JASA ARTICLE

Response patterns to vowel formant perturbations in children

Stephanie T. Cheung,[1,a)] Kristen Thompson,[1] Joyce L. Chen,[2,b)] Yana Yunusova,[3,c)] and Deryk S. Beal[1,d)]
[1]*Holland Bloorview Kids Rehabilitation Hospital, Toronto, Ontario, M4G 1R8, Canada*
[2]*Faculty of Kinesiology and Physical Education, University of Toronto, Toronto, Ontario, M5S 2W6, Canada*
[3]*Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario, M5G 1V7, Canada*

ABSTRACT:
Auditory feedback is an important component of speech motor control, but its precise role in developing speech is less understood. The role of auditory feedback in development was probed by perturbing the speech of children 4–9 years old. The vowel sound /ɛ/ was shifted to /æ/ in real time and presented to participants as their own auditory feedback. Analyses of the resultant formant magnitude changes in the participants' speech indicated that children compensated and adapted by adjusting their formants to oppose the perturbation. Older and younger children responded to perturbation differently in F_1 and F_2. The compensatory change in F_1 was greater for younger children, whereas the increase in F_2 was greater for older children. Adaptation aftereffects were observed in both groups. Exploratory directional analyses in the two-dimensional formant space indicated that older children responded more directly and less variably to the perturbation than younger children, shifting their vowels back toward the vowel sound /ɛ/ to oppose the perturbation. Findings support the hypothesis that auditory feedback integration continues to develop between the ages of 4 and 9 years old such that the differences in the adaptive and compensatory responses arise between younger and older children despite receiving the same auditory feedback perturbation.
 https://doi.org/10.1121/10.0006567

(Received 7 December 2020; revised 9 September 2021; accepted 17 September 2021; published online 12 October 2021)
[Editor: Anders Lofqvist] Pages: 2647–2654

I. INTRODUCTION

Sensory feedback monitoring is an integral component of the speech production system. One's own speech signal, in the form of auditory feedback, is known to be an important source of sensory information (Borden, 1979; Postma, 2000). In early development, for example, children with profound hearing loss, who lack the input of auditory feedback, demonstrate challenges in acquiring speech (Geffner, 1980; Sterne and Goswami, 2000). Auditory feedback provides ongoing information about the outcomes of articulatory movements so that errors can be immediately corrected, and stored associations between movement and sound can be adaptively maintained over time through motor learning (Borden, 1979; Postma, 2000). Here, we use the term *compensation* to refer to online, immediate changes that reduce or eliminate introduced errors; we use *adaptation* to refer to the motor control learning process informed by exposure to sensory feedback over time (Houde and Jordan, 2002).

The compensatory and adaptive roles of auditory feedback can be understood within computational models of speech such as the directions into velocities of articulators (DIVA) computational model (Guenther and Hickok, 2015; Tourville and Guenther, 2011). In this model, auditory, motor, and somatosensory targets are established on the activation of a speech sound map that contains stored associations between these targets (Guenther and Hickok, 2015; Tourville and Guenther, 2011). A feedforward controller executes motor commands to achieve the motor target, whereas auditory and somatosensory feedback are continuously processed to gauge the difference between the actual result and intended sensory target (Guenther and Hickok, 2015; Tourville and Guenther, 2011). Mismatches between the actual and intended speech sounds generate sensory error signals; compensatory corrective commands are generated to address the error. Auditory feedback also provides important information about the relation between articulation and speech output, allowing for feedforward commands to be adaptively calibrated over time (Guenther and Hickok, 2015; Tourville and Guenther, 2011).

There is theoretical speculation on the relative importance of feedforward and feedback loops and the particular role of auditory feedback in developing speech (Callan *et al.*, 2000; Feng *et al.*, 2011; Guenther and Hickok, 2016; Lametti *et al.*, 2012). Pitch perturbation studies have

[a)]Also at: Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada.
Electronic mail: stephaniet.cheung@mail.utoronto.ca, ORCID: 0000-0001-7006-3160.
[b)]Also at: Sunnybrook Health Sciences Centre, Toronto, Ontario, M4N 3M5, Canada.
[c)]Also at: Rehabilitation Sciences Institute, University of Toronto, Toronto, Ontario M5G 1V7, Canada; Toronto Rehabilitation Institute, University Health Network, Toronto, Canada; and Sunnybrook Research Institute, Toronto, Ontario M4N 3M5, Canada; ORCID: 0000-0002-2353-2275.
[d)]Also at: Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario, M5G 1V7, Canada; Rehabilitation Sciences Institute, University of Toronto, Toronto, Ontario, M5G 1V7, Canada; Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario, M5S 3G9, Canada; ORCID: 0000-0001-9488-4326.

revealed the ability of children to integrate such auditory information correctively for laryngeal control: In such studies, the voice fundamental frequency F_0 is increased or decreased while the participant speaks. Children aged 3–12 years old have been shown to compensate for pitch perturbations by subsequently shifting their voice fundamental frequencies correctively in the opposite direction (Liu *et al.*, 2010a; Liu *et al.*, 2013; Liu *et al.*, 2010b; Scheerer *et al.*, 2016; Scheerer *et al.*, 2013). In several studies, younger participants exhibited significantly longer response latencies; children's compensatory responses had later timing onsets than those of both young adults and adults above the age of 18 years old (Liu *et al.*, 2010a; Liu *et al.*, 2013; Liu *et al.*, 2010b; Scheerer *et al.*, 2013). This may indicate that the auditory feedback loop becomes more efficient with age as a result of improved synaptic connectivity, enabling the system to produce a faster corrective response (Scheerer *et al.*, 2013). However, comparatively little has been empirically established regarding the developmental trajectory of auditory feedback integration for articulatory control.

