Sensorimotor Control of Vocal Production in Early Childhood

Nichole E. Scheerer
Wilfrid Laurier University and Simon Fraser University,
Burnaby, British Columbia, Canada

Danielle S. Jacobson
Wilfrid Laurier University and University of Toronto

Jeffery A. Jones
Wilfrid Laurier University

Children maintain fluent speech despite dramatic changes to their articulators during development. Auditory feedback aids in the acquisition and maintenance of the sensorimotor mechanisms that underlie vocal motor control. MacDonald, Johnson, Forsythe, Plante, and Munhall (2012) reported that toddlers' speech motor control systems may "suppress" the influence of auditory feedback, since exposure to altered auditory feedback regarding their formant frequencies did not lead to modifications of their speech. This finding is not parsimonious with most theories of motor control. Here, we exposed toddlers to perturbations to the pitch of their auditory feedback as they vocalized. Toddlers compensated for the manipulations, producing significantly different responses to upward and downward perturbations. These data represent the first empirical demonstration that toddlers use auditory feedback for vocal motor control. Furthermore, our findings suggest toddlers are more sensitive to changes to the postural properties of their auditory feedback, such as fundamental frequency, relative to the phonemic properties, such as formant frequencies.

Keywords: speech production, speech development, sensorimotor control, vocal pitch, fundamental frequency

Throughout development, children maintain relatively fluent speech despite dramatic changes to the shape, size, and musculature of the articulators involved in speech production (Callan, Kent, Guenther, & Vorperian, 2000; Guenther, 1994). The vocal tract of a child is not simply a smaller version of an adult vocal tract, as there are many differences in the structures supporting

speech production in children, relative to adults (Callan et al., 2000; Civier, Tasko, & Guenther, 2010). In both humans and songbirds, auditory feedback has been suggested to not only aid in the acquisition of vocal motor control, but also act as an adaptive signal to guide movements of the articulators during development (Callan et al., 2000; Doupe & Kuhl, 1999; Guenther, 1994; Kelly & Sober, 2014; Perkell et al., 1997; Sakata & Brainard, 2008).

The importance of auditory feedback for monitoring and correcting ongoing vocalizations has been demonstrated by utilizing the frequency altered feedback (FAF) paradigm to synthetically alter speakers' auditory feedback (Burnett, Freedland, Larson, & Hain, 1998; Civier et al., 2010; Scheerer, Jacobson, & Jones, 2016; Scheerer & Jones, 2012, 2014, 2018a, 2018b; Scheerer, Liu, & Jones, 2013). When a speaker's auditory feedback is manipulated by changing properties such as the fundamental frequency (F0; Burnett et al., 1998; Civier et al., 2010; Scheerer et al., 2013, 2016; Scheerer & Jones, 2012, 2014, 2018a, 2018b), or the formant frequencies (Cai et al., 2012; Houde & Jordan, 1998; Purcell & Munhall, 2006; Villacorta, Perkell, & Guenther, 2007), the speaker reflexively responds by opposing the manipulation. These compensatory responses demonstrate that when a speaker detects changes in their auditory feedback, they use information from the deviant auditory feedback to modify their ongoing vocalization.

Scheerer and colleagues (2013) investigated the developmental trajectory of the speech motor control system by exposing children and adults to changes to their F0 using the FAF paradigm. The results of that study indicated that children as young as four years old produce compensatory responses to changes in their voice F0, consistent with the notion that auditory feedback is important for

monitoring and correcting for speech production errors. Similarly, Scheerer and colleagues (2016) found that children as young as three years old show similar rates of sensorimotor adaptation following changes to the F0 of their auditory feedback, relative to older adults. Surprisingly, MacDonald and colleagues (2012) reported that children and adults, but not toddlers, compensate for manipulations of the formant frequencies of their auditory feedback. The results of that study led to the suggestion that the auditory feedback component of the speech motor control system may be "suppressed" in toddlers. However, given the dramatic changes occurring to the acoustic properties of the vocal tract during development, it is hard to imagine how toddlers can preserve speech fluency without the aid of auditory feedback.

