Published in final edited form as:

J Exp Psychol Hum Percept Perform. 2014 August ; 40(4): 1308–1315. doi:10.1037/a0036660.

Auditory-Perceptual Learning Improves Speech Motor Adaptation in Children

Douglas M. Shiller^{1,2,3} and Marie-Lyne Rochon¹

¹Research Center, Sainte-Justine Hospital, Université de Montréal, Montreal, QC, Canada

²School of Speech-Language Pathology and Audiology, Université de Montréal, Montreal, QC, Canada

³Centre for Research on Brain, Language & Music, McGill University, Montreal, QC, Canada

Abstract

Auditory feedback plays an important role in children's speech development by providing the child with information about speech outcomes that is used to learn and fine-tune speech motor plans. The use of auditory feedback in speech motor learning has been extensively studied in adults by examining oral motor responses to manipulations of auditory feedback during speech production. Children are also capable of adapting speech motor patterns to perceived changes in auditory feedback, however it is not known whether their capacity for motor learning is limited by immature auditory-perceptual abilities. Here, the link between speech perceptual ability and the capacity for motor learning was explored in two groups of 5–7-year-old children who underwent a period of auditory perceptual training followed by tests of speech motor adaptation to altered auditory feedback. One group received perceptual training on a speech acoustic property relevant to the motor task while a control group received perceptual training on an irrelevant speech contrast. Learned perceptual improvements led to an enhancement in speech motor adaptation (proportional to the perceptual change) only for the experimental group. The results indicate that children's ability to perceive relevant speech acoustic properties has a direct influence on their capacity for sensory-based speech motor adaptation.

Keywords

speech perception; speech production; speech development; sensorimotor learning

Introduction

It has long been known that hearing ability is essential for speech development, as indicated by the significant impact of hearing impairment on the quality and quantity of early speech output (Eilers & Oller, 1994; Eisenberg, 2007; Kent et al., 1987; McGowan et al., 2008; Oller & Eilers, 1988; Rvachew et al., 1999). Auditory input related to a child's own speech

production (i.e., feedback) is also believed to play a critical role (Mowrer, 1952). In current models of speech acquisition, auditory feedback contributes to speech development by providing the child with information about speech outcomes that is used to learn and subsequently fine-tune processes for speech motor planning (Callan et al., 2000; Perkell et al., 1997). It remains unknown, however, whether children's immature speech perceptual abilities limit their ability to use auditory feedback for such motor learning processes. In other domains of voluntary motor control, such as visually guided limb motion, immaturity in sensory representations has been linked with a limited capacity in young children (4–6 years of age) to adapt their motor patterns to changes in sensory feedback (Contreras-Vidal et al., 2004). Very little is known, however, about whether limitations in the auditory perception of speech sounds similarly constrain children's capacity for speech-motor learning. Understanding the link between children's auditory perceptual and speech motor function may not only help us to better understand the role of auditory feedback in children's speech motor development, but also to better grasp the factors underlying the enormous variability in speech production ability observed among children.

The idea that auditory feedback drives changes in speech motor control has been extensively studied in adults by evaluating adaptive changes in speech motor planning in response to systematic, predictable manipulations of auditory feedback during speech production. In these studies, partial or complete speech motor adaptation has been observed using a range of acoustic manipulations, including fundamental frequency (Jones & Munhall, 2000, 2003), vowel formant frequency (Cai et al., 2010; Houde & Jordan, 1998, 2002; Purcell & Munhall, 2006; Rochet-Capellan & Ostry, 2011; Villacorta et al., 2007), and fricative spectral properties (Shiller et al., 2009). Very few studies of auditory-motor adaptation in speech have been carried out in children. Shiller et al. (2010) demonstrated successful speech motor adaptation in a group of 9–11-year-olds to a manipulation of auditory feedback during the production of the fricative /s/. More recently, it has been demonstrated that children as young as 4 years of age will adapt speech motor patterns to perceived changes in vowel auditory feedback to maintain the accuracy of their speech output (MacDonald et al., 2012). These studies support a role for auditory feedback in children's speech motor function. Little is known, however, about the efficiency of such auditory-motor processes in children. In particular, speech adaptation to changes in auditory feedback is likely predicated upon the availability of accurate sensory information about acoustic speech outcomes. There is evidence from a single study in adult talkers involving vowel production that naturally occurring differences in auditory acuity are correlated with the degree of speech adaptation to altered auditory feedback (Villacorta et al., 2007). It remains unknown whether children's immature auditory perceptual abilities similarly influence their capacity to effectively use auditory feedback in order to improve speech motor outcomes.