Vowel formant perturbation methods readily provide a means of investigating articulatory speech motor control in adults (Cai *et al.*, 2012; Cai *et al.*, 2011; Houde and Nagarajan, 2011; Purcell *et al.*, 2006; van den Bunt *et al.*, 2017) and children (Daliri *et al.*, 2018; MacDonald *et al.*, 2012; Terband and van Brenk, 2015). Using this approach, a speaker's vowel formants are shifted in real time so that they are newly associated with a different vowel sound. Thus, the integration of auditory feedback can be probed by examining the compensatory and adaptive changes to speech under perturbed conditions.

Vowel perturbation studies have predominantly focused on perturbations of the vowel /ɛ/, which allows for manipulations of vowels along the /i/ to /ɑ/ plane (e.g., Daliri *et al.*, 2018; Houde and Nagarajan, 2011; MacDonald *et al.*, 2012; Purcell *et al.*, 2006; Terband and van Brenk, 2015). The majority of vowel formant perturbation research has been performed in adult participants. When encountering perturbed vowels, adult speakers respond by modifying their formants to correctively oppose the manipulation applied (Cai *et al.*, 2008; Cai *et al.*, 2010; Villacorta *et al.*, 2007). The responses are thought to be automatic; no differences in the response are observed when speakers are told to attend to the manipulation or ignore it (Munhall *et al.*, 2009). However, a small number of speakers appears to follow rather than oppose the perturbation directions, enhancing the perturbation in a manner that would be unexpected in a negative feedback loop that should correct for the manipulation. After removal of the perturbation, formant aftereffects persist; this indicates that previous auditory feedback has been incorporated by speakers through adaptive motor learning. (Cai *et al.*, 2010; Munhall *et al.*, 2009; Villacorta *et al.*, 2007).

Only five published studies to date have investigated typically developing children's responses to online vowel perturbation in participants aged 2–4 (MacDonald *et al.*, 2012), 4–9 (Terband *et al.*, 2014; Terband and van Brenk, 2015; van Brenk and Terband, 2020), and 7 to 12 years old (Daliri *et al.*, 2018). The existing research has demonstrated that children exhibit greater compensatory vowel shifts in response to perturbation than adults (Daliri *et al.*, 2018; Terband and van Brenk, 2015; van Brenk and Terband, 2020), but that compensation does not occur before the age of 4 years old (MacDonald *et al.*, 2012). However, none of these studies was designed with the primary goal of probing the presence of developmental changes by examining the similarities and differences between the responses of younger and older children.

The goal of this study was to examine the role of auditory feedback in children's speech production to gain insight into the systems underlying the development of speech motor control. In particular, this study was designed to examine the role of auditory feedback in developing speech by comparing how younger and older children compensate and adapt to vowel perturbation. We hypothesized that if auditory feedback monitoring is fully developed in the early stages of speech development, younger and older children are likely to compensate and adapt in a similar manner. However, if auditory feedback monitoring is not fully mature in the early stages of speech development, younger children may show weaker or more variable responses than older children.

II. METHODS

The study protocols and procedures were reviewed and approved by the Research Ethics Board at Holland Bloorview Kids Rehabilitation Hospital.

A. Participants

Typically developing children were recruited through community postings at Holland Bloorview Kids Rehabilitation Hospital. We recruited children who were aged 4–9 years old inclusive. We chose to include participants in this age range because they have acquired all vowel and consonant sounds (McLeod and Crowe, 2018) but still undergo changes to the vocal tract anatomy (Vorperian and Kent, 2007) and articulation (Smith and Zelaznik, 2004), presenting an interesting opportunity to explore articulatory control. Exclusion criteria included a history of learning difficulties and head injury and speech or language problems. These criteria were determined through reports by parents. Additionally, children were required to pass a basic hearing screen at the time of the study appointment. To meet this criterion, children were required to have pure tone hearing thresholds below 20 dB hearing level (HL) at 1, 2, and 4 kHz. Twenty-six children were enrolled in the study. One child was withdrawn after a speech difficulty was detected, one child was withdrawn for not completing the hearing screen, and one child was withdrawn for not completing the study task. Thus, data from the remaining 23 participants were analyzed.

https://doi.org/10.1121/10.0006567

B. Protocol

Children spoke one of three words per trial when prompted by a picture with accompanying text on the computer screen. The three words were "head," "bed," and "Ted," all of which contain the target vowel /ɛ/. This vowel was selected because of its common use in previous vowel perturbation studies (Daliri *et al.*, 2018; Houde and Nagarajan, 2011; MacDonald *et al.*, 2012; Purcell *et al.*, 2006; Terband and van Brenk, 2015). One word was presented per trial in a randomized order. Children were familiarized with the task during a short familiarization phase ($\leq$ 15 trials) prior to the beginning of the study proper. During the study setup, the experimenter described the desired word associated with each prompt, then asked the child to say the word indicated by the prompt to validate the child's understanding. The experimenter did not proceed with the study until determining that the child understood the task.