Models of speech motor control, such as the direction into the velocity of articulators model (DIVA), suggest that for prelinguistic children, random articulatory movements, or babbling, lead to the detection of novel sounds (Guenther, 2006; Guenther & Vladusich, 2012). These novel speech sounds activate neurons in the ventral premotor cortex and the posterior inferior frontal cortex, creating an auditory target that encodes the allowable variability in the acoustic signal for that novel sound (Guenther, 2006; Guenther & Vladusich, 2012; Tourville, Reilly, & Guenther, 2008). Once this neural representation for the auditory target is established, it can be activated for the production of the associated speech sound (Guenther, 2006). When active, this neural representation sends projections containing time-varying signals that encode the articulator velocities for the production of the auditory target to the articulatory control regions of the motor cortex. Projections containing time-varying auditory expectations associated with the activated neural representation are also sent to the auditory regions dedicated to the processing of the expected auditory feedback for the auditory target. At the same time, these auditory regions are receiving the auditory feedback resulting from the articulatory movements. Early on, these neural representations are not well tuned, thus the auditory expectation may not match the actual auditory feedback resulting from the articulatory movements. This difference between expected and actual auditory feedback constitutes an auditory error. This auditory error is then transformed into a corrective motor command that is sent to the articulatory control region of the motor cortex to modify the ongoing movement, and is also used to update the neural representation so that future productions of that sound are more accurate. Following repeated production of a speech sound, a somatosensory target for the sound is also learned. This somatosensory target encodes the expected tactile and proprioceptive sensations associated with the production of the sound, and allows somatosensory feedback-based error correction to occur in a similar manner to that described for auditory feedback. Thus, the DIVA model proposes that auditory feedback plays a fundamental role in the learning and maintenance of speech motor commands throughout development (Guenther, 2006; Guenther & Vladusich, 2012; Tourville et al., 2008).

Auditory feedback is important for maintaining both the phonemic and postural settings of speech, which ensure that phonemic distinctions are preserved, and that speech remains intelligible in dynamic acoustic environments, respectively (Perkell et al., 1997). F0, a property of auditory feedback related to postural control, has been shown to be more sensitive to the loss of auditory feedback, relative to formant frequency, a property of phonemic control (Perkell et al., 1997). For example, as reported by Perkell and

colleagues (1997), when there is a change in hearing status, as would happen when a cochlear implant is temporarily turned off, postural parameters like F0, average duration, and sound pressure level (SPL) change rapidly. On the other hand, phonemic setting, such as formant patterns, stay relatively stable. Since the phonemic and postural settings of speech appear to be regulated independently, it is also possible that these speech properties develop at different rates.

In the current study, we utilized the FAF paradigm to manipulate the F0 of toddlers' auditory feedback. Responses, or a lack of a response, to this manipulation will further our understanding of whether toddlers are capable of extracting information from their auditory feedback in order to modify their speech motor commands during ongoing speech, or whether the auditory feedback component of the speech motor control system is in fact "suppressed." Although previous research suggests that toddlers do not produce compensatory responses to manipulations of the formant frequencies of their auditory feedback (MacDonald et al., 2012), formant frequency and F0 control have been shown to be regulated independently (Perkell et al., 1997). For this reason, we expected that toddlers would produce compensatory responses to the F0 manipulations, demonstrating that auditory feedback plays a role in the ongoing monitoring and maintenance of the toddlers' speech motor control. Furthermore, coupled with the findings of MacDonald and colleagues (2012), these results would suggest that toddlers are more proficient at regulating the postural aspects of their speech, such as F0, relative to the phonemic aspects, such as formant frequencies.

Method

Participants

Twenty-five toddlers between the ages of 2 and 3 years ($M = 36.15$ months, $SD = 6.05$ months; 11 females) participated in this study. However, five of the toddlers did not complete the study, as they refused to produce the vocalizations. The remaining 20 toddlers were divided into two groups: younger toddlers who were between 24 and 35 months ($n = 11$, $M = 31.55$ months, $SD = 3.64$ months) and older toddlers who were between 40 and 46 months ($n = 9$, $M = 41.78$ months, $SD = 2.33$ months). We have previously investigated the effect of FAF on vocal responses in children and found that the difference in vocal responses across two different shift sizes yields an effect size of $n^2 = .55$ (Scheerer et al., 2013). Using G-Power (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007), it was determined that a sample size of $n = 13$ would provide 95% power to detect a difference across two shift magnitudes. Given the toddlers who participated in this study were younger, and we know that variability decreases with age (Scheerer et al., 2013), we boosted our sample to 20 to ensure we would maintain adequate power. All participants were Canadian-English speakers who did not speak a tonal language, and had no formal vocal training. Informed consent was obtained from all participants, as well as from a parent or guardian of each participant. All procedures were approved by the Wilfrid Laurier Research Ethics Board and were in accordance with the World Medical Association, 2013 Declaration of Helsinki.