A limitation of correlational methods such as those used by Villacorta et al. (2007) is that the direction of influence remains unclear. That is, it remains unknown whether differences in perceptual ability in fact constrain speech motor adaptation performance, or whether limitations in the capacity for auditory-motor learning have constrained speech perceptual development, for example by limiting the production of speech sound contrasts. In the present study, we hoped to overcome this limitation by exploring, in 5–7 year old children, the link between speech auditory perceptual ability and the capacity for speech motor

learning through a combination of speech motor adaptation to altered auditory feedback and a period of auditory perceptual *training*. Subjects first underwent a test of speech adaptation involving a manipulation of auditory feedback during production of the vowel /ε/ (as in “bed”). The manipulation involved an increase in the first formant frequency (F1) of the vowel acoustic signal --- an acoustic-spectral property inversely related to the height of the tongue within the vocal tract and a major determinant of vowel identity (Ladefoged, 2000). The ability of children to *perceive* a change in F1 for this vowel was subsequently enhanced through a period of perceptual training, immediately prior to carrying out a second test of speech adaptation to altered auditory feedback. We hypothesized that improvements in F1 perception following perceptual training would increase the magnitude of perceived error under conditions of altered auditory feedback, resulting in greater speech motor adaptation. The results, when compared with those of a separate group of children who underwent perceptual training on an unrelated phoneme contrast, support this hypothesis, demonstrating that changes in children’s ability to perceive relevant speech acoustic properties (such as formant frequencies) can have a direct impact on their capacity for sensory-based speech learning related to those properties.

Materials & Methods

Subjects

22 English speaking children, 5–7 years of age, were tested. Participants were divided into two perceptual training conditions: 1) the *experimental* condition (EXP-Group; n=11; 5 females; mean age = 6.4 years) in which the children underwent perceptual training on a phonemic contrast directly related to the speech motor adaptation task (the vowels /ε/ vs. /æ/), and 2) the *sham* condition (SHAM-Group; n=11; 6 females; mean age = 6.2 years) in which the children underwent perceptual training on a phonemic contrast that was not related to the test of speech motor adaptation (the consonants /b/ vs. /d/). All subjects were native English speakers, with no history of speech, language, or hearing disorder. Hearing status was confirmed by a pure-tone hearing screening prior to testing.

Procedures

Subjects in both groups underwent a sequence of tasks that included 1) a baseline evaluation of speech production, 2) a pre-test of speech motor adaptation to altered auditory feedback 3) a speech perception pre-test, 4) a period of auditory perceptual training, 5) a speech perception post-test, and 6) a post-test of speech motor adaptation to altered auditory feedback (see Figure 1 for schematic). Speech was recorded in a quiet testing room using a head-mounted microphone (C520, AKG, Germany) and digitized at 16-bit / 44.1 kHz on a PC using custom software written in Matlab (v.2010b, Mathworks, MA). Auditory speech signals were presented to subjects using circumaural headphones (880 pro, Beyerdynamic, Germany).

1. Baseline speech production—The first task involved the repeated production of two nonsense words under conditions of normal auditory feedback. For children in the EXP-Group, the words were “Beb” (/bεb/, containing the target vowel /ε/, as in “head”) and “Bab” (/bæb/ containing the target vowel /æ/, as in “had”), produced 15 times each in

randomized order. The task was implemented as a child-friendly computer-based activity in which the children were instructed to say the names of two different cartoon characters (“Beb” and “Bab” for the EXP-group) when they appeared on a computer screen. The correspondence between the two cartoon characters and the two target words was counterbalanced within each group.

2. Speech Motor Learning Test 1—The baseline production task was followed by an initial test of speech motor adaptation involving 100 productions of the target word “Beb”. As in prior studies of speech adaptation to altered auditory feedback (Purcell & Munhall, 2006; Rochet-Capellan & Ostry, 2011; Shiller, et al., 2009; Villacorta, et al., 2007), subjects underwent four auditory feedback conditions in the following sequence: 1) unaltered feedback (30 trials, *null* phase), 2) ramp up to maximum shift (20 trials, *ramp* phase), 3) maintained at maximum shift (40 trials, *hold* phase), and 4) return to unaltered feedback (10 trials, *after-effect* phase). The auditory feedback manipulation corresponded to a 25% increase in F1 frequency (in Hz), which resulted in a vowel that was perceived to be more like the vowel /æ/ (i.e., “Bab”; see *Real-time alteration of speech* for details).

3. Speech Perception Test 1—Following this initial motor adaptation test, the children in the EXP-Group underwent an initial test of speech perception (labeled PT1, in Figure 1) involving the identification of speech sounds that lie along a continuum from /ε/ to /æ/. In order to maintain a reasonable testing time for these child subjects, only 4 stimulus steps along the continuum were used to evaluate the children’s perception of this contrast. The stimuli were generated uniquely for each subject on the basis of each child’s own speech. The 4 stimulus steps consisted of unaltered recordings of the child’s production of “Beb” (step 1) in addition to three altered versions of those stimuli in which F1 frequency (in Hz) was increased by 15% (step 2), 25% (step 3) or 35% (step 4) relative to the original value in “Beb”. These three steps, involving proportional increases in F1 frequency, were chosen in order to provide for each subject a linear continuum that was closely matched to the manipulation of auditory feedback carried out in the speech motor learning test. Five different productions of “Beb”, drawn from the baseline task, were used to construct the stimuli, thus providing a degree of naturally-occurring token-to-token variation in the stimuli used for the perceptual testing. In the perception test, each stimulus step was presented 10 times (the 5 different tokens presented two times each) in a fully randomized order, yielding 40 trials in total. Upon hearing each stimulus, the child had to identify it as “Beb” or “Bab” by touching the appropriate cartoon character on a touch-sensitive computer screen.