The F_1 and F_2 changes applied by Audapter (Cai *et al.*, 2008) to the feedback speech signal during the experimental protocol are outlined in Fig. 1. During the start phase (15 trials), children performed the study task while listening to unaltered auditory feedback of their own voices. During the ramp phase (18 trials), perturbation of the auditory feedback was gradually increased until it reached a maximum of a 25% increase in F_1 and 12.5% decrease in F_2. The shift approximated the vowel /æ/ so that the shifted words sounded more like "had," "bad," and "tad." During the stay phase (18 trials), this maximum perturbation was held. In the last phase, the end phase (15 trials), the perturbation was removed and participants once again heard their unaltered voices. Audapter data for each trial, including F_1 and F_2 input and output traces, were automatically saved for analysis. The stay phase responses illustrated the participants' online compensation for the perturbation, whereas the end phase responses represented lingering adaptive effects of motor learning from the previous perturbed state.

Children were not informed that their speech would be perturbed, but were told that they would hear the sound of their own voices as they spoke. At the end of the study session, the nature of the experimental deception and its purpose were then revealed and debriefed with the participants.

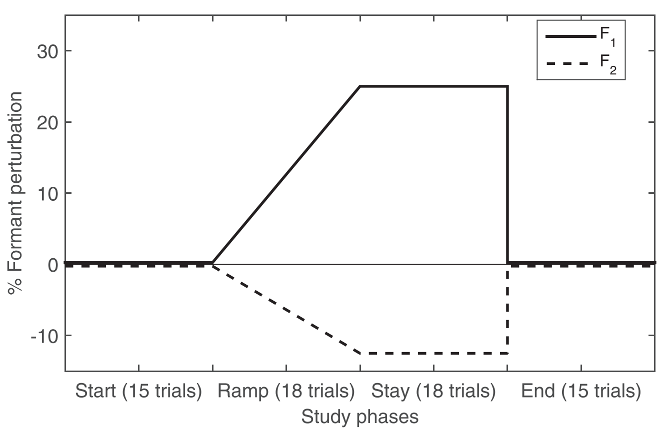

FIG. 1. The vowel formant perturbation study was conducted over four consecutive phases: start, ramp, stay, and end.

C. Equipment

Study sessions were conducted in a clinic sound booth (Controlled Acoustical Environments, Industrial Acoustics Company, Ontario, Canada) at Holland Bloorview Kids Rehabilitation Hospital. The auditory perturbation protocol was administered using Audapter software (Cai *et al.*, 2008), a MATLAB package designed for use in auditory perturbation research studies. Audapter manipulates audio signals in near real time by measuring the formant trace of the signal and applying a perturbing shift. The participants sat in front of a computer monitor on which trial prompts were visually displayed. The participants spoke into a mounted microphone (Apex515, Apex Electronics, Ontario, Canada). An external audio system (MOTU Microbook II, MOTU, Cambridge, MA) interfaced inputs and outputs with a Dell Latitude 3550 laptop (Intel CORE i7 CPU, Santa Clara, CA) on which the Audapter software was run. Audapter output audio signals were amplified using an external audio mixer (Xenyx 502, Behringer Corporation, Willich, Germany), then delivered to participants through over-the-ear headphones (HD 280 pro, Sennheiser Electronic Corporation, Wedemark, Hanover Region, Lower Saxony, Germany). Audapter latency has previously been described as near real time, operating at 14-ms delay (Cai *et al.*, 2008; Cai *et al.*, 2010).

Using built-in parameters in the Audapter study software, the experimenter adjusted the number of linear prediction coefficient filters during a setup phase of the study, examining spectrographic outputs for each participant's vocalizations to ensure that formants were properly tracked. Tracked formant values across the duration of the vowel (the "trace") were automatically saved for offline analysis.

We did not attempt to remove somatosensory feedback and bone-conducted feedback as it was not our aim to investigate the role of auditory feedback in isolation, and our interest was in probing the role of the auditory signal through manipulation without removing other forms of feedback.

D. Analysis

Audapter calculates formant traces from audio signals using linear predictive coding methods. For each trial, F_1 and F_2 were determined from a segment of the formant trace spanning from 40% to 60% of the Audapter output as in Daliri *et al.* (2018). This allowed us to extract steady-state portions of the vowel without capturing transitions between consonant and vowel sounds. Next, the formant traces were smoothed using a quadratic polynomial locally weighted smoothing function, implemented with the loess setting of the MATLAB function smooth. The outliers in each phase were removed using Tukey's rule so that extreme values exceeding 1.5 times the interquartile range were discarded. All outputs, including outlier outputs, were plotted on an F_1-

F_2 vowel space representation, and data were then visually inspected to ensure that the outliers were properly removed. In total, of 1518 trials recorded from the 23 participants, 165 trials were discarded as outliers (11%), which is comparable to the 8% of discarded trials previously reported in another auditory perturbation study with child participants (Daliri *et al.*, 2018). Of the outlier trials, 34 were from the start phase, 28 were from the ramp phase, 36 were from the stay phase, and 67 were from the end phase.