Procedure

Participants were seated in front of a computer monitor that displayed an interactive farm game. As part of the game, a prerecorded voice produced a "baa" sound and the child was asked to mimic the "baa" sound in order to progress through the game. The prerecorded "baa'" sound played prior to each vocalization served as a reminder of how to produce the "baa" sound. The researcher also coached the toddler throughout the experiment, demonstrating how to hold the sound steady, and for an appropriate amount of time. This process was repeated until approximately 100 vocalizations were obtained.

Participants were asked to vocalize at a loud, but comfortable volume. Vocalizations were played back to the participants in real time via headphones. Since the participants were quite young, sound level was not precisely monitored. However, the experimenter provided online feedback throughout the experiment if the participant was vocalizing too loudly or too softly to help regulate the participants' volume. This volume monitoring technique was deemed adequate as previous research has shown that the relative loudness of auditory feedback has no influence on the amplitude or latency of responses to FAF (Burnett et al., 1998). Participants were also told that during the experiment their voice might sound different, but they should ignore these differences and continue to produce their vocalizations at a consistent pitch.

The experiment consisted of four blocks of 25 trials, and lasted approximately 20 min. During each vocalization the participant's voice was perturbed one semitone (100 cents) upward or downward three times. Each perturbation lasted for 200 ms, with an interstimulus interval of 1000–1200 ms (see Figure 1). The perturbation direction was held constant in each block, but alternated across blocks. The block orders were counterbalanced, so that half of the participants were initially exposed to the upward shift, while the other half were initially exposed to the downward shift.
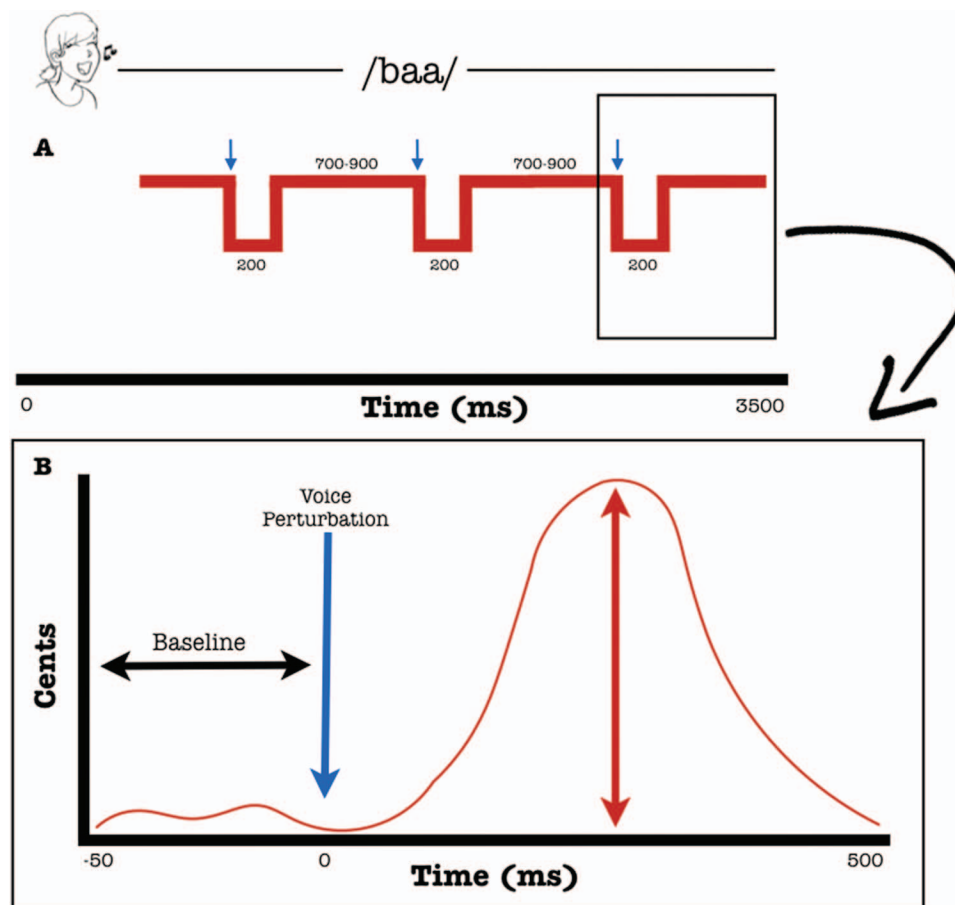
Figure 1. Experimental Paradigm. (A) Vocalization timeline: Participants were asked to vocalize a /baa/ sound for approximately 3.5 s while their F0 was perturbed either upward or downward one semitone (± 100 cents) three times per vocalization for 200 ms, with an interstimulus interval of 900–1100 ms. This figure represents an example of a downward perturbation trial. (B) Perturbation timeline: Each vocalization was segmented relative to the voice perturbation (blue line), with the 50 ms prior to the voice perturbation serving as a baseline to normalize the vocalization. The response magnitude was calculated by determining the maximum deviation from the baseline in the 500 ms following the voice perturbation (red line). This figure represents an example of a typical response to a downward perturbation. See the online article for the color version of this figure.