4. Auditory perceptual training—For children in the EXP-Group, the initial speech perception test was followed by four blocks of perceptual training trials (5–10 minutes each, labeled *perceptual training 1–4* in Figure 1) designed to improve the child’s perception of the change in F1 frequency associated with the contrast between /ε/ and /æ/. The training involved a series of tasks in which the child had to identify stimuli as either “Beb” or “Bab”, as in the perception test, only in this case involving a restricted stimulus set that included only two stimulus steps: 1) stimuli from step 1 (unaltered F1) and 2) stimuli from one other step (F1-altered by 25% or 35%). The stimuli were presented in random order, and

following each response the child was provided feedback about whether their response was correct or incorrect (where the *correct* response always corresponded to the unaltered stimulus being identified as “Beb” and the F1-altered stimulus being identified as “Bab”). To minimize ceiling effects, the stimulus step used for perceptual training was adjusted dynamically based upon the child’s performance, increasing the difficulty of the task as needed to maintain a level of performance below 85% correct. Hence, following the first block of trials (involving the largest contrast between stimulus steps: Level 1 and Level 4), the contrast distance in subsequent blocks would be decreased by one step (Level 1 vs. Level 3) if the child’s performance was greater than 85%. Each training block consisted of approximately 40 stimulus presentations.

The training tasks were implemented as computer games designed to motivate the children to minimize errors. For example, one such game involved a race between a cartoon character controlled by the child and another controlled by the computer. For each correct response, the child’s character would advance toward the finish line, whereas for each incorrect response the child’s opponent would advance toward the finish line.

5. Speech Perception Test 2—A perceptual test identical to the *Speech Perception Test 1* was carried out immediately following the final perceptual training task in order to evaluate the effect of the training.

6. Speech Motor Learning Test 2—Finally, to evaluate the impact of perceptual training on speech motor learning, the children underwent a second test of speech motor adaptation involving altered auditory feedback (identical to the *Speech Motor Learning Test 1*).

SHAM Group

The testing sequence used for the SHAM group was identical to that used for the EXP-group, the only difference being the target sounds produced or perceived during the *baseline* production task and the perceptual tests. Rather than targeting the vowel contrast between /ε/ and /æ/, these tasks involved a completely unrelated sound contrast: /b/ (“Beb”) vs. /d/ (“Deb”). Importantly, the test of speech motor learning involved a target word (“Beb”) and auditory feedback manipulation (25% increase in F1 frequency for /ε/) that was identical to the EXP-group. For Speech Perception Tests 1 and 2, the stimuli consisted of each child’s own baseline recordings of “Beb” and “Deb”, using five different productions of each word. No continuum was used between the two stimuli, rather 50% of the presented words consisted of “Beb” and 50% consisted of “Deb”. For the perceptual training task, all of the perceptual training procedures were identical to those used in the EXP-Group, only using the various exemplars of “Beb” and “Deb”.

The purpose of this control condition was to verify that any observed changes in motor adaptation performance in the EXP-group were in fact due to the change in vowel perception, and not simply to the fact that the motor adaptation task had been carried out twice (with 30 minutes of intervening speech-perception activities). The presence of naturally occurring acoustic variability in the /b–d/ stimuli used by the SHAM-group (combined with the young age of the participants) ensured that identification performance

for this contrast was not perfect, and hence amenable to improvement with feedback-based training.

Real-time alteration of speech

The manipulation of auditory feedback involved a 25% increase in the first formant (F1) of the vowel acoustic signal (mean: 208 Hz). The alteration of F1 was carried out using a system that has been reported previously (Rochet-Cappelan & Ostry, 2011; Shume et al., 2011; Lametti, et al., 2012; Mollai et al., 2013). The microphone input was amplified and split into two channels: one providing an unprocessed signal and the other altered using a digital signal processor to increase the frequency of all vowel formants (VoiceOne, TC Helicon). The vowel alteration was restricted to F1 by splitting both signals into non-overlapping low- and high-frequency components (Wavetek model 753a, 7th order elliptic filter, 115dB/octave roll-off), and then mixing the low-frequency portion of the processed signal with the high-frequency portion of the unprocessed signal. The filter cutoff used to separate the two signals was set at 1400Hz, which lies roughly half-way between the first and second formant values for the production of the vowel /ε/ for children in this age range (based upon pilot studies). The total signal processing delay was less than 15 ms.

