for the suddenly introduced perturbations, and (2) responses to formant-shift perturbations positively correlated with responses to formant-clamp perturbations for the suddenly (but not gradually) introduced perturbations.

Conclusions: These results showed that the speech motor system responds to task-relevant errors and task-irrelevant errors more differently when errors are introduced gradually than suddenly. Overall, the speech motor system evaluates the relevance of errors and uses its evaluation to modulate its adaptive responses to errors.

Keywords: speech, task-relevance, prediction error, adaptation

Relevance of Auditory Errors Decreases When Errors Are Introduced Suddenly

Speech monitoring plays a fundamental role in speech acquisition during childhood and ensures speech production accuracy during adulthood (Callan et al., 2000; Guenther, 2016). As the speech motor system (SMS) prepares speech movements, it also predicts somatosensory and auditory outcomes of the movements (Daliri, 2021; Daliri et al., 2014; Guenther, 2016; Hickok, 2012; Houde & Nagarajan, 2011; Parrell et al., 2019). The SMS then executes the planned movements and monitors the sensory outcomes of the movements to ensure that the movements are executed accurately. For this purpose, the SMS compares the predicted sensory outcomes of the movements with the actual sensory feedback from the movements. During this speech monitoring process, the SMS may find a mismatch between the predicted outcome and sensory feedback (i.e., *prediction error*). The SMS uses the sensory prediction error to modify (*adapt*) its motor programs so that future speech movements more accurately reach the desired sensory targets (Daliri, 2021; Kearney et al., 2020). Thus, error detection and error adaptation are crucial processes that ensure the accuracy of speech movements.

Prediction errors typically occur due to structural changes in speech articulators (e.g., rapid developmental changes during childhood or structural changes during the aging process) or sensory feedback (e.g., hearing impairment) (Daliri et al., 2013; Perkell, 2012; Perkell et al., 1992; Pittman et al., 2018; Vorperian et al., 2005, 2009). One could generate prediction errors in the laboratory setting by experimentally manipulating somatosensory feedback (Daliri et al., 2013; Lametti et al., 2012; Tremblay et al., 2003) or auditory feedback (Abur et al., 2018; Ballard et al., 2018; Daliri et al., 2017; Daliri & Max, 2018; Houde & Jordan, 1998; Kim et al., 2020; Kothare et al., 2020; Scott et al., 2020; Stepp et al., 2017). Due to practical reasons and the importance of auditory feedback for speech production, many studies have used auditory

perturbations to induce auditory prediction errors during speaking (for a review, see Fuchs et al., 2019). In auditory perturbation paradigms, acoustic characteristics of a speaker’s speech, such as formants, are shifted in near real-time and played back to the speaker (e.g., the speaker produces the word “bed” but hears a word that sounds like “bad”). The *formant-shift* procedure results in an experimentally generated auditory prediction error—a mismatch between the speaker’s expected auditory feedback (“bed”) and actual auditory feedback (“bad”). In response to repeated exposure to prediction errors, the speaker gradually adapts his/her motor commands to reduce the error. Examining adaptive responses provides essential insights into the error detection and error adaptation processes that the SMS uses to integrate prediction errors.

Previous studies have shown that characteristics of prediction errors contribute to the magnitude of adaptive responses (Daliri & Dittman, 2019; Kothare et al., 2020; MacDonald et al., 2010; Max & Maffett, 2015; Mitsuya et al., 2017). For example, adaptive responses to substantially large errors (MacDonald et al., 2010) or errors that are introduced with a delay (Max & Maffett, 2015; Mitsuya et al., 2017) are minimal or entirely absent. We have interpreted such findings as evidence that the SMS evaluates errors and responds less to less task-relevant errors (Daliri & Dittman, 2019). We recently developed a novel experimental procedure (i.e., *formant-clamp*) to directly manipulate the relevance of auditory errors. In the formant-clamp procedure, the induced error does not correspond with the speaker’s speech (i.e., task-irrelevant error), and the speaker cannot generate adaptive responses to reduce the experienced error. We found that adaptive responses to formant-clamp perturbations were significantly smaller than responses to formant-shift perturbations (Daliri & Dittman, 2019). In that study, we applied the perturbations gradually; thus, the SMS was exposed to small errors that incrementally increased over several trials, allowing the SMS to develop adaptive responses to the experienced errors.

Studies of limb motor learning suggest that different neural processes may prepare adaptive responses to gradually and suddenly introduced errors (Doya, 2000; Kim et al., 2020; Smith & Shadmehr, 2005; Venkatakrisnan et al., 2011; Werner et al., 2010). By examining adaptive responses to formant-shift perturbations, previous studies have shown that the SMS generates similar adaptive responses to gradually and suddenly introduced auditory errors (Kearney et al., 2020; Kim et al., 2020; MacDonald et al., 2010). However, it remained unclear how the SMS would evaluate and respond to suddenly introduced formant-clamp perturbations.