Mean F_1 and F_2 values were calculated separately for each phase and each participant, and analyses were performed on these data. As in Terband *et al.* (2014), only the first 12 end phase trials were used to calculate adaptive response to avoid capturing de-adaptation.

1. Formant magnitude analyses

The magnitude changes in F_1 and F_2 under perturbation were examined. For each participant, formant values in each trial were normalized against their baseline start phase mean. The mean and standard deviations (SDs) of each participant's formant values in each phase were then calculated. The mean formant ratios were derived by calculating the percentage changes in the mean formant values from the start phase baselines as in Terband *et al.* (2014), Terband and van Brenk (2015), and Daliri *et al.* (2018). Each participant's compensation data were examined to categorize the speaker as a nonresponder or a responder to the perturbation. The responders were defined as participants whose mean response formants for F_1, F_2, or both differed significantly from their baseline formants as determined by *t*-tests. Further, responses for those formants were required to exceed the estimation error as established in Cheung (2020; 60 Hz for F_1, 180 Hz for F_2) for participants to be classified as responders. The data from the responders were averaged to compute and plot the group mean responses. The nonresponder data were not included in the mean as we were interested in the magnitude of compensation to the perturbation and in that context, the magnitude of a nonresponse has no meaning. However large or small in magnitude, a response that does not compensate for the perturbation is a nonresponse. Our approach was to further classify the responders whose vowels followed rather than opposed perturbation as followers, but no such responses were observed.

As in Terband and van Brenk (2015), linear mixed-effects models were fit separately for F_1 and F_2 data. To fit the model, we used fixed effects of the age group (younger versus older, categorical variable), phase (start, stay, and end, categorical), and phase by group interactions. The random effect slopes were included in the model for phase and participant as well as for group and participant.

2. Exploratory paired directional analyses

To further investigate the results, we conducted supplemental exploratory analyses to visualize the paired F_1-F_2 vowel shift of the compensatory responses as a directional change relative to the F_1-F_2 vowel shift of the applied perturbation. The responses to perturbation can be interpreted directionally by representing the (F_1, F_2) pair as a point in two-dimensional formant space rather than as independent one-dimensional measures. An angular representation of the data can be attained by visualizing both the experimental manipulation of formants F_1 and F_2 for vowel perturbation as well as consequent F_1 and F_2 changes in the participants' speech as unit vectors on the F_1-F_2 plane. This method allows for speakers' associations between F_1 and F_2 to be retained and enables exploratory circular statistics to be applied to the study of directness patterns in angles of the compensatory and adaptive response.

The method used to apply circular analyses to compensatory and adaptive vowel changes under perturbation was as follows: Compensation was calculated by subtracting formants from the stay phase, during which perturbation was held at its maximum, from formants of the start phase, the baseline phase during which no perturbation was applied,

$$\Delta F_{\text{compensation}} = F_{\text{stay}} - F_{\text{start}}. \quad (1)$$

The perturbation was vectorized as

$$\vec{v}_{\text{perturbation}} = (F_{1\,\text{perturbation}}) + j(F_{2\,\text{perturbation}}). \quad (2)$$

An alignment step was performed to rotate all formant data on a participant-by-participant basis so that $\vec{v}_{\text{perturbation}}$ was aligned at 0 rad for each participant.

The unit vectors, representing directions of vowel change in response to perturbation, were plotted on a unit circle for exploratory visual analysis, where each participant's compensation vectors were rotated such that the perturbation directions across participants were aligned. This resulted in one compensation vector $\vec{v}_{\text{compensation}}$ for each participant. Following the methods of Berens *et al.* (2009) and using the MATLAB toolbox `CircStat`, unit vectors $\hat{\vec{v}}_{\text{compensation}}$ were plotted as open circle markers on unit circles for compensatory responses. As the nonresponses cannot be vectorized and visualized as directions, only the compensatory responders were included in the directional analyses of compensation. Each marker represented the compensation data from one participant. This visualization of the data is useful: The π rad angle represents the direction of perfect opposition to perturbation. The 0 rad angle represents the direction of perturbation itself. The angles $\pi/2$ and $3\pi/2$ represent the responses that are equally as distant to both the 0 and π rad angles, that is, equally as opposite to the perturbation angle as aligned with it.

III. RESULTS

The age and sex demographics for responders are reported in Table I. We aimed to study the presence of developmental differences in response to the vowel formant perturbation as an indicator of auditory feedback integration, therefore, examination of the group differences was of interest. To examine potential age effects on the response, analyses were performed with participants separated into younger

TABLE I. The age and sex demographic information for the participants.

	Group characteristics	
	Mean age (SD)	Sex
Age 4 ($N = 1$)	4.87 yr	1 F
Age 5 ($N = 4$)	5.40 (0.26) yr	3 F, 1 M
Age 6 ($N = 6$)	6.51 (0.32) yr	2 F, 4 M
Age 7 ($N = 2$)	7.68 (0.26) yr	2 F
Age 8 ($N = 4$)	8.64 (0.06) yr	1 F, 3 M
Age 9 ($N = 1$)	9.15 yr	1 F

(<7 years old) and older ($\geq$ 7 years old) groups. A linear effects model was applied to examine the main effects and interactions to compare the responses between the two age groups. It is important to note that this particular division of participants by age only served as an indicator of the presence of developmental differences between younger and older children within the range of 4–9 years of age. Our participant grouping is not designed to imply any timescale for developmental observations, particularly in such a small sample.