Apparatus

Participants were seated in front of a computer monitor and fitted with a headset-microphone (Sennheiser HMD 280–13 Pro, Sennheiser Electronics, Germany). Presentation of the visual stimuli were controlled by Powerpoint, while the onset and offset of the auditory feedback perturbations were controlled by Max/MSP (v.5.0, Cycling '74). During the experiment, vocalizations were sent to a Mac Mini (Apple, Cupertino, CA) where they were mixed using Studio One Software (Presonus, Baton Rouge, LA) and then sent to a digital signal processor (VoiceOne, T.C. Hellicon, Victoria, BC), which allow the pitch of the participant's voice to be altered. This process introduced approximately a 10-ms delay in the feedback signal, which was then presented back to the participant as auditory feedback. Both the altered and the unaltered voice signals were digitally recorded (Presonus FireStudio Project, Baton Rouge, LA) at a sampling rate of 44.01 kHz for later analysis.

Data Analysis

The digital recording of the vocalizations was segmented into separate utterances, and F0 values were calculated for each utterance. Utterances were then segmented on the basis of the pitch-shift onset. F0 values for each of the three segments were normalized to the baseline (50 ms of speech prior to the onset of the pitch shift) by converting hertz values to cents using the following formula:

$$\text{Cents} = 100(12 \log_2 F/B).$$

In this formula, F is the F0 value in hertz and B is the mean frequency of the baseline period. Cents values were calculated for 100 ms before and 500 ms after the perturbation. Since participant's vocalizations were much shorter than anticipated (less than 2 s), the majority of vocalizations only contained one perturbation. An averaged F0 trace was constructed for each shift magnitude (-100 cents, $+100$ cents), for each participant. The magnitude of each vocal response, the latency of the peak of the response, as well as participants' vocal variability (SD of the baseline period), were evaluated. The magnitude of each participant's compensatory response was determined by finding the point at which the averaged F0 trace deviated maximally from the baseline mean, the latency was calculated as the time at which this maximal deviation occurred, while the SD of the baseline period was calculated on the basis of the 50 ms of unaltered voice prior to the pitch shift.