Subjects were encouraged to maintain a constant speaking volume throughout the task through the use of a VU meter presented on the computer display (showing current and peak acoustic signal level during each trial). Subjects were instructed to maintain a target level on the display, which was adjusted at the beginning of the experiment to correspond to a comfortable speaking volume. The subject's perception of his or her own air/bone conducted speech acoustic signal was reduced by mixing the auditory feedback signal (presented at approximately 75 dB SPL) with speech-shaped masking noise (presented at 60 dB SPL).

Acoustic Analyses

For each of the 230 productions of the target words (30 baseline, and 100 during each of the two speech adaptation tests), a 30 ms segment centered about the midpoint of the vowel was extracted. Mean F1 frequency for each segment was then estimated using LPC analysis in Matlab. LPC parameters were chosen on a per-subject basis in order to minimize the occurrence of clearly spurious formant values. Following Villacorta et al. (2007) and others (Rochet-Cappelan & Ostry, 2011; Shum et al., 2011; Lametti et al., 2012; Mollai et al., 2013), changes in vowel production were computed as the proportion change in F1 frequency relative to the mean F1 values during the *null* phase of the adaptation task (averaged over trials 16–30). Such normalized units convey changes in formant values while accounting for individual differences in baseline acoustic properties. Speech adaptation was evaluated at three time-points: 1) the beginning of the Hold phase (averaged over trials 51–60) under conditions of altered auditory feedback, 2) the end of the Hold phase (trials 81–90) under conditions of altered auditory feedback, and 3) during the After-Effect phase following removal of the feedback manipulation (trials 91–100).

Results

Mean baseline production values of F1 frequency were comparable for the two groups, averaging 805.6 Hz (SD: 78.5) for the EXP-group and 844.7 Hz (SD: 73.8) for the SHAM-group. The difference was not statistically reliable ($t[20]=1.20$, $p > 0.05$).

The changes in perception of the /ε-æ/ contrast following the perceptual training task in the EXP-group are shown in Figure 2A. The figure shows the mean proportion of /æ/ responses for each of the four stimulus steps immediately prior to and following the period of perceptual training (blue and red bars respectively). An increase in the proportion of /æ/ responses can be observed following training for the three F1 shifted stimulus conditions (increasing on average by 11.8% for step 2, 20.1% for step 3, and 10.9% for step 4, on average). The un-shifted stimulus (labeled “Null” in Figure 2) shows a much smaller change (averaging 2.3%). The changes in perception of the F1 shifted stimuli were evaluated using a 2-way repeated-measures ANOVA, with STEP (four stimulus steps along the continuum) and CONDITION (pre- vs. post-test) as factors. The main effect of STEP was highly significant ($F[3,30]=27.94$, $p < 0.001$), as was the main effect of CONDITION ($F[1,30]=8.62$, $p < 0.01$), and the interaction effect ($F[3,20]=2.99$, $p < 0.05$). Because of the reliable interaction effect, post-hoc pair-wise comparisons using the Holm-Bonferroni method were carried out to examine the difference between pre- and post-tests at each stimulus step. The tests revealed a reliable change in perception for the 25% shifted stimulus ($t[10]=3.13$, $p < 0.05$), while the remaining three comparisons showed no reliable effect following the correction for multiple comparisons (Null: $t[10]=0.43$, $p > 0.05$; 15% shifted: $t[10]=2.08$, $p > 0.05$; 35% shifted: $t[10]=1.29$, $p > 0.05$). Hence, the ability of the children to perceive the 25%-formant shifted stimuli as different from /ε/ was reliably improved following the perceptual training task.

Measures of speech output during the two speech motor adaptation tasks for children in the EXP-group are shown in Figure 3A. The mean change in F1 frequency relative to baseline is shown over successive blocks of 10 trials. In the motor learning pre-test (prior to perceptual training; blue line) the children show a change in their production of the target vowel following the introduction of the auditory feedback manipulation. The change is compensatory, involving a systematic decrease in F1 frequency (by 6.9% on average at the end of the Hold phase) that partially counteracts the perceived increase in F1 frequency. The change in F1 also shows a degree of persistence following the sudden restoration of normal feedback (trials 91–100), indicating that the compensation effect was due, at least in part, to a learned change in speech motor planning. Following auditory perceptual training (red line), the speech motor adaptation task again shows a compensatory decrease in F1 output, however in this case with increased magnitude (averaging 12.2% at the end of the Hold phase). A 2-way repeated-measures ANOVA was carried out to evaluate the significance of the training effect, with PHASE (Early Hold, Late Hold, and After-Effect) and CONDITION (pre-test vs. post-test) as factors. The results show no reliable effect of PHASE ($F[2,20]=0.39$; $p > 0.05$), but a significant main effect of CONDITION ($F[1,20]=5.26$; $p < 0.05$), and no significant interaction effect ($F[2,20]=0.65$; $p > 0.05$). The children in the EXP group, therefore, showed a reliable increase in the magnitude of speech motor adaptation following perceptual training.