Under our criteria for categorizing compensatory responses to perturbation, 5 participants ($N = 1$ aged 5 years old, $N = 1$ aged 8 years old, $N = 3$ aged 9 years old, 22% of the total) were classified as nonresponders and 18 participants (78%) were classified as responders. Only the data from the responders are included in the analyses reported here. Thus, the included data are from 11 participants in our younger group (<7 years old) and 7 participants in our older group ($\geq$ 7 years old).

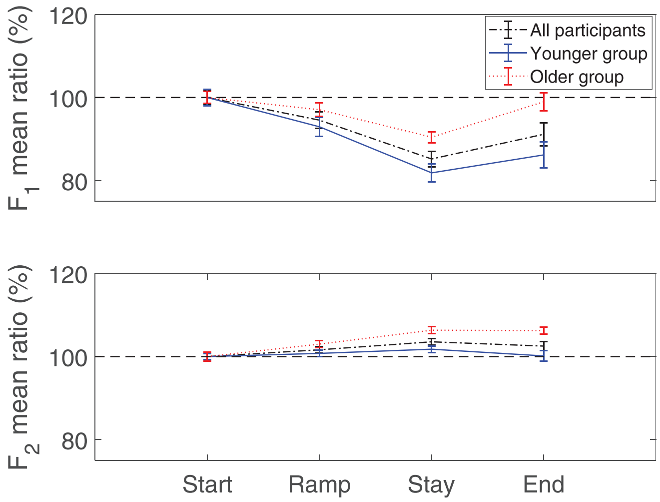

FIG. 2. (Color online) The group mean formant ratios indicated the stay phase compensation and end phase adaptive aftereffects in all participants as shown in a younger group ($\leq$ 7 years old, $N = 11$) and an older group ($\geq$ 7 years old, $N = 7$).

A. Formant magnitude analyses

The group mean formant ratios were obtained for all participants, for the younger group (<7 years old, $N = 11$), and the older group ($\geq$ 7 years old, $N = 7$), and plotted against the phase (Fig. 2). The detailed mean formant ratios and their SDs by phase are detailed in Table II. These results revealed that as groups, younger children and older children compensated for the perturbation with formant shifts that generally opposed the experimental manipulation.

1. Linear mixed effect modelling

For the F_1 ratio, a significant effect was found for the phase [$F(2,759) = 13.416$, $p < 0.01$]: The F_1 was lower under perturbation than at the start phase baseline. A significant effect was also found for the group [$F(2,759) = 11.18$, $p < 0.01$]: Children in the younger group exhibited larger F_1 changes under perturbation. The group by phase interaction for F_1 was also significant [$F(2,759) = 6.29$, $p < 0.01$]: Children in the younger group had larger F_1 changes under perturbation than did children in the older group.

For the F_2 ratio, a significant effect was found for the phase [$F(2,759) = 12.80$, $p < 0.01$]: F_2 was higher under perturbation than at the baseline. The effect of the group was also significant for F_2 [$F(2,759) = 12.05$, $p < 0.01$]: Older children exhibited larger F_2 changes under perturbation. The group by phase interaction was also significant for the F_2 ratio [$F(2,759) = 5.77$, $p < 0.01$]: Older children had larger F_2 changes under perturbation than younger children.

B. Exploratory paired directional analyses

Results from the exploratory directional analyses suggest that older children responded more directly to vowel perturbation than younger children. As π rad indicates an opposing response to perturbation or, in other words, a compensatory response that creates the opposite vowel shift to the perturbation applied, back toward the original vowel sound, it appears that older children generally responded in

TABLE II. The mean formant ratios across the four study phases.

		Mean formant ratios (%)			
		Start	Ramp	Stay	End
All participants	F_1	0.00 (SD = 1.77)	−5.45 (2.01)	−14.80 (1.86)	−8.84 (2.77)
	F_2	0.00 (0.86)	1.64 (0.79)	3.56 (0.81)	2.53 (1.13)
Younger group	F_1	0.00 (1.98)	−7.05 (2.27)	−18.14 (2.20)	−13.81 (3.14)
	F_2	0.00 (0.68)	0.77 (0.78)	1.78 (0.79)	0.16 (1.29)
Older group	F_1	0.00 (1.43)	−2.94 (1.61)	−9.55 (1.33)	−1.02 (2.20)
	F_2	0.00 (1.14)	3.00 (0.81)	6.35 (0.85)	6.25 (0.89)

a manner that opposed perturbation more directly than younger children. Older children's responses clustered more closely to π rad, whereas younger children's responses were more dispersed and variable:(Fig. 3). Older children's compensation vectors had a range of 0.52 rad (2.86–3.38 rad). For younger children, the spread of compensatory responses was larger with a range of 1.74 rad (3.03–4.77 rad).

IV. DISCUSSION

Our study examined the role of auditory feedback in developing speech by comparing how younger and older children compensate and adapt to vowel formant perturbation. Two hypotheses were presented for this study: If auditory feedback monitoring is fully developed in early speech development, younger and older children would likely compensate and adapt similarly. However, if auditory feedback monitoring has yet to mature in early speech development, younger children would show weaker or more variable responses compared to older children.

