J Neurophysiol 122: 2076–2084, 2019.
First published September 11, 2019; doi:10.1152/jn.00028.2019.

RESEARCH ARTICLE | *Sensory Processing*

1972

Recalibration of auditory perception of speech due to orofacial somatosensory inputs during speech motor adaptation

Hiroki Ohashi[1,2] and **Takayuki Ito**[2,3]
[1]*Department of Psychology, McGill University, Montreal, Quebec, Canada;* [2]*Haskins Laboratories, New Haven, Connecticut; and* [3]*Centre National de la Recherche Scientifique, GIPSA-Lab, Grenoble Institute of Technology, University of Grenoble-Alpes, Saint Martin d'Heres, France*

Submitted 15 January 2019; accepted in final form 9 September 2019

Ohashi H, Ito T. Recalibration of auditory perception of speech due to orofacial somatosensory inputs during speech motor adaptation. *J Neurophysiol* 122: 2076–2084, 2019. First published September 11, 2019; doi:10.1152/jn.00028.2019.—Speech motor control and learning rely on both somatosensory and auditory inputs. Somatosensory inputs associated with speech production can also affect the process of auditory perception of speech, and the somatosensory-auditory interaction may play a fundamental role in auditory perception of speech. In this report, we show that the somatosensory system contributes to perceptual recalibration, separate from its role in motor function. Subjects participated in speech motor adaptation to altered auditory feedback. Auditory perception of speech was assessed in phonemic identification tests before and after speech adaptation. To investigate a role of the somatosensory system in motor adaptation and subsequent perceptual change, we applied orofacial skin stretch in either a backward or forward direction during the auditory feedback alteration as a somatosensory modulation. We found that the somatosensory modulation did not affect the amount of adaptation at the end of training, although it changed the rate of adaptation. However, the perception following speech adaptation was altered depending on the direction of the somatosensory modulation. Somatosensory inflow rather than motor outflow thus drives changes to auditory perception of speech following speech adaptation, suggesting that somatosensory inputs play an important role in tuning of perceptual system.

NEW & NOTEWORTHY This article reports that the somatosensory system works not equally with the motor system, but predominantly in the calibration of auditory perception of speech by speech production.

auditory-somatosensory integration; speech motor learning; speech perception; speech production; speech sound acquisition

INTRODUCTION

The nature of the interaction between production and perception mechanisms is a central question in speech science. The mechanisms of production and perception are presumably shaped not only individually by, but also through the interaction between, sensory inflow and motor outflow. Alterations of auditory and somatosensory feedback during speech training result in a recalibration of sensorimotor mapping (Houde and Jordan 1998; Tremblay et al. 2003). Intriguingly, the recalibration associated with speech motor training induces changes to the speaker's auditory map for speech perception (Lametti et al. 2014; Nasir and Ostry 2009; Schuerman et al. 2017; Shiller et al. 2009). These results support the idea that the tuning of sensory perception is integrated into the process of motor learning and adaptation. Although the recalibration of sensory perception occurs in the context of both sensory inflow and motor outflow, it is not known whether these two work together or independently in the processing of recalibration.

The involvement of motor system in auditory perception of speech has been examined extensively (Hickok et al. 2011; Liberman et al. 1967; Liberman and Mattingly 1985). Disruption of premotor cortex and the lip representation in the motor cortex by transcranial magnetic stimulation impairs speech categorical perception (Meister et al. 2007; Möttönen and Watkins 2009). Listening to speech sounds modulates human motor cortical excitation (Fadiga et al. 2002; Pulvermüller et al. 2006). From this perspective, the audio-motor interaction that originates from changes in motor outflow can be a source of changes to auditory perception of speech in speech motor adaptation.