In contrast with the EXP group, children in the SHAM group underwent a series of perceptual training tasks on a phoneme contrast unrelated to the auditory feedback manipulation (/b/ vs. /d/). Figure 2B shows the results of the perception test prior to and following training. While identification performance was quite good prior to training (90% correct for “Beb”, and 81.2% correct for “Deb”), a significant improvement in perception of the contrast was observed following training, with the overall proportion of correct identification responses increasing from 0.86 to 0.95 ($t[10]=4.33$, $p < 0.01$). Speech adaptation was evaluated prior to and following the perceptual training (Figure 3B). During the pre-test (red line), the children were found to adapt their speech production to the change in vowel feedback, averaging 8.1% change in F1 at the end of the Hold phase --- a magnitude that did not differ reliably from the adaptation effect observed during the pre-test for the EXP group ($t[20] = 0.37$, $p > 0.05$). In contrast with the children in the EXP group, however, children in this group showed no reliable improvement in motor adaptation performance following the perceptual training, averaging 8.5% F1 change in the post-test. A repeated-measures ANOVA was carried out with CONDITION (pre-test vs. post-test) and PHASE (early-hold, late-hold and after-effect) as factors. A reliable effect of PHASE was observed ($F[2,20]=4.43$; $p < 0.05$), however there was no reliable main effect of CONDITION ($F[1,20]=0.38$; $p > 0.05$) and no reliable interaction effect ($F[2,20]=0.28$; $p > 0.05$).

Because of the within-subjects design, it was possible to evaluate the relationship between changes in vowel perception and changes in speech motor learning among the children in the EXP-group (Figure 4). An index of perceptual improvement was calculated for each subject as the proportional change in /æ/ responses for the formant-shifted stimuli following perceptual training. An index of motor adaptation improvement was calculated for each subject as the change in the F1 adaptation response at the end of the hold phase following perceptual training (where more positive values correspond to a greater compensatory response). The Pearson product-moment correlation coefficient was computed between these variables, and found to be positive and statistically reliable ($r[9] = 0.68$, $p < 0.05$). Hence, subjects showing greater improvement in speech perception also showed a greater improvement in speech motor adaptation.

Discussion

In the present study, we examined whether changes in perception of a speech sound contrast alter the capacity for speech motor learning in 5–7 year old children by combining a brief period of intensive auditory perceptual training with tests of speech motor adaptation to altered auditory feedback. After approximately 20 minutes of perceptual training, children’s likelihood of identifying a change in F1 frequency for the vowel /ε/ as a different vowel category (/æ/) was demonstrably increased. This perceptual improvement resulted in a significant increase in the children’s speech motor adaptation to a manipulation in F1 frequency related to the production of that vowel --- an effect whose magnitude was found to be positively correlated with the degree of perceptual improvement. A second group of children underwent an identical series of training and testing procedures, however in their case the auditory perceptual training involved a phoneme contrast completely unrelated to the auditory feedback manipulation. This control group showed no corresponding

improvement in speech motor adaptation to a change in vowel auditory feedback, further supporting the link between the perceptual improvement observed in the experimental group and their improved motor adaptation performance.

The present study adds to a very small body of research on sensorimotor adaptation in speech production in pediatric populations (MacDonald et al., 2012), indicating that young children are indeed capable of adapting their speech motor patterns to changes in auditory feedback. Our results suggest, however, that while children's auditory perceptual abilities allow for speech motor adaptation, they may not yet be fully optimized for the task. Young children have been shown in numerous studies to have greater difficulty than older children and adults in making accurate perceptual judgments of speech, in particular when the acoustic signal is distorted or degraded (Edwards et al., 2002; Hazan & Barrett, 2000; Elliott et al., 1987; Munson, 2001b; Nittrouer, 1992; Ryalls & Pisoni, 1997; Walley, 1988), indicating that attention and/or sensitivity to certain types of speech acoustic information continues to develop through to late childhood. This is consistent with the more general finding that children's phoneme category representations (i.e., the slopes of phoneme identification functions) become more clearly defined with age (see, e.g., Walley & Flege, 1999; Nittrouer, 2002). In the present study, the 5–7-year-old children were capable of adapting to the altered auditory feedback even prior to perceptual training, indicating that their perceptual abilities were sufficient for such sensory-error-based learning. Adaptation performance further improved, however, in conjunction with improvements in their ability to perceive a relevant acoustic property (F1 frequency) following vowel perceptual training, suggesting that their prior auditory perceptual ability had been limiting their speech motor adaptation performance.