Our analysis indicated that children within the range of 4–9 years old compensated and adapted for perturbation. However, we observed differences in the compensatory and adaptive responses between younger and older children. We note that in compensatory and adaptive responses, younger children showed a greater change in the F_1 mean formant ratio as a result of perturbation than older children, but older children showed a greater change in the F_2 mean formant ratio than younger children. A similar age-by-formant interaction has been observed by Terband and van Brenk (2015), who found that children generally had larger vowel perturbation responses than adults for F_1 but not F_2. Although spectral analysis methods enable the two formants to be analyzed independently, it is important to note that the measures are closely associated in vowel production. In effect, the age group by formant interaction we observed appears to indicate that younger and older children's corrective compensations differed in where they lay within the vowel space.

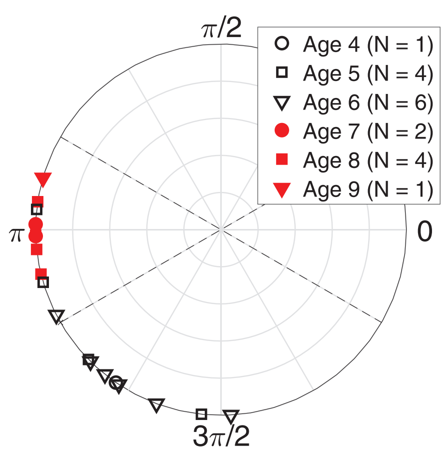

FIG. 3. (Color online) The circular visualization of the directions of the compensatory response to perturbation. The results are displayed on a unit circle.

Although preliminary and exploratory, paired directional analyses indeed indicated a directional difference in the compensatory response between younger and older children, where older children compensated in a manner that more directly opposed the shift in perturbation. The preliminary results suggest that directionality might be a useful way to understand the differences in children's responses to perturbation beyond magnitude formant changes. We note that to fully harness the advantages of circular statistical methods for these data, including inferential methods not explored in the present study to observe perturbation responses on a finer age gradation, a larger number of participants would be needed.

Participants had the task of integrating perturbed auditory feedback with unaltered somatosensory feedback; the differences in the integration of the two sensory sources of information might give rise to the differences in the response between younger and older children. Whereas the relative importance of auditory and somatosensory feedback in development is unclear, immaturity of the proper generation of corrective commands from one or both sensory feedback control loops in the DIVA model may hold an explanatory value for the indirect responses in our study. It could be that in younger children as compared to older children, the integration of altered auditory feedback and unaltered somatosensory feedback may be weighted differently. If younger participants relied more on unperturbed somatosensory feedback than did older children, it is reasonable to posit that they might have exhibited poorer corrective control for the perturbed auditory feedback, presenting in indirect responses to the auditory experimental manipulation. Modifications to the DIVA model of speech (Guenther and Vladusich, 2012; Tourville and Guenther, 2011) may be required to fully capture the developmental changes in childhood, perhaps with the inclusion of gain factors within the auditory and somatosensory feedback loops that are modulated in part by age. It might be that somatosensory feedback takes a higher priority in younger children than in older children or greater sensitivity to auditory information is achieved with age.

It is also possible that the sensory domains are not weighted differently by age, but auditory goals are simply better defined in older than in younger children. If somatosensory goals are developed first and auditory goals only later, compensation to auditory perturbation might be more precise in older children, resulting in a more direct corrective response in compensation compared to younger children. If directness of the compensatory response is indeed a useful indicator of the appropriateness of corrective compensation, the exploratory directional approach to analysis might allow investigators to futher compare how well (i.e., how directly) younger and older children compensate.

The role of the differences in motor skill development is also worth noting: Although consonants are acquired by the age of 5 years old (McLeod and Crowe, 2018), young children still undergo nonlinear changes to the vocal tract length and geometry well into late childhood (i.e., aged 7–12 years old; Vorperian and Kent, 2007), and motor coordination of vocal tract structures in children is more variable than in adults (Smith, 2010; Smith and Zelaznik, 2004). However, motor ability cannot be considered independently from the maturity of auditory feedback integration; a developmental change in the motor skill itself suggests an ongoing role for sensory feedback in the maintenance of speech motor mapping and precise auditory targets, one that might change over the course of development as motor skills mature.

A. Individual differences in nonresponse and following patterns

In our study, 22% of participants did not respond to auditory perturbation at all and were classified as nonresponders. This proportion of nonresponders in the overall sample is within the range of previously reported proportions of child participants displaying "inconsistent or insignificant" formant changes in another vowel formant perturbation study (Terband and van Brenk, 2015). It is unclear what nonresponse might signify. Lametti *et al.* (2012) have found that some adults seem to exhibit larger responses to either one of the auditory perturbation or somatosensory pertubation compared to the other. Perhaps such preferences are individual differences that underlie nonresponse. A link between developmental maturity and the ability to generate compensatory responses is also plausible but warrants further exploration: In a comparison of adult and child speech under perturbation, van Brenk and Terband (2020) found that fewer than 30% of adults exhibit no response compared to 30%–40% of children.