The primary goal of the present study is to determine how the SMS evaluates and responds to task-relevant vs. task-irrelevant errors that are introduced suddenly or gradually. We used an auditory-motor adaptation paradigm similar to our previous study (Daliri & Dittman, 2019), in which we applied formant-shift and formant-clamp perturbations. We used gradually introduced perturbations in our previous study (Daliri & Dittman, 2019); thus, we used suddenly introduced formant-shift and formant-clamp perturbations in the current study. We designed the perturbations based on participant-specific vowel configurations such that a participant's first and second formants of /ε/ were perturbed toward their /æ/. The magnitude of the perturbation was 80% of the participant-specific ε–æ distance. We measured formant changes within 0–100 ms from the vowel onset to estimate adaptive responses. We previously showed that these early responses are more appropriately suited to study adaptive changes (Daliri, 2021). We also re-analyzed the published data from participants in our previous study in which we applied gradual perturbations (Daliri & Dittman, 2019). Overall, we compared adaptive responses to suddenly introduced formant-shift and formant-clamp perturbations (data from the current study) with responses to gradually introduced formant-shift and formant-clamp perturbations (data from Daliri & Dittman, 2019). Suddenly introduced perturbations are more noticeable than gradually

introduced perturbations; thus, at least initially, the SMS may evaluate suddenly introduced errors as less task-relevant errors. We hypothesized that the SMS responds to task-relevant errors (induced by formant-shift perturbations) and task-irrelevant errors (induced by formant-clamp perturbations) more differently when errors are introduced gradually than suddenly.

Method

Participants

In this study, we included data from two groups of participants with similar participant characteristics: (1) a new group of participants and (2) a group of participants from our published study (Daliri & Dittman, 2019). The first group included 30 healthy adult participants (age: $M = 24.4$ years, $SD = 5.3$ years; age range: 20.7–45.8 years). Most participants were recruited from a pool of undergraduate students at Arizona State University. All participants met the following inclusion criteria: (1) being a native speaker of American English, (2) not having a current or a history of speech-language, psychological, and neurological disorders (self-reported), (3) not taking medications that influence the nervous system, and (4) having normal binaural hearing thresholds (≤ 20 dB HL at all octave frequencies of 250–8000 Hz) as tested with a standard pure-tone hearing screening (ASHA, 1997). The study protocols were approved by the Institutional Review Board of Arizona State University, and participants signed a written consent form before the experiment.

The second group included 30 healthy adult participants (age: $M = 23.7$ years, $SD = 3.7$ years; age range: 18.5–32.9 years). We have already published the data from these participants (Daliri & Dittman, 2019). However, we re-analyzed their data and included the analysis in the present study to compare adaptive responses to suddenly introduced (current study) with

gradually introduced (previous study) auditory perturbations. More information about these participants can be found in our previous study (Daliri & Dittman, 2019).

Apparatus

As shown in Figure 1A, we used the same experimental apparatus as our previous study (Daliri & Dittman, 2019). Participants sat comfortably in a chair in front of a computer monitor. We recorded speech using a microphone (Shure SM58), placed 15 cm away from the participant's mouth at a 45° angle. The microphone signal was amplified (ART Tubeopto 8) and fed into an audio interface (MOTU Ultralite Mk3 hybrid). This audio interface was connected to a computer that controlled all aspects of the experiment, such as presenting visual stimuli to the participant and manipulating auditory feedback. The audio interface's output was amplified (Eurorack Pro Rx1602) and binaurally played back to the participant via insert earphones (ER-1, Etymotic Research Inc.). The amplification level of the microphone and headphones amplifiers was adjusted so that the headphones signals were played back at 5 dB higher than the microphone signal (Daliri & Max, 2015, 2016; McGuffin et al., 2020; Merrikhi et al., 2018).

For auditory feedback manipulation, we used the Audapter software package (Cai, 2015). Audapter provides various routines to track and manipulate formant frequencies in near real-time. We used the following set of parameters for Audapter: sampling rate of 48 kHz, downsampling factor of 3, buffer size of 96 bytes, and a linear prediction model order of 17 for male participants and 15 for female participants (~18 ms input-output delay). In the present study, we used two types of formant perturbations: formant-shift and formant-clamp. In the formant-shift procedure, Audapter applies a relative shift to formants such that it conserves the overall formant trajectories (e.g., 10% increase in formants). In our previous study (Daliri & Dittman, 2019), we modified Audapter's source code to apply formant-clamp perturbations. In

the formant-clamp procedure, Audapter changes formants to user-defined formant values, regardless of the original formant values (e.g., the first formant would be shifted to 600 Hz regardless of the value of the formant). Because in the formant-shift perturbations, the shifted formants (that the participant hears in the earphones) are dependent on the participant's formants, the participant can change his/her formants to change what he/she hears in the earphones. However, in the formant-clamp perturbations, the formants that the participant hears in the earphones are independent of the participant's formants; thus, the participant cannot change his/her formants to change what he/she hears in the earphones. In other words, there is a close correspondence between the participant's formants and the formants he/she hears in the earphones in the formant-shift perturbations (i.e., correspondence between motor commands and auditory feedback). However, this is not the case for the formant-clamp perturbations. Overall, the auditory feedback in the formant-shift perturbation is task-relevant, whereas the auditory feedback in the formant-clamp perturbation is task-irrelevant.