It is worth noting that the improvement in speech perception and adaptation performance observed in the present study may have been somewhat enhanced by the choice of vowel contrast (between /ε/ and /æ/), which does not rely as heavily on differences in steady-state formant values as other, more perceptually distinct contrasts such as the contrast between the vowels /ε/ and /I/ (see, e.g., Hillenbrand et al., 1995). It is also worth noting that while, in the present study, the perceptual training involved children making explicit judgments on auditory stimuli, that aspect of the training task may not have in fact been necessary. It is possible that simply by receiving greater exposure to stimuli involving the relevant vowel contrast, improvements in perception, and hence speech motor adaptation, would have been observed. Future studies examining different phoneme contrasts and perceptual training methods will be necessary to determine how critical such factors are to the observed relationship between perception and motor adaptation performance.

The present findings have implications for our understanding of developmental speech disorders of unknown origin (e.g., developmental phonological or articulation disorder). It has been demonstrated that children with speech sound disorders have difficulty with speech perception when listening to live-voice or recorded natural speech (Cohen & Diehl, 1963; Hoffman et al., 1983; Kronvall & Diehl, 1952; Marquardt & Saxman, 1972; Rvachew et al., 2003; Sherman & Geith, 1967; Smit & Bernthal, 1983), digitally altered natural speech (Edwards et al., 2002; Monnin & Huntington, 1974; Raaymakers & Crul, 1988), and synthesized speech (Broen et al., 1983; Hoffman et al., 1985; Rvachew & Jamieson, 1989).

The nature of the perceptual deficits appears to be in the realm of acoustic-phonetic representations (i.e., knowledge of the specific acoustic cues that permit identification of the relevant phonetic categories), rather than in the domain of more abstract knowledge of the speech sound categories that comprise their native language (see, e.g., Edwards, Fourakis, Beckman, & Fox, 1999). The idea that perceptual deficits *cause* speech disorders is further supported by a number of studies in which children with speech disorders were trained to better perceive the difference between speech errors and normally articulated sounds (Jamieson & Rvachew, 1992; Rvachew, 1994; Rvachew et al., 2004). The use of such perceptual training (alone or in combination with articulation therapy) was shown to have a facilitating effect on the children's speech production outcomes, providing evidence that perceptual deficits were contributing directly to their speech difficulties. The results of the present study complement this work nicely, demonstrating that improvements in children's auditory perceptual abilities do not simply improve motor performance, but alter the capacity for auditory-feedback based speech motor learning --- a process that is arguably central to the clinical treatment of speech production disorders. Future studies examining speech motor adaptation in children with speech disorders in combination with perceptual training procedures may further clarify the sensorimotor mechanisms underlying childhood speech disorders, as well as provide important clues to how to optimally incorporate auditory-perceptual training into their treatment.

Acknowledgments

We would like to thank Shari Baum for her valuable comments on the manuscript. This work was supported by the Natural Sciences and Engineering Research Council of Canada and the National Institutes of Health (NIDCD-R01DC012502).

References

- Broen PA, Strange W, Doyle SS, Heller JH. Perception and production of approximant consonants by normal and articulation-delayed preschool children. *Journal of Speech, Language, and Hearing Research*. 1983; 26 (4):601–608.
- Cai S, Ghosh SS, Guenther FH, Perkell JS. Adaptive auditory feedback control of the production of formant trajectories in the Mandarin triphthong /iau/ and its pattern of generalization. *Journal of the Acoustical Society of America*. 2010; 128 (4):2033–2048. [PubMed: 20968374]
- Callan DE, Kent RD, Guenther FH, Vorperian HK. 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*. 2000; 43 (3):721–736.
- Cohen JH, Diehl CF. Relation of speech sound discrimination ability to articulation-type speech defects. *Journal of Speech and Hearing Disorders*. 1963; 28:187–190. [PubMed: 14022055]
- Contreras-Vidal JL, Bo J, Boudreau JP, Clark JE. Development of visuomotor representations for hand movement in young children. *Experimental Brain Research*. 2005; 162:155–164. [PubMed: 15586275]
- De Boysson-Bardies B, Halle P, Sagart L, Durand C. A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*. 1989; 16(1):1–17. [PubMed: 2925806]
- Edwards J, Fourakis M, Beckman ME, Fox RA. Characterizing knowledge deficits in phonological disorders. *Journal of Speech, Language, and Hearing Research*. 1999; 42:169–186.
- Edwards J, Fox RA, Rogers CL. Final consonant discrimination in children: effects of phonological disorder, vocabulary size, and articulatory accuracy. *Journal of Speech, Language, and Hearing Research*. 2002; 45:231–242.