We observed no following responses among our participants. In one previous vowel formant perturbation study, 3 of 15 children (20%) followed the perturbation in F_1 and 5 of 15 children (33%) followed the perturbation in F_2 (Terband *et al.*, 2014). In yet another study, follower-type responses were observed in as many as 30% of children (van Brenk and Terband, 2020). Terband *et al.* (2014) have posited that the internal models that map the relations between motor command and sensory output might be "impaired" in these participants if phonemic targets are not well formed or if perturbation outputs are interpreted as updated speech targets rather than external feedback. This could result in a "target drift" on which perturbed vowels are taken as updated targets for speech rather than erroneous outcomes (Terband *et al.*, 2014). It is not immediately apparent why we would not have observed any following responses at all in our sample of 23 children.

We chose to include participants in this age range because they have acquired all vowel and consonant sounds (McLeod and Crowe, 2018) but still undergo changes to the vocal tract anatomy (Vorperian and Kent, 2007) and articulation (Smith and Zelaznik, 2004), presenting an interesting opportunity to explore articulatory control.

B. Limitations and future recommendations

It is important to note that the separation of participants into two age groups, one consisting of children aged 4–6 years old and the other consisting of children aged 7–9 years old, does not imply that differences in the response differentiate children between 6 and 7 years of age or any such differences are expected. This analytical method enabled a gross comparison of the response differences with developmental maturation by introducing a grouping factor of chronological age, but results should not be taken to suggest that the quantification of the difference is developmentally meaningful only that a difference appears to exist. Further research is required to probe the age-related differences observed in this study and understand the extent to which chronological age affects response. A larger number of participants would enable the regression analyses to be performed on the relation between age and response magnitude. Further, circular statistical methods might allow the relation between age and direction of response to be probed further: For example, circular-linear tests of correlation could identify relations between the circular variable of the response angle and linear variable of the chronological age.

V. CONCLUSIONS

This study was directed at investigating the role of auditory feedback integration in the speech of developing children. This was accomplished by comparing the compensatory and adaptive responses to vowel formant perturbation in younger and older children. Results largely indicated differences in the use of auditory feedback in younger children relative to older children, which could indicate that auditory feedback monitoring and integration continue to mature even within the age range of 4–9 years old. Future research, incorporating circular analytical methods, is needed to further explore the finding that younger and older children compensate differently for vowel formant perturbations. In particular, the nature of the group variability in directional responses requires additional evidence. A thorough understanding of the developing interplay between somatosensory feedback integration and auditory feedback integration would provide valuable context for the interpretation of findings arising from this research study.

ACKNOWLEDGMENTS

This paper contains excerpts from the doctoral dissertation by S.T.C. This research was funded by Wildcat Graduate Scholarships and the Eleanor Cate Allen Fellowship from the University of Toronto (held by S.T.C.), as well as the Kimel Family Graduate Student Scholarship from Holland Bloorview Kids Rehabilitation Hospital (held by S.T.C.). The contribution from K.T. was supported by the Ward family through a studentship at Holland Bloorview Kids Rehabilitation Hospital. The funding for the experimental study was supported by a Discovery Grant through the Natural Sciences and Engineering Research Council of Canada (held by D.S.B., Grant No. RGPIN-2015–05803).