Procedure

Participants completed three experimental tasks: one block of a practice task, one block of a pre-test task, and two blocks of an adaptation task. The entire experimental session lasted less than one hour. In the practice task, participants produced ten repetitions of consonant-vowel-consonant words ("bed," "Ted," and "head"). In each trial, one of the target words (in random order) appeared on the monitor for 2.5 s, followed by an intertrial interval of 1–2 s. Participants read the words aloud and received normal unperturbed auditory feedback. After each production, we provided visual feedback regarding the duration and intensity of the produced word. This task was designed to train participants to produce the target words within an intensity range of 70–80 dB SPL and a duration range of 400–600 ms.

The pre-test task was similar to the practice task, except that participants produced 30 repetitions of consonant-vowel-consonant words containing front vowels (/bɪd/, /bɛd/, /bæd/). Immediately after the completion of this task, we calculated the average first formant (F1) and second formant (F2) for each production (from the middle 20% of formant trajectories estimated by Audapter). We then calculated the centroid of each vowel and visually inspected the centroids. We used the distance and angle between the centroids of /ɛ/ and /æ/ (in the F1-F2 coordinates) to determine the perturbation magnitude and angle for each participant. We also used F1 and F2 of the centroid of /ɛ/ as initial values to improve Audapter's formant tracking algorithms.

As mentioned before, we included data from two groups of participants: a new group of participants and a group of participants from our published study (Daliri & Dittman, 2019). Each group of participants completed two blocks of the adaptation task. Each block of the adaptation task consisted of four phases in the following order: *baseline*, *change*, *hold*, and *end*. In each trial of the adaptation task, the participant produced one of the target words ("bed," "Ted," and "head") while receiving unperturbed or perturbed auditory feedback. In the baseline phase (30 trials), participants received unperturbed (normal) auditory feedback. In the change phase (30 trials), participants received perturbed auditory feedback that either started in the first trial of this phase (sudden) or gradually increased over the trials of this phase until it reached the maximum perturbation level (gradual). Participants from our previous study received gradual perturbations (Figure 1B), and the new group of participants received sudden perturbations (Figure 1C). The maximum perturbation was 80% of the participant-specific ϵ - α distance, such that F1 and F2 of /ɛ/ were shifted toward /æ/ (i.e., an increase in F1 and a decrease in F2). In the hold phase (60 trials), participants received the auditory perturbations with the maximum magnitude (80% of ϵ - α). In the end phase (30 trials), participants received unperturbed (normal) auditory feedback.

Each group of participants completed two blocks of the adaptation task. We used formant-shift perturbation in one block and formant-clamp perturbation in the second block. We counterbalanced the order of the blocks within each group of participants.

Data analysis

We used a custom-written MATLAB script to manually define the onset and offset of the vowel in each trial of the adaptation task. Trials that contained mispronunciations or formant tracking errors were excluded from the analysis. We extracted F1 and F2 trajectories starting from the onset time and ending at the offset time. Because the early portion of the vowel is primarily influenced by the feedforward control system (Daliri, 2021; Kearney et al., 2020), we averaged formant values within the first 100 ms of each vowel production to calculate the adaptive response. We designed the perturbations based on participant-specific vowel configurations such that the perturbations were parallel to the participant-specific ϵ - æ line (i.e., perturbation line). Similar to our previous studies (Chao et al., 2019; Daliri et al., 2020; Daliri & Dittman, 2019), we projected the average F1 and F2 of each trial to the ϵ - æ perturbation line (with the centroid of / ϵ / as the reference point). Positive projected responses were in the direction toward / æ /. Given that the target words were in the context of /C ϵ d/ (“bed,” “Ted,” and “head”), the starting formant transition (from the consonant to the vowel) of the three target words were different. Therefore, we corrected the projected responses based on the average responses in the baseline phase on a word-specific basis to minimize the effects of formant transitions on the responses (Daliri, 2021; Daliri et al., 2020). Finally, to directly compare adaptive responses across participants, we divided the responses by the participant-specific ϵ - æ distance. We calculated these adaptive responses for the adaptation tasks and entered them into the statistical analysis.

Statistical analysis

We performed all statistical analyses using R version 4.1.0 (R Core Team, 2021). As a dependent variable, we calculated the average adaptive responses based on (1) the 30 trials of the change phase, (2) the last 30 trials of the hold phase, (3) and the 30 trials of the end phase (see gray-shaded areas in Figure 2 A and B). The average responses in each phase and for each perturbation type and group are listed in Table 1. Because the responses were adjusted based on the average of the baseline trials, we did not enter the average responses of the baseline in the analysis. To statistically examine the adaptive responses, we used a linear mixed-effect model with perturbation group (sudden and gradual), perturbation type (formant-shift and formant-clamp), and phase (change, hold, and end phases) as fixed factors. We also entered participant as a random intercept in the linear mixed-effect model. We used the lmerTest package with Satterthwaite's method for approximating the degrees of freedom to determine the statistical significance of the main effects and their interactions (Kuznetsova et al., 2017). For post-hoc analysis of statistically significant interactions, we used the emmeans package and Bonferroni method to correct for multiple comparisons (Lenth, 2019). Additionally, to evaluate the relationship between adaptive responses to formant-shift and formant-clamp in different phases, we used the Psych package (Revelle, 2018) to calculate Pearson's correlation coefficients.