- Eilers RE, Oller DK. Infant vocalizations and the early diagnosis of severe hearing impairment. *Journal of Pediatrics*. 1994; 124 (2):199–203. [PubMed: 8301422]
- Eisenberg LS. Current state of knowledge: speech recognition and production in children with hearing impairment. *Ear and Hearing*. 2007; 28 (6):766–772. [PubMed: 17982364]
- Elliott LL, Hammer MA, Evan KE. Perception of gated, highly familiar spoken monosyllabic nouns by children, teenagers, and older adults. *Perception Psychophysics*. 1987; 42 (2):150–157. [PubMed: 3627935]
- Guenther FH. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*. 1995; 102 (3):594–621. [PubMed: 7624456]
- Hazan V, Barrett S. The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*. 2000; 28:377–396.
- Hillenbrand J, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*. 1995; 97:3099–3111. [PubMed: 7759650]
- Hoffman PR, Daniloff RG, Bengoa D, Schuckers G. Misarticulating and normally articulating children's identification and discrimination of synthetic [r] and [w]. *Journal of Speech and Hearing Disorders*. 1985; 50:46–53. [PubMed: 3974212]
- Hoffman PR, Stager S, Daniloff RG. Perception and production of misarticulated /r/. *Journal of Speech and Hearing Disorders*. 1983; 48 (2):210–215. [PubMed: 6621013]
- Houde JF, Jordan MI. Sensorimotor adaptation in speech production. *Science*. 1998; 279 (5354):1213–1216. [PubMed: 9469813]
- Houde JF, Jordan MI. Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech, Language and Hearing Research*. 2002; 45 (2):295–310.
- Jamieson DG, Rvachew S. Remediation of speech production errors with sound identification training. *Journal of Speech-Language Pathology and Audiology*. 1992; 16:201–210.
- Jones JA, Munhall KG. Perceptual calibration of F0 production: evidence from feedback perturbation. *Journal of the Acoustical Society of America*. 2000; 108 (3 Pt 1):1246–1251. [PubMed: 11008824]
- Jones JA, Munhall KG. Remapping auditory-motor representations in voice production. *Current Biology*. 2005; 15 (19):1768–1772. [PubMed: 16213825]
- Kent RD, Osberger MJ, Netsell R, Hustedde CG. Phonetic development in identical twins differing in auditory function. *Journal of Speech and Hearing Disorders*. 1987; 52 (1):64–75. [PubMed: 3807347]
- Kronvall EL, Diehl CF. The relationship of auditory discrimination to articulatory defects of children with no known organic impairment. *Journal of Speech and Hearing Disorders*. 1952; 19:335–338. [PubMed: 13212839]
- Kuhl PK, Meltzoff AN. Infant vocalizations in response to speech: vocal imitation and developmental change. *Journal of the Acoustical Society of America*. 1996; 100 (4 Pt 1):2425–2438. [PubMed: 8865648]
- Ladefoged, P. *A Course in Phonetics*. 4. Belmont: Wadsworth Publishing; 2000.
- Lametti DR, Nasir SM, Ostry DJ. Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *The Journal of Neuroscience*. 2012; 32:9351–9358. [PubMed: 22764242]
- MacDonald EN, Johnson EK, Forsythe J, Plante P, Munhall KG. Children's development of self-regulation in speech production. *Current biology*. 2012; 22 (2):113–117. [PubMed: 22197241]
- Marquardt TP, Saxman JH. Language comprehension and auditory discrimination in articulation deficient kindergarten children. *Journal of Speech and Hearing Research*. 1972; 15:382–389. [PubMed: 5047876]
- McGowan RS, Nittrouer S, Chenausky K. Speech production in 12-month-old children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*. 2008; 51 (4):879–888.
- Mollaei F, Shiller DM, Gracco VL. Sensorimotor adaptation of speech in Parkinson's disease. *Movement disorders : official journal of the Movement Disorder Society*. (In Press).
- Monnin LM, Huntington DA. Relationship of articulatory defects to speech-sound identification. *Journal of Speech and Hearing Research*. 1974; 17:352–366. [PubMed: 4424729]