Berens, P. (**2009**). "CircStat: A MATLAB toolbox for circular statistics," J. Stat. Softw. **31**(10), 1–21.
Borden, G. J. (**1979**). "An interpretation of research on feedback interruption in speech," Brain Lang. **7**(3), 307–319.
Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., and 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.
Cai, S., Boucek, M., Ghosh, S. S., Guenther, F. H., and Perkell, J. S. (**2008**). "A system for online dynamic perturbation of formant trajectories and results from perturbations of the Mandarin triphthong /iau/," in *Proceedings of the 8th International Seminar on Speech Production*, p. 4, available at https://issp2008.loria.fr/Proceedings/PDF/issp2008-10.pdf (Last viewed 2020-03-31).
Cai, S., Ghosh, S. S., Guenther, F. H., and Perkell, J. S. (**2010**). "Adaptive auditory feedback control of the production of formant trajectories in the Mandarin triphthong /iau/ and its pattern of generalization," J. Acoust. Soc. Am. **128**(4), 2033–2048.
Cai, S., Ghosh, S. S., Guenther, F. H., and Perkell, J. S. (**2011**). "Focal manipulations of formant trajectories reveal a role of auditory feedback in the online control of both within-syllable and between-syllable speech timing," J. Neurosci. **31**(45), 16483–16490.
Callan, D. E., Kent, R. D., Guenther, F. H., and 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," J. Speech, Lang. Hear. Res. **43**(3), 721–736.
Cheung, S. T. (**2020**). "Methodological considerations and findings of a vowel perturbation study in typically developing children," Ph.D. thesis, University of Toronto, available at https://tspace.library.utoronto.ca/handle/1807/100930 (Last viewed 2020-03-31).
Daliri, A., Wieland, E. A., Cai, S., Guenther, F. H., and Chang, S.-E. (**2018**). "Auditory-motor adaptation is reduced in adults who stutter but not in children who stutter," Dev. Sci. **21**, e12521.
Feng, Y., Gracco, V. L., and Max, L. (**2011**). "Integration of auditory and somatosensory error signals in the neural control of speech movements," J. Neurophysiol. **106**(2), 667–679.
Geffner, D. (**1980**). "Feature characteristics of spontaneous speech production in young deaf children," J. Commun. Disord. **13**(6), 443–454.
Guenther, F. H., and Hickok, G. (**2015**). "Chapter 9—Role of the auditory system in speech production," in *Handbook of Clinical Neurology*, edited by M. J. Aminoff, F. Boller, and D. F. Swaab (Elsevier, Netherlands), Vol. 129, pp. 161–175.
Guenther, F. H., and Hickok, G. (**2016**). "Chapter 58—Neural models of motor speech control," in *Neurobiology of Language*, edited by G. Hickok and S. L. Small (Academic, San Diego), pp. 725–740.
Guenther, F. H., and Vladusich, T. (**2012**). "A neural theory of speech acquisition and production," J. Neurolinguist. **25**(5), 408–422.
Houde, J. F., and Jordan, M. I. (**2002**). "Sensorimotor adaptation of speech I: Compensation and adaptation," J. Speech, Lang., Hear. Res. **45**(2), 295–310.
Houde, J. F., and Nagarajan, S. S. (**2011**). "Speech production as state feedback control," Front. Human Neurosci. **5**, 1–14.
Lametti, D. R., Nasir, S. M., and Ostry, D. J. (**2012**). "Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback," J. Neurosci. **32**(27), 9351–9358.
Liu, H., Russo, N. M., and Larson, C. R. (**2010a**). "Age-related differences in vocal responses to pitch feedback perturbations: A preliminary study," J. Acoust. Soc. Am. **127**(2), 1042–1046.
Liu, P., Chen, Z., Jones, J. A., Wang, E. Q., Chen, S., Huang, D., and Liu, H. (**2013**). "Developmental sex-specific change in auditory–vocal integration: ERP evidence in children," Clin. Neurophysiol. **124**(3), 503–513.
Liu, P., Chen, Z., Larson, C. R., Huang, D., and Liu, H. (**2010b**). "Auditory feedback control of voice fundamental frequency in school children," J. Acoust. Soc. Am. **128**(3), 1306–1312.
MacDonald, E. N., Johnson, E. K., Forsythe, J., Plante, P., and Munhall, K. G. (**2012**). "Children's development of self-regulation in speech production," Curr. Biol. **22**(2), 113–117.
McLeod, S., and Crowe, K. (**2018**). "Children's consonant acquisition in 27 languages: A cross-linguistic review," Am. J. Speech-Lang. Pathol. **27**(4), 1546–1571.
Munhall, K. G., MacDonald, E. N., Byrne, S. K., and Johnsrude, I. (**2009**). "Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate," J. Acoust. Soc. Am. **125**(1), 384–390.
Postma, A. (**2000**). "Detection of errors during speech production: A review of speech monitoring models," Cognition **77**(2), 97–132.
Purcell, D. W., and Munhall, K. G. (**2006**). "Compensation following real-time manipulation of formants in isolated vowels," J. Acoust. Soc. Am. **119**(4), 2288–2297.
Scheerer, N., Jacobson, D., and Jones, J. (**2016**). "Sensorimotor learning in children and adults: Exposure to frequency-altered auditory feedback during speech production," Neuroscience **314**, 106–115.
Scheerer, N. E., Liu, H., and Jones, J. A. (**2013**). "The developmental trajectory of vocal and event-related potential responses to frequency-altered auditory feedback," Eur. J. Neurosci. **38**(8), 3189–3200.
Smith, A. (**2010**). "Development of neural control of orofacial movements for speech," in *The Handbook of Phonetic Sciences* (Wiley, New Jersey), Chap. 7, pp. 251–296.
Smith, A., and Zelaznik, H. N. (**2004**). "Development of functional synergies for speech motor coordination in childhood and adolescence," Dev. Psychobiol. **45**(1), 22–33.
Sterne, A., and Goswami, U. (**2000**). "Phonological awareness of syllables, rhymes, and phonemes in deaf Children," J. Child Psychol. Psychiatry Allied Discip. **41**(5), 609–625.
Terband, H., van Brenk, F., and van Doornik-van der Zee, A. (**2014**). "Auditory feedback perturbation in children with developmental speech sound disorders," J. Commun. Disord. **51**, 64–77.
Terband, H. R., and van Brenk, F. J. (**2015**). "Compensatory and adaptive responses to real-time formant shifts in adults and children," in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, available at https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS1017.pdf (Last viewed 2020-03-31).
Tourville, J. A., and Guenther, F. H. (**2011**). "The DIVA model: A neural theory of speech acquisition and production," Lang. Cognit. Process. **26**(7), 952–981.
van Brenk, F., and Terband, H. (**2020**). "Compensatory and adaptive responses to real-time formant shifts in adults and children," J. Acoust. Soc. Am. **147**(4), 2261–2270.
van den Bunt, M. R., Groen, M. A., Ito, T., Francisco, A. A., Gracco, V. L., Pugh, K. R., and Verhoeven, L. (**2017**). "Increased response to altered auditory feedback in dyslexia: A weaker sensorimotor magnet implied in the phonological deficit," J. Speech, Lang. Hear. Res. **60**(3), 654–667.
Villacorta, V. M., Perkell, J. S., and Guenther, F. H. (**2007**). "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception," J. Acoust. Soc. Am. **122**(4), 2306–2319.
Vorperian, H. K., and Kent, R. D. (**2007**). "Vowel acoustic space development in children: A synthesis of acoustic and anatomic data," J. Speech, Lang. Hear. Res. **50**(6), 1510–1545.