As an alternate point of view, changes in somatosensory inflow can be a potential source of the change of auditory perception of speech induced by speech motor adaptation, since the somatosensory system per se contributes to auditory perception of speech. Psychophysical studies have found that the identification of phonemic contrasts is altered by somatosensory inputs (Gick and Derrick 2009; Ito et al. 2009). This idea is supported by evidence beyond the speech literature. Somatosensory stimulation activates auditory cortex, and auditory stimulation activates somatosensory cortex (Perez-Bellido et al. 2018; Schürmann et al. 2006). Transcranial magnetic stimulation to somatosensory cortex modulates auditory frequency discrimination (Convento et al. 2018). Anatomical connections between primary auditory cortex and primary and secondary somatosensory cortex have also been documented (Ro et al. 2013). This link between the somatosensory and auditory systems may contribute to the system of auditory perception of speech on its own or in combination with the mechanisms of motor outflow.

The present study examined how somatosensory inputs affect the changes in auditory perception of speech that are

---

Address for reprint requests and other correspondence: T. Ito, CRNS, GIPSA-Lab, Grenoble INP, University of Grenoble-Alpes, 11 rue des Mathématiques, Grenoble Campus BP46, F-38402 Saint Martin d'Heres Cedex, France (e-mail: takayuki.ito@gipsa-lab.grenoble-inp.fr).

induced by speech motor adaptation. As an experimental model of speech motor adaptation and perceptual changes, a procedure similar to that of Shiller et al. (2009) was carried out in which auditory feedback of consonant sounds was altered. Before and after adaptation training, consonant identification tests were carried out. During the adaptation training, we applied a somatosensory perturbation involving facial skin deformation with the assumption that the facial skin deformation provides kinesthetic information normally associated with speech articulatory movement (Ito and Gomi 2007; Ito and Ostry 2010). This technique enables us to test a direct contribution of somatosensory inflow during speech training to auditory perception of speech by modifying somatosensory inflow alone without changing motor outflow.

MATERIALS AND METHODS

Subjects and procedure. Eighty-nine native speakers of American English aged 18–35 yr old were tested, 65 in the main experiment (41 men and 24 women) and 24 in a follow-up study (13 men and 11 women). Subjects had no reported impairment of hearing or speech. The Human Investigation Committee of Yale University approved the experimental protocol. Subjects provided written, informed consent.

The main experiment aimed to examine how somatosensory inputs can modulate perceptual shifts that occur in conjunction with speech motor adaptation. The experiment consisted of three sessions. In the first session, subjects performed a perceptual test to measure baseline perception for a contrast between fricative consonants. The second session involved speech adaptation training with altered auditory feedback (Houde and Jordan 1998; Lametti et al. 2012, 2014; Purcell and Munhall 2006; Schuerman et al. 2017; Shiller et al. 2009; Villacorta et al. 2007). The third session was a perceptual test to measure the aftereffects of speech motor adaptation on auditory perception of speech. These three sessions were carried out in a row without breaks and took 25 min in total. To examine effects of somatosensory inflow, we applied somatosensory perturbations using facial skin deformation (Ito et al. 2009) during the speech adaptation training. The perceptual shifts were assessed by comparing perceptual thresholds before and after speech adaptation training.

We focused on the fricative contrast between /ʃ/ and /s/, as in the work by Shiller et al. (2009), for both production and perceptual testing. This contrast is primarily characterized by a difference in spectral centroid values, which is a frequency of the spectral center of gravity (Maniwa et al. 2009). We manipulated the spectral centroid to alter auditory feedback for adaptation training and to synthesize the speech continuum between /ʃ/ and /s/ for perceptual testing (see below for detail). These consonants were embedded in the carrier phrase "a _ed." The carrier phrase stabilizes temporal variation of target consonants across trials.

The direction of skin stretch was based on articulatory characteristics of fricative production in which the production of /ʃ/ involves more lip protrusion than /s/ (see Fig. 1*A*). The direction of the somatosensory perturbation was either backward or forward and was only applied during the production of /ʃ/. The backward skin stretch was expected to attenuate somatosensation arising from the lip protrusion for /ʃ/, and the forward skin stretch is expected to increase it.