Results

Difference between adaptive responses to formant-shift and formant-clamp perturbations

The primary goal of this study was to examine whether the sudden and gradual introduction of perturbations can influence adaptive responses to task-relevant (formant-shift) and task-irrelevant (formant-clamp) errors. The normalized and baseline corrected adaptive responses for the gradual and sudden groups are shown in Figure 2 A and B, respectively. We

found statistically significant main effects of phase, $F(2, 1010) = 66.951, p < .001$, and perturbation type, $F(1, 1010) = 327.180, p < .001$; however, the main effect of perturbation group was not statistically significant, $F(1, 58) = 2.036, p = .159$. We also found phase $\times$ perturbation type interaction, $F(2, 1010) = 6.094, p = .002$, phase $\times$ perturbation group interaction, $F(2, 1010) = 21.558, p < .001$, and perturbation type $\times$ perturbation group interaction, $F(1, 1010) = 20.846, p < .001$. However, the phase $\times$ perturbation type $\times$ perturbation group interaction was not statistically significant, $F(2, 1010) = 0.679, p = .507$. The average responses in each phase and for each perturbation type and perturbation group are listed in Table 1. As shown in Figure 2 C and D, these interactions indicated that (1) the difference between adaptive responses to formant-shift and formant-clamp perturbations was the smallest when the perturbations were introduced suddenly ($p = .031$), (2) the difference between adaptive responses to formant-shift and formant-clamp perturbations was the smallest in the change phase ($p < .016$), and (3) regardless of the perturbation type, adaptive responses to gradually introduced perturbations were smaller than responses to suddenly introduced perturbations in the change phase ($p < .001$) but not in the hold and end phases ($p > .822$).

Correlation between adaptive responses to formant-shift and formant-clamp perturbations

We conducted a series of correlational analyses to determine the relationship between adaptive responses to formant-shift and formant-clamp perturbations in different phases (change, hold, and end) and for different perturbation groups (sudden and gradual). Figure 3 shows the results of these correlational analyses. For the suddenly introduced perturbations, we found positive correlations between adaptive responses to formant-shift and formant-clamp perturbations in all three phases (change phase: $r = .665, p < .001$; hold phase: $r = .691, p < .001$; end phase: $r = .425, p = .019$); however, this was not the case for the gradually introduced

perturbations. In other words, participants responded more similarly to formant-shift and formant-clamp perturbations when the perturbations were applied suddenly.

Discussion

The primary goal of the present study was to determine how the SMS evaluates and responds to task-relevant vs. task-irrelevant errors that are introduced suddenly and gradually. In a previous study (Daliri & Dittman, 2019), we developed an experimental procedure to directly manipulate the relevance of errors (i.e., formant-clamp). In the current study, we used an auditory-motor adaptation paradigm in which we applied formant-shift perturbations (task-relevant errors) or formant-clamp perturbations (task-irrelevant errors). We compared adaptive responses to suddenly introduced formant-shift and formant-clamp perturbations (data from the current study) with responses to gradually introduced formant-shift and formant-clamp perturbations (data from the previous study). Suddenly introduced perturbations are more noticeable than gradually introduced perturbations; thus, at least initially, the SMS may evaluate suddenly introduced errors as less task-relevant errors. We hypothesized that the SMS responds to task-relevant errors (induced by formant-shift perturbations) and task-irrelevant errors (induced by formant-clamp perturbations) more differently when the errors are introduced gradually.

We found that (1) the difference between adaptive responses to formant-shift and formant-clamp perturbations was the smallest when the perturbations were introduced suddenly, and (2) adaptive responses to formant-shift perturbations positively correlated with responses to formant-clamp perturbations for suddenly introduced perturbations but not gradually introduced perturbations. These results indicated that the SMS evaluates errors and uses its evaluations to determine its adaptive responses. Generally, suddenly introduced errors are more pronounced,

and the brain is more likely to assign suddenly introduced errors to external sources (Berniker & Körding, 2008; Krakauer et al., 2019; Shadmehr et al., 2010). Therefore, the SMS may evaluate suddenly introduced formant-shift perturbations as less task-relevant and integrate them like task-irrelevant errors induced by formant-clamp perturbations. While this interpretation can explain within-subject analysis (i.e., responses to formant-shift vs. formant-clamp perturbations in each group), it does not explain the results of between-subject analyses. Specifically, if the SMS evaluates suddenly introduced errors as less task-relevant, the responses to sudden perturbations should be smaller than responses to gradual perturbations. However, our between-subject analysis showed that responses to gradual perturbations were like sudden perturbations (except in the change phase) for formant-shift and formant-clamp perturbations. The results for gradually vs. suddenly introduced formant-shift perturbations are consistent with the results of previous studies (Kearney et al., 2020; Kim et al., 2020; MacDonald et al., 2010). These results may be related to the fact that the within-subject analysis is more sensitive to detect potential effects (for a review, see Charness et al., 2012) of the sudden and gradual introduction of perturbations. Future studies can examine this issue by designing experiments in which each participant experiences gradually and suddenly introduced formant-shift and formant-clamp perturbations. Overall, our results indicated that how errors are introduced modulates the SMS's evaluation of the relevance of the errors and the SMS's adaptive responses to the errors.