- Mowrer OH. Speech development in the young child. I. The autism theory of speech development and some clinical applications. *Journal of Speech and Hearing Disorders*. 1952; 17 (3):263–268. [PubMed: 13053547]
- Munson B. Relationships between vocabulary size and spoken word recognition in children aged 3 to 7. *Contemporary Issues in Communication Science and Disorders*. 2001b; 28:20–29.
- Nittrouer S. Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics*. 1992; 20:1–32.
- Nittrouer S. Learning to perceive speech: how fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America*. 2002; 112:711–719. [PubMed: 12186050]
- Oller DK, Eilers RE. The role of audition in infant babbling. *Child Development*. 1988; 59 (2):441–449. [PubMed: 3359864]
- Perkell JS, Matthies ML, Lane H, Guenther FH, Wilhelms-Tricarico R, Wozniak J, et al. Speech motor control: acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*. 1997; 22:227–250.
- Perkell JS, Matthies ML, Tiede M, Lane H, Zandipour M, Marrone N, Stockmann E, Guenther FH. The distinctness of speakers' /s-/S/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech Language and Hearing Research*. 2004; 47:1259–69.
- Perkell JS, Guenther FH, Lane H, Matthies ML, Stockmann E, Tiede M, Zandipour M. The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *Journal of the Acoustical Society of America*. 2004; 116:2338–44. [PubMed: 15532664]
- Purcell DW, Munhall KG. Adaptive control of vowel formant frequency: evidence from real-time formant manipulation. *Journal of the Acoustical Society of America*. 2006; 120 (2):966–977. [PubMed: 16938984]
- Raaymakers EM, Crul TAM. Perception and production of the final /s-ts/ contrast in Dutch by misarticulating children. *Journal of Speech and Hearing Disorders*. 1988; 53:262–270. [PubMed: 3398480]
- Rochet-Capellan A, Ostry DJ. Simultaneous acquisition of multiple auditory-motor transformations in speech. *Journal of Neuroscience*. 2011; 31(7):2657–2662. [PubMed: 21325534]
- Rvachew S. Speech perception training can facilitate sound production learning. *Journal of Speech and Hearing Research*. 1994; 37:347–357. [PubMed: 8028316]
- Rvachew S. Phonological processing and reading in children with speech sound disorders. *American Journal of Speech-Language Pathology*. 2007; 16:260–270. [PubMed: 17666551]
- Rvachew S, Jamieson DG. Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders*. 1989; 54:193–208. [PubMed: 2709838]
- Rvachew S, Nowak M, Cloutier G. Effect of phonemic perception training on the speech production and phonological awareness skills of children with expressive phonological delay. *American Journal of Speech-Language Pathology*. 2004; 13:250–263. [PubMed: 15339234]
- Rvachew S, Ohberg A, Grawburg M, Heyding J. Phonological awareness and phonemic perception in 4-year-old children with delayed expressive phonology skills. *American Journal of Speech-Language Pathology*. 2003; 12:463–471. [PubMed: 14658998]
- Rvachew S, Slawinski EB, Williams M, Green CL. The impact of early onset otitis media on babbling and early language development. *The Journal of the Acoustical Society of America*. 1999; 105:467–475. [PubMed: 9921670]
- Ryalls BO, Pisoni DB. The effect of talker variability on word recognition in preschool children. *Developmental Psychology*. 1997; 33 (3):441–452. [PubMed: 9149923]
- Sherman D, Geith A. Speech sound discrimination and articulation skill. *Journal of Speech and Hearing Research*. 1967; 10:277–280. [PubMed: 6082354]
- Shiller DM, Sato M, Gracco VL, Baum SR. Perceptual recalibration of speech sounds following speech motor learning. *Journal of the Acoustical Society of America*. 2009; 125 (2):1103–1113. [PubMed: 19206885]
- Shiller DM, Gracco VL, Rvachew S. Auditory-motor learning during speech production in 9–11-year-old children. *PloS One*. 2010; 5(9):e12975. [PubMed: 20886033]

- Shum M, Shiller DM, Baum SR, Gracco VL. Sensorimotor integration for speech motor learning involves the inferior parietal cortex. *The European journal of neuroscience*. 2011; 34 (11):1817–1822. [PubMed: 22098364]
- Smit AB, Bernthal JE. Performance of articulation-disordered children on language and perception measures. *Journal of Speech and Hearing Research*. 1983; 26:124–136. [PubMed: 6865368]
- Stoel-Gammon C, Williams K, Buder E. Cross-language difference in phonological acquisition: Swedish and American /t/. *Phonetica*. 1994; 51:146–158. [PubMed: 8052669]
- Villacorta VM, Perkell JS, Guenther FH. Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *Journal of the Acoustical Society of America*. 2007; 122 (4):2306–2319. [PubMed: 17902866]
- Walley A. Spoken word recognition by children and adults. *Cognitive Development*. 1988; 3:137–165.
- Walley AC, Flege JE. Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics*. 1999; 27:307–332.

Figure 1.
Schematic of testing sequence.

Figure 2.

A. Mean responses for the tests of vowel perception prior to and following perceptual training for the EXP-group. Data show the mean proportion of /æ/ responses for each of the four perceptual stimuli. Following training, the children's ability to identify stimuli whose F1 value has been increased is improved, as indicated by an increase in the proportion of /æ/ responses for the formant-shifted stimuli. Error bars show 1 SEM. **B.** Mean responses for the perception tests for the SHAM-group (proportion of "Deb" responses). The children showed a reliable improvement in the proportion of correct identification responses following perceptual training. Error bars show 1 SEM.

Figure 3.

Results of the test of speech motor learning. 3A. Mean F1 frequency produced by subjects in the EXP-group averaged over successive blocks of 10 trials, prior to perceptual training (blue) and immediately following the training (red). The perceptual training resulted in greater compensation to altered auditory feedback of the vowel sound. 3B. For subjects in the SHAM-group, who underwent perceptual training unrelated to the auditory feedback manipulation, no such improvement in motor adaptation is observed between the pre-test (blue) and post-test (red). Error bars show ± 1 SEM.

Figure 4.

Relationship between perceptual improvement and speech motor adaptation. A significant positive correlation was observed between the degree of vowel perceptual improvement and the degree of motor adaptation improvement for children in the EXP-group.