The 65 subjects were assigned to one of four groups. The first group received altered auditory feedback and a backward skin stretch perturbation during the production of /aʃɛd/ (AS$_{B}$). The second group received altered auditory feedback and a forward skin stretch perturbation (AS$_{F}$). The third group received altered auditory feedback during the production of /aʃɛd/ without skin stretch (A). The fourth group produced /aʃɛd/ without any perturbations (CTL). As in previous studies with altered auditory feedback (Houde and Jordan 1998; Lametti et al. 2012, 2014; Nasir and Ostry 2009; Purcell and Munhall 2006; Shiller et al. 2009; Tremblay et al. 2003; Villacorta et al. 2007), not all subjects can adapt. For purposes of the present study, to examine the effects of somatosensory feedback manipulations on

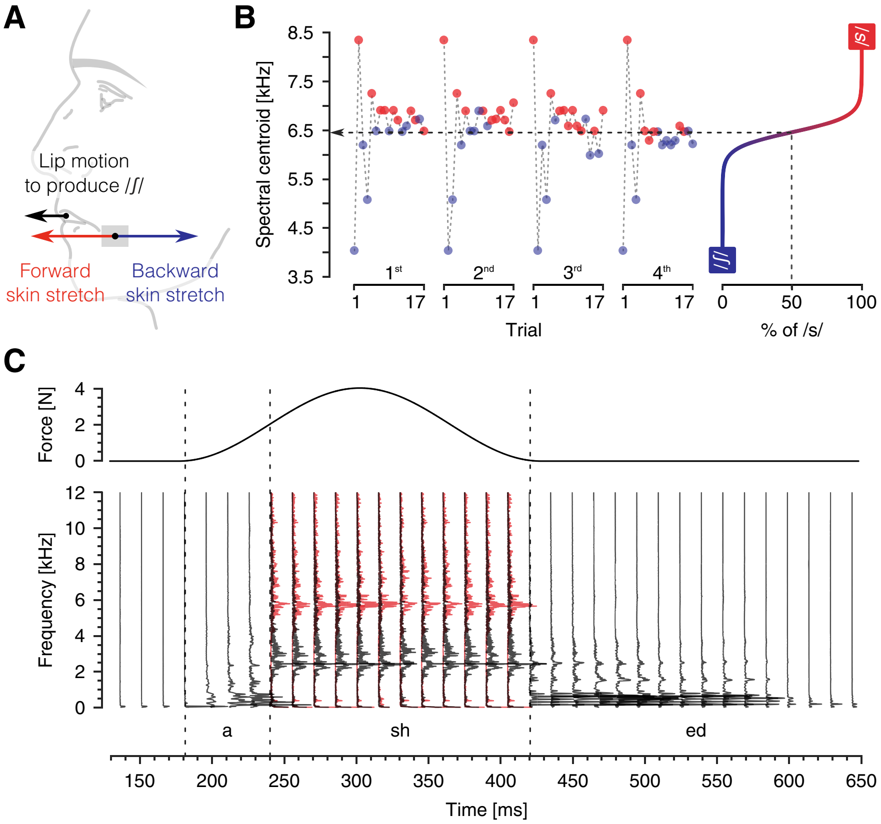

Fig. 1. Experimental setup. *A*: illustration of skin stretch direction and the predominant direction of movement for the consonant /ʃ/. *B*: representative example of the maximum-likelihood procedure (*left*) and psychometric function (*right*). Red and blue circles indicate trials in which a subject identified the stimulus as /s/ and /ʃ/, respectively. Dashed line represents the perceptual threshold between /s/ and /ʃ/, measured as the 50% point of the fitted psychometric function. *C*: temporal profile of the force applied to the facial skin (*top*) and spectrogram of sounds in "a shed" (*bottom*). The first vertical dashed line represents the onset of vocalization. This event triggered the somatosensory perturbation. The second and third dashed lines represent the onset and offset timing of the fricative sound, respectively. In this period, the spectrum of produced sounds (black) was shifted toward higher frequencies (red).

perceptual changes that are induced by auditory feedback adaption, we excluded from analysis subjects that failed to adapt to altered auditory feedback. Following removal of nonadapting subjects, there were