We also found that (1) adaptive responses to gradual perturbations were smaller than responses to sudden perturbations in the change phase, and (2) the difference between adaptive responses to formant-shift and formant-clamp perturbations was the smallest in the change phase. We attribute both results to the introduction of the perturbations. The adaptive responses to the gradual perturbations gradually increased throughout the change phase (30 trials) until

they reached a stable level at the end of the change phase. However, adaptive responses to the sudden perturbations increased rapidly, and after ~9 trials, they reached a relatively stable level. It should be noted that these patterns of responses were similar for both formant-shift and formant-clamp perturbations. Because we averaged responses in all 30 trials of the change phase, the overall responses were larger for the sudden perturbations than the gradual perturbations. Additionally, because the overall responses to formant-shift and formant-clamp perturbations are smaller in the change phase (compared with responses in other phases), the difference between these responses is also smaller. These results may indicate that, although the SMS evaluates formant-shift and formant-clamp perturbations differently, it uses similar procedures to respond to the errors. One limitation of the current study is that the maximum perturbation was 80% of the participant-specific ϵ - æ distance; therefore, it is unclear whether the perturbation magnitude differently modulates adaptive responses to formant-shift and formant-clamp perturbations. Future studies can determine the effects of perturbation magnitude by examining adaptive responses to formant-shift and formant-clamp perturbations with various magnitudes.

Our results have important theoretical and clinical implications. Most models of speech posit that the SMS calculates prediction errors and uses the errors to develop adaptive responses (Daliri, 2021; Hickok, 2012; Houde & Nagarajan, 2011; Kearney et al., 2020; Parrell et al., 2019). However, as our results indicated, the SMS evaluates the relevance of errors and uses its evaluation to modulate its adaptive responses to the experienced errors. Recently, we developed a state-space model to estimate the SMS's sensitivity to errors and predict the SMS's responses to various errors (Daliri, 2021). Like other models, our model used the calculated error in each trial of an adaptation task to generate a response (i.e., multiplying the calculated error by error sensitivity). Given the results of this study, we suggest that a future version of the model can

include an error evaluation stage in which the SMS assigns a weight to the error based on the relevance of the error (e.g., a weight between 0 and 1, with 1 being the highest level of task-relevance). Then, the adaptive responses can be calculated based on the experienced errors, the sensitivity to the errors, and the relevance of the errors. Our results may also have clinical implications. For example, we have previously shown that adaptive responses to auditory perturbations are significantly small (or nonexistent) for adults who stutter (Daliri et al., 2017; Daliri & Max, 2018; Kim et al., 2020; Max & Daliri, 2019). Previously, we interpreted the small adaptive responses of people who stutter as lower sensitivity to errors. However, based on the current results, one could speculate that adults who stutter may evaluate errors as less task-relevant (i.e., assign a lower weight to the experienced errors). Overall, our results can be used to modify current theories of speech production and theories of disorders of speech production.

In summary, we examined the SMS's adaptive responses to task-relevant vs. task-irrelevant errors when the errors are introduced gradually and suddenly. We used an experimental procedure to directly manipulate the relevance of errors (i.e., formant-clamp). Two groups of participants completed auditory-motor adaptation tasks; the first group received gradually introduced formant-shift and formant-clamp perturbations; the second group received suddenly introduced formant-shift and formant-clamp perturbations. We found that (1) the difference between adaptive responses to formant-shift and formant-clamp perturbations was the smallest when the perturbations were introduced suddenly, and (2) adaptive responses to formant-shift perturbations positively correlated with responses to formant-clamp perturbations for the sudden but not the gradual perturbations. These results indicated that the SMS evaluates the relevance of errors and uses its evaluation to modulate the magnitude of its adaptive responses to errors.

356

Acknowledgments

357

This work was supported by a grant from the National Institutes of Health awarded to A.

358

Daliri (R21 DC017563). We thank Kimiya Kasraeian and Sreelakshmi Gurralla for their

359

contributions to participant recruitment for this project.