The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Results

Averaged F0 traces were created for each participant and averaged across the younger toddlers and older toddlers (see Figure 2). Both the younger and older toddlers compensated for the auditory feedback perturbation by changing their vocal pitch in the opposite direction of the manipulation. To confirm these observations, response magnitudes were subjected to a repeated measures analysis of variance (RM-ANOVA) with perturbation direction (upward, downward) and age (younger toddler, older toddler) as factors. A main effect of perturbation direction, $F(1, 18) = 60.228$,
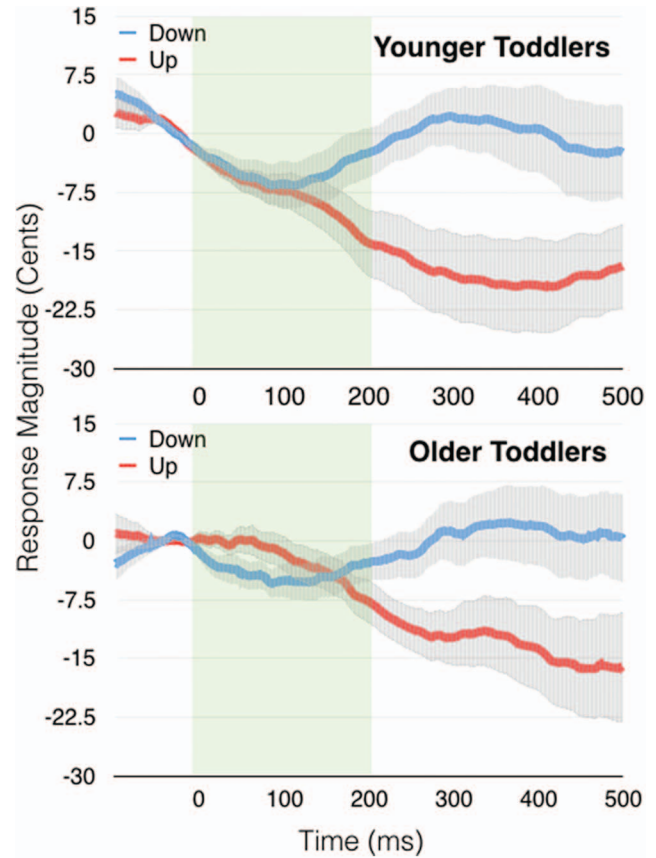

Figure 2. Averaged F0 Trace. Averaged F0 trace for both the upward perturbation (red) and downward perturbation (blue) conditions plotted separately for younger toddlers (top) and older toddlers (bottom). Error bars represent the standard error of the mean. The shaded portion (0–200 ms) indicates the perturbed portion of the trace. See the online article for the color version of this figure.

$p < .001$, $n^2 = .770$, confirmed that participants compensated for the perturbations by increasing their F0 in response to downward perturbations ($M = 10.59$ cents, $SD = 12.95$), and decreasing their F0 in response to upward perturbations ($M = -22.93$ cents, $SD = 17.13$; see Figure 3). However, the main effect of age, $F(1, 18) = .529$, $p = .479$, $n^2 = .029$, and the interaction between perturbation direction and age, $F(1, 18) = .659$, $p = .427$, $n^2 = .035$, were not significant.

A RM-ANOVA with perturbation direction (upward, downward) and age (younger toddler, older toddler) was also conducted on vocal response latencies. The main effect of perturbation direction, $F(1, 18) = .035$, $p = .854$, $n^2 = .002$, age, $F(1, 18) = .028$, $p = .870$, $n^2 = .002$, and the interaction between perturbation direction and age, $F(1, 18) = .059$, $p = .811$, $n^2 = .003$, were all not significant, as the timing of the peak response to the upward perturbation ($M = 354.35$, $SD = 121.48$) was similar to the timing of the peak response to the downward perturbation ($M = 343.40$, $SD = 154.77$), and the timing of the responses produced by the younger toddlers ($M = 346.07$, $SD = 134.79$) was similar to the timing of responses produced by the older toddlers ($M = 352.31$, $SD = 145.28$).
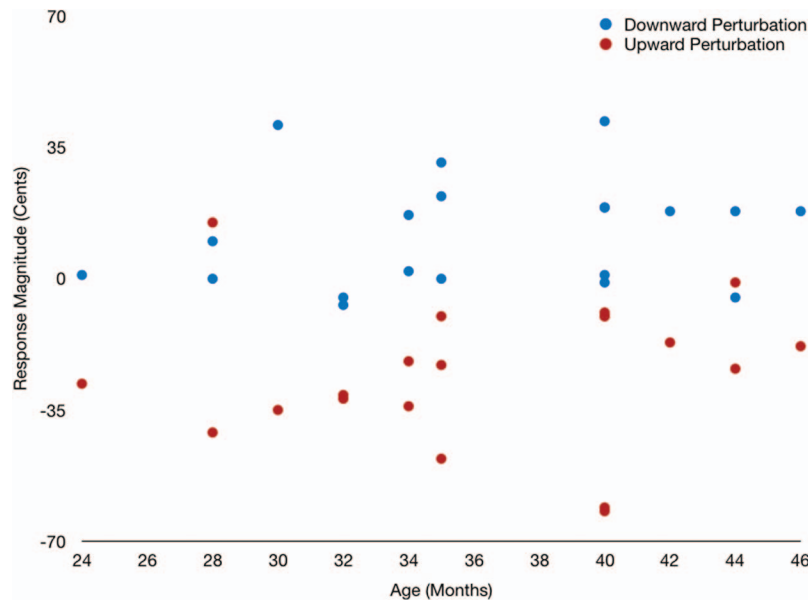
Figure 3. Compensatory Responses by Participant. Averaged F0 responses to the upward (red) and downward (blue) perturbations are plotted separately for each participant as a function of their age. See the online article for the color version of this figure.