360

References

- Abur, D., Lester-Smith, R. A., Daliri, A., Lupiani, A. A., Guenther, F. H., & Stepp, C. E. (2018). Sensorimotor adaptation of voice fundamental frequency in Parkinson's disease. *PLoS ONE*, 13(1), e0191839. <https://doi.org/10.1371/journal.pone.0191839>
- ASHA, A. S.-L.-H. A. (1997). *Guidelines for audiologic screening: Panel on audiologic assessment*. American Speech-Language-Hearing Association.
- Ballard, K. J., Halaki, M., Sowman, P. F., 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>
- Berniker, M., & Körding, K. P. (2008). Estimating the sources of motor errors for adaptation and generalization. *Nature Neuroscience*, 11(12), 1454–1461. <https://doi.org/10.1038/nn.2229>
- Cai, S. (2015). *Audapter*. https://github.com/shanqing-cai/audapter_matlab
- Callan, D. E., Kent, R. D., Guenther, F. H., & Vorperian, H. K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research*, 43(3), 721–736. <https://doi.org/10.1044/jslhr.4303.721>
- Chao, S., Ochoa, D., & Daliri, A. (2019). Production variability and categorical perception of vowels are strongly linked. *Frontiers in Human Neuroscience*, 13, 96. <https://doi.org/10.3389/FNHUM.2019.00096>
- Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81, 1–8. <https://doi.org/10.1016/j.jebo.2011.08.009>

- 384 Daliri, A. (2021). A Computational Model for Estimating the Speech Motor System's Sensitivity
385 to Auditory Prediction Errors. *Journal of Speech, Language, and Hearing Research*, 64(6),
386 1841–1854.
- 387 Daliri, A., Chao, S., & Fitzgerald, L. C. (2020). Compensatory responses to formant
388 perturbations proportionally decrease as perturbations increase. *Journal of Speech,*
389 *Language, and Hearing Research*, 63(10), 3392–3407.
390 https://doi.org/10.1044/2020_JSLHR-19-00422
- 391 Daliri, A., & Dittman, J. (2019). Successful auditory motor adaptation requires task-relevant
392 auditory errors. *Journal of Neurophysiology*, 122(2), 552–562.
393 <https://doi.org/10.1152/jn.00662.2018>
- 394 Daliri, A., & Max, L. (2015). Modulation of auditory processing during speech movement
395 planning is limited in adults who stutter. *Brain and Language*, 143, 59–68.
396 <https://doi.org/10.1016/j.bandl.2015.03.002>
- 397 Daliri, A., & Max, L. (2016). Modulation of auditory responses to speech vs. nonspeech stimuli
398 during speech movement planning. *Frontiers in Human Neuroscience*, 10, 1–9.
399 <https://doi.org/10.3389/fnhum.2016.00234>
- 400 Daliri, A., & Max, L. (2018). Stuttering adults' lack of pre-speech auditory modulation
401 normalizes when speaking with delayed auditory feedback. *Cortex*, 99, 55–68.
402 <https://doi.org/10.1016/j.cortex.2017.10.019>
- 403 Daliri, A., Prokopenko, R. A., Flanagan, J. R., & Max, L. (2014). Control and prediction
404 components of movement planning in stuttering versus nonstuttering adults. *Journal of*
405 *Speech, Language, and Hearing Research*, 57(6), 2131–2141.
406 https://doi.org/10.1044/2014_JSLHR-S-13-0333

- 407 Daliri, A., Prokopenko, R. A., & Max, L. (2013). Afferent and efferent aspects of mandibular
408 sensorimotor control in adults who stutter. *Journal of Speech, Language, and Hearing*
409 *Research*, 56(6), 1774–1778. [https://doi.org/10.1044/1092-4388\(2013/12-0134\)](https://doi.org/10.1044/1092-4388(2013/12-0134))
- 410 Daliri, A., Wieland, E. A., Cai, S., Guenther, F. H., & Chang, S.-E. (2017). Auditory-motor
411 adaptation is reduced in adults who stutter but not in children who stutter. *Developmental*
412 *Science*, 21(2), e12521. <https://doi.org/10.1111/desc.12521>
- 413 Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor
414 control. *Current Opinion in Neurobiology*, 10(6), 732–739. [https://doi.org/10.1016/S0959-](https://doi.org/10.1016/S0959-4388(00)00153-7)
415 4388(00)00153-7
- 416 Fuchs, S., Cleland, J., & Rochet-Capellan, A. (Eds.). (2019). *Speech production and perception:*
417 *Learning and memory*. Peter Lang. <https://doi.org/10.3726/b15982>
- 418 Guenther, F. H. (2016). *Neural control of speech*. MIT Press.
- 419 Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews*
420 *Neuroscience*, 13(2), 135–145. <https://doi.org/10.1038/nrn3158>
- 421 Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*,
422 279(5354), 1213–1216. <https://doi.org/10.1126/science.279.5354.1213>
- 423 Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers*
424 *in Human Neuroscience*, 5, 82. <https://doi.org/10.3389/fnhum.2011.00082>
- 425 Kearney, E., Nieto-Castañón, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., Ballard, K.
426 J., Chang, S.-E., Chao, S.-C., Murray, E. S. H., Scott, T. L., & Guenther, F. H. (2020). A
427 simple 3-parameter model for examining adaptation in speech and voice production.
428 *Frontiers in Psychology*, 10(2995), 1–16. <https://doi.org/10.3389/fpsyg.2019.02995>
- 429 Kim, K. S., Daliri, A., Flanagan, J. R., & Max, L. (2020). Dissociated development of speech and

- 430 limb sensorimotor learning in stuttering: Speech auditory-motor learning is impaired in both
431 children and adults who stutter. *Neuroscience*, 451, 1–21.
432 <https://doi.org/10.1016/j.neuroscience.2020.10.014>
- 433 Kothare, H., Raharjo, I., Ramanarayanan, V., Ranasinghe, K., Parrell, B., Johnson, K., Houde, J.
434 F., & Nagarajan, S. S. (2020). Sensorimotor adaptation of speech depends on the direction
435 of auditory feedback alteration. *The Journal of the Acoustical Society of America*, 148,
436 3682–3697. <https://doi.org/10.1121/10.0002876>
- 437 Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., & Haith, A. M. (2019). Motor learning.
438 *Comprehensive Physiology*, 9(2), 613–663. <https://doi.org/10.1002/cphy.c170043>
- 439 Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in
440 linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
441 <https://doi.org/doi:10.18637/jss.v082.i13>
- 442 Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production
443 revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of*
444 *Neuroscience*, 32(27), 9351–9358. <https://doi.org/10.1523/JNEUROSCI.0404-12.2012>
- 445 Lenth, R. (2019). *emmeans: Estimated marginal means* (1.3.3). R package. [https://cran.r-](https://cran.r-project.org/package=emmeans)
446 [project.org/package=emmeans](https://cran.r-project.org/package=emmeans)
- 447 MacDonald, E. N., Goldberg, R., & Munhall, K. G. (2010). Compensations in response to real-
448 time formant perturbations of different magnitudes. *The Journal of the Acoustical Society of*
449 *America*, 127(2), 1059–1068. <https://doi.org/10.1121/1.3278606>
- 450 Max, L., & Daliri, A. (2019). Limited pre-speech auditory modulation in individuals who stutter:
451 Data and hypotheses. *Journal of Speech, Language, and Hearing Research*, 62(8S), 3071–
452 3084. https://doi.org/10.1044/2019_JSLHR-S-CSMC7-18-0358

- 453 Max, L., & Maffett, D. G. (2015). Feedback delays eliminate auditory-motor learning in speech
454 production. *Neuroscience Letters*, 591, 25–29. <https://doi.org/10.1016/j.neulet.2015.02.012>
- 455 McGuffin, B. J., Liss, J. M., & Daliri, A. (2020). The orofacial somatosensory system is
456 modulated during speech planning and production. *Journal of Speech, Language, and*
457 *Hearing Research*, 63(8), 2637–2648.
- 458 Merrikhi, Y., Ebrahimpour, R., & Daliri, A. (2018). Perceptual manifestations of auditory
459 modulation during speech planning. *Experimental Brain Research*, 236(7), 1963–1969.
460 <https://doi.org/10.1007/s00221-018-5278-3>
- 461 Mitsuya, T., Munhall, K. G., & Purcell, D. W. (2017). Modulation of auditory-motor learning in
462 response to formant perturbation as a function of delayed auditory feedback. *The Journal of*
463 *the Acoustical Society of America*, 141(4), 2758–2767. <https://doi.org/10.1121/1.4981139>
- 464 Parrell, B., Ramanarayanan, V., Nagarajan, S. S., & Houde, J. F. (2019). The FACTS model of
465 speech motor control : fusing state estimation and task-based control. *PLoS Computational*
466 *Biology*, 15(9), e1007321. <https://doi.org/10.1371/journal.pcbi.1007321>
- 467 Perkell, J. S. (2012). Movement goals and feedback and feedforward control mechanisms in
468 speech production. *Journal of Neurolinguistics*, 25(5), 382–407.
469 <https://doi.org/10.1016/j.jneuroling.2010.02.011>
- 470 Perkell, J. S., Lane, H., Svirsky, M., & Webster, J. (1992). Speech of cochlear implant patients: A
471 longitudinal study of vowel production. *The Journal of the Acoustical Society of America*,
472 91(5), 2961–2978. <https://doi.org/10.1121/1.402932>
- 473 Pittman, A. L., Daliri, A., & Meadows, L. (2018). Vocal biomarkers of mild-to-moderate hearing
474 loss in children and adults: Voiceless sibilants. *Journal of Speech, Language, and Hearing*
475 *Research*, 61(11), 2814. https://doi.org/10.1044/2018_JSLHR-H-17-0460