Since a relationship between the age and vocal variability has previously been identified (Scheerer et al., 2013), a RM-ANOVA was also conducted to investigate the effect of perturbation direction (upward, downward) and age (younger toddler, older toddler) on vocal variability. This analysis revealed a main effect of age, $F(1, 18) = 4.60$, $p = .046$, $n^2 = .204$, as the younger toddlers ($M = 2.14$ cents, $SD = 1.19$) were more variable than the older toddlers ($M = 1.32$ cents, $SD = 0.99$). However, the main effect of perturbation direction, $F(1, 18) = 1.716$, $p = .207$, $n^2 = .087$, and the interaction between perturbation direction and age, $F(1, 18) = .277$, $p = .605$, $n^2 = .015$, were not significant.

Discussion

The aim of the current study was to investigate whether toddlers are capable of extracting information from their auditory feedback in order to regulate their speech motor control system during ongoing speech. The results indicate that toddlers produced compensatory responses to brief perturbations of the F0 of their auditory feedback. The responses produced following upward perturbations were significantly different than those produced following downward perturbations. These findings indicate that toddlers can in fact extract information from their auditory feedback to modify their ongoing vocalizations.

MacDonald and colleagues (2012) reported that children and adults compensate for alterations of the frequency of the first and second formants of their auditory feedback, but toddlers do not. Based on these findings, the authors concluded that the auditory feedback component of the speech motor control system may be suppressed in toddlers, or may not develop until closer to the age of four. The results of the current study argue against these conclusions, as they clearly demonstrate that toddlers under the age of four are actively using auditory feedback to monitor and

correct for errors in their ongoing speech. The current study recruited children of the same age range as MacDonald and colleagues (2012), who also spoke Canadian English. Similarly to MacDonald and colleagues (2012), this study also used an interactive game to elicit vocalizations from the toddlers. However, it is important to note that MacDonald and colleagues (2012) manipulated the formant frequencies of the toddlers' auditory feedback, while the current study manipulated the F0 of the toddlers' auditory feedback. Auditory feedback plays an important role in maintaining the phonemic settings of speech, in order to ensure phonemic distinctions are preserved (Perkell et al., 1997). However, auditory feedback is also important for maintaining the postural settings of speech, which ensure that speech remains intelligible in dynamic acoustic environments (Perkell et al., 1997). Formant frequency, a property of auditory feedback related to phonemic control, has been shown to be less sensitive to the loss of auditory feedback, relative to F0, a property of postural control (Perkell et al., 1997). This difference may explain why toddlers were able to modify their F0 using information from the deviant auditory feedback in the current study, but were unaffected by the formant frequency manipulations imposed by MacDonald and colleagues (2012). Aside from manipulating F0, rather than formant frequencies, all other experimental parameters were kept consistent across these two studies. Importantly, these differences suggest that toddlers may be more proficient at regulating the postural properties of their speech, relative to the phonemic settings, using auditory feedback.