- 476 R Core Team. (2021). *R: A language and environment for statistical computing*. (4.1.0). R
477 Foundation for Statistical Computing. <https://www.r-project.org/>
- 478 Revelle, W. (2018). *psych: Procedures for psychological, psychometric, and personality research*
479 (1.8.12). Northwestern University. <https://cran.r-project.org/package=psych>
- 480 Scott, T. L., Haenchen, L., Daliri, A., Chartove, J., Guenther, F. H., & Perrachione, T. K. (2020).
481 Noninvasive neurostimulation of left ventral motor cortex enhances sensorimotor adaptation
482 in speech production. *Brain and Language*, 209, 104840.
483 <https://doi.org/10.1016/j.bandl.2020.104840>
- 484 Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and
485 adaptation in motor control. *Annual Review of Neuroscience*, 33(1), 89–108.
- 486 Smith, M. A., & Shadmehr, R. (2005). Intact ability to learn internal models of arm dynamics in
487 Huntington’s disease but not cerebellar degeneration. *Journal of Neurophysiology*, 93(5),
488 2809–2821. <https://doi.org/10.1152/jn.00943.2004>
- 489 Stepp, C. E., Lester-Smith, R. A., Abur, D., Daliri, A., Noordzij, P. J., & Lupiani, A. A. (2017).
490 Evidence for auditory-motor impairment in individuals with hyperfunctional voice
491 disorders. *Journal of Speech, Language, and Hearing Research*, 60(6), 1545–1550.
492 https://doi.org/10.1044/2017_JSLHR-S-16-0282
- 493 Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production.
494 *Nature*, 423(6942), 866–869. <https://doi.org/10.1038/nature01710>
- 495 Venkatakrishnan, A., Banquet, J. P., Burnod, Y., & Contreras-vidal, J. L. (2011). Parkinson’s
496 disease differentially affects adaptation to gradual as compared to sudden visuomotor
497 distortions. *Human Movement Science*, 30(4), 760–769.
498 <https://doi.org/10.1016/J.HUMOV.2010.08.020>

- 499 Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S.
500 (2005). Development of vocal tract length during early childhood: A magnetic resonance
501 imaging study. *The Journal of the Acoustical Society of America*, 117(1), 338–350.
502 <https://doi.org/10.1121/1.1835958>
- 503 Vorperian, H. K., Wang, S., Chung, M. K., Schimek, E. M., Durtschi, R. B., Kent, R. D., Ziegert,
504 A. J., & Gentry, L. R. (2009). Anatomic development of the oral and pharyngeal portions of
505 the vocal tract: An imaging study. *The Journal of the Acoustical Society of America*, 125(3),
506 1666–1678. <https://doi.org/10.1121/1.3075589>
- 507 Werner, S., Bock, O., Gizewski, E. R., Schoch, B., & Timmann, D. (2010). Visuomotor adaptive
508 improvement and aftereffects are impaired differentially following cerebellar lesions in
509 SCA and PICA territory. *Experimental Brain Research*, 201, 429–439.
510 <https://doi.org/10.1007/s00221-009-2052-6>
511
512

513 **Tables**

514 *Table 1.* Average adaptive responses (and standard deviations inside the parentheses) to formant-
 515 shift and formant-clamp perturbations in different phases (change, hold, and end) for both
 516 perturbation groups (gradual and sudden).

517

	Gradual		Sudden	
	Formant-Shift	Formant-Clamp	Formant-Shift	Formant-Clamp
Change Phase	-9.70 (8.22)	0.83 (9.29)	-16.45 (12.21)	-10.30 (11.49)
Hold Phase	-25.35 (11.61)	-8.40 (12.31)	-22.01 (15.29)	-12.89 (14.38)
End Phase	-17.92 (12.46)	-2.31 (13.21)	-15.05 (12.13)	-4.59 (8.79)

518

519

Figures

Figure 1. The experimental setup for applying formant perturbations (A). Each participant completed two auditory-motor adaptation blocks in which he/she received formant-shift and formant-clamp perturbations. For one group of participants, the perturbations were introduced gradually (B), and for the second group of participants, the perturbations were introduced suddenly (C). In all cases, the perturbation shifted the participant's first formant (F1) and second formant (F2) of /ε/ toward /æ/ (increase in F1 and decrease in F2).

Figure 2. The group-average adaptive responses to formant-shift and formant clamp perturbations when perturbations were introduced gradually (A) and suddenly (B). Color-shaded areas in panels A and B correspond to the standard error of the mean. We calculated the average adaptive responses in the change, hold, and end phases (gray-shaded areas in panels A and B). Panel C shows the individual responses to different formant perturbation types (formant-shift and formant-clamp), perturbation groups (gradual and sudden), and in different phases (change, hold, and end). We found that the overall responses to formant-clamp perturbations were smaller than responses to formant-shift perturbations. However, this effect was modified by the phase factor and the group factor. The interactions indicated that the difference between adaptive responses to formant-shift and formant-clamp perturbations was the smallest when the perturbations were introduced suddenly ($p = .031$). Panel D shows the difference between responses to formant-shift and formant-clamp perturbations (responses to formant-shift perturbations minus responses to formant-clamp perturbations) for different phases and different groups. We also found that the difference between adaptive responses to formant-shift and formant-clamp perturbations was the smallest in the change phase ($p < .016$). Note that we published the data from participants of the

gradual group in a previous study (Daliri & Dittman, 2019); however, in this study, we re-analyzed the data to calculate adaptive responses within 0–100 ms of the vowel onset. Error bars in panels C and D correspond to the standard error of the mean.

Figure 3. We examined the relationship between adaptive responses to formant-shift and formant-clamp perturbations in different phases (change, hold, and end) and for different perturbation groups (sudden and gradual). We found positive correlations between responses to formant-shift and formant-clamp perturbations in all three phases of the suddenly introduced perturbations; however, this was not the case for the gradually introduced perturbations. The gray dashed line is the identity line (i.e., a line with the slope of 1).