Toddlers' increased proficiency at regulating the postural properties of speech, specifically F0, may also be a consequence of exposure to infant directed speech (IDS). Early in development, caregivers communicate with infants using IDS, and this IDS has been shown to provide an effective way to communicate affect

with prelinguistic children (Fernald, 1993; Saint-Georges et al., 2013). Since recognizing changes in F0 provides young children with a means for comprehending the affect and intentions of caregivers, prior to becoming linguistically proficient (Saint-Georges et al., 2013), IDS may increase the saliency of the prosodic features of speech such as F0. On the other hand, since nonprosodic features of the speech signal, such as formant frequencies, are less beneficial for speech comprehension in prelinguistic children (Saint-Georges et al., 2013), these features may be less salient. For this reason, we suggest that exposure to IDS may facilitate infants' ability to process prosodic aspects of speech, which later manifests itself as an increased proficiency at monitoring and correcting for prosodic changes in their own speech.

Considering the results of the current study, and those reported by MacDonald and colleagues (2012), it is reasonable to conclude that F0 and formant control develop at different rates. However, an alternative explanation may be that MacDonald and colleagues (2012) were unable to capture the compensatory responses being produced by the toddlers in their study as these responses were being masked by the variability in the toddlers' productions. The variability measure reported for the toddler group, standard deviation of the baseline, was larger in magnitude than the compensatory responses reported for both the adult and young children groups in the study conducted by MacDonald and colleagues (2012). On the other hand, when considering the results of the current study, vocal variability ($M = 1.73$) was only a fraction of the size of the average absolute compensatory response ($M = 16.76$). Thus, it is possible that these differences in vocal variability underlie the different experimental outcomes reported across these two studies.

Although the younger and older toddlers in this study produced similar sized compensatory responses to the FAF, younger toddlers were found to be more variable than the older toddlers. This finding is in line with previous studies that have demonstrated that vocal variability decreases with age (Scheerer et al., 2013). Scheerer and colleagues (Scheerer et al., 2013; Scheerer & Jones, 2012, 2014) have suggested that vocal variability is related to the proficiency of speech motor control. As vocal stability improves through vocal training or development, speakers become less dependent on auditory feedback for closed-loop speech motor control (Scheerer et al., 2013; Scheerer & Jones, 2012, 2014). Instead, speakers transition to a predominantly open-loop speech motor control system, where speech motor commands are generated from stored representations (Chen et al., 2013; Doupe & Kuhl, 1999; Guenther, 2006; Guenther & Vladusich, 2012; Max, Guenther, Gracco, Ghosh, & Wallace, 2004; Sakata & Brainard, 2008; Scheerer & Jones, 2012; Tourville et al., 2008). The decreased variability with age observed in this study may reflect a gradual maturation of the speech motor control system throughout development.

Although the results of this study demonstrate young children are capable of using auditory feedback to regulate their ongoing vocalizations, it is important to acknowledge some limitations of the study as well as some cautions when interpreting these results. Since we anticipated that it would be difficult to solicit these vocalizations from young children, we opted to collect responses to upward and downward perturbations, without collecting data during a baseline (or unaltered) vocalization condition, in order to minimize the number of vocalizations produced by the toddlers.

We took this approach knowing that different responses to upward and downward perturbations would be evidence that the direction of the manipulation was influencing the direction of the response, thus auditory feedback was being used to regulate ongoing vocalizations. That being said, the downward slope observed across both the upward and downward responses was unexpected, and without a baseline condition it is difficult for us to explain this slope. It is possible that this slope may be related to the short duration of the vocalizations (generally less than 2 s). While we intended to collect much longer vocalizations, with three perturbations per vocalization, this task proved to be too difficult for most toddlers, and for the majority of vocalizations (72%) only one perturbation occurred. While this downward slope in the compensatory responses makes it difficult to interpret the size of the compensatory responses, particularly in the upward shift condition, latency values indicate that responses peaked at approximately 350 ms. Since this response peak is not at the end of the vocalization, we can be confident that these data represent a compensatory response, and not just a downward trend in the data. Furthermore, the large variability in the response latency, a standard deviation of almost 140 ms, will have the effect of attenuating peaks in the average response, thus making this average waveform less representative of individual responses. Future work will be geared toward understanding the specific properties of this compensatory response in toddlers.

In summary, the current results demonstrate that toddlers as young as two years old produce compensatory responses to changes in their auditory feedback. These results provide empirical support for the long-held assumption that auditory feedback functions as an adaptive signal to guide movements of the articulators in order to preserve speech fluency while the articulators are restructured during development. In addition, these results suggest that toddlers may be more sensitive to changes to the postural properties of their auditory feedback, such as fundamental frequency, relative to the phonemic properties, such as formant frequencies.

References

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America, 103,* 3153–3161. http://dx.doi.org/10.1121/1.423073

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., & Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: Evidence for abnormal auditory-motor transformation. *PLoS ONE, 7,* e41830. http://dx.doi.org/10.1371/journal.pone.0041830

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,* 721–736. http://dx.doi.org/10.1044/jslhr.4303.721

Chen, Z., Jones, J. A., Liu, P., Li, W., Huang, D., & Liu, H. (2013). Dynamics of vocalization-induced modulation of auditory cortical activity at midutterance. *PLoS ONE, 8,* e60039. http://dx.doi.org/10.1371/journal.pone.0060039

Civier, O., Tasko, S. M., & Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: Simulations of stuttering and fluency-inducing conditions with a neural model of speech

production. *Journal of Fluency Disorders, 35,* 246–279. http://dx.doi.org/10.1016/j.jfludis.2010.05.002

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience, 22,* 567–631. http://dx.doi.org/10.1146/annurev.neuro.22.1.567

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149–1160. http://dx.doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64,* 657–674. http://dx.doi.org/10.2307/1131209

Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics, 72,* 43–53. http://dx.doi.org/10.1007/BF00206237

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39,* 350–365. http://dx.doi.org/10.1016/j.jcomdis.2006.06.013

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics, 25,* 408–422. http://dx.doi.org/10.1016/j.jneuroling.2009.08.006

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

Kelly, C. W., & Sober, S. J. (2014). A simple computational principle predicts vocal adaptation dynamics across age and error size. *Frontiers in Integrative Neuroscience, 8,* 75. http://dx.doi.org/10.3389/fnint.2014.00075

MacDonald, E. N., Johnson, E. K., Forsythe, J., Plante, P., & Munhall, K. G. (2012). Children's development of self-regulation in speech production. *Current Biology, 22,* 113–117. http://dx.doi.org/10.1016/j.cub.2011.11.052

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders, 31,* 105–122. http://dx.doi.org/10.1044/cicsd_31_S_105

Perkell, J., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22,* 227–250. http://dx.doi.org/10.1016/S0167-6393(97)00026-5

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America, 120,* 966–977. http://dx.doi.org/10.1121/1.2217714

Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., . . . Cohen, D. (2013). Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *PLoS ONE, 8*(10), e78103. http://dx.doi.org/10.1371/journal.pone.0078103

Sakata, J. T., & Brainard, M. S. (2008). Online contributions of auditory feedback to neural activity in avian song control circuitry. *The Journal of Neuroscience, 28,* 11378–11390. http://dx.doi.org/10.1523/JNEUROSCI.3254-08.2008

Scheerer, N. E., Jacobson, D. S., & Jones, J. A. (2016). Sensorimotor learning in children and adults: Exposure to frequency-altered auditory feedback during speech production. *Neuroscience, 314,* 106–115. http://dx.doi.org/10.1016/j.neuroscience.2015.11.037

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

Scheerer, N. E., & Jones, J. A. (2014). The predictability of frequency-altered auditory feedback changes the weighting of feedback and feed-forward input for speech motor control. *European Journal of Neuroscience, 40,* 3793–3806. http://dx.doi.org/10.1111/ejn.12734

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

Scheerer, N. E., & Jones, J. A. (2018b). The role of auditory feedback at vocalization onset and mid-utterance. *Frontiers in Psychology, 9,* 2019. http://dx.doi.org/10.3389/fpsyg.2018.02019

Scheerer, N. E., Liu, H., & Jones, J. A. (2013). The developmental trajectory of vocal and event-related potential responses to frequency-altered auditory feedback. *European Journal of Neuroscience, 38,* 3189–3200. http://dx.doi.org/10.1111/ejn.12301

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage, 39,* 1429–1443. http://dx.doi.org/10.1016/j.neuroimage.2007.09.054

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America, 122,* 2306–2319. http://dx.doi.org/10.1121/1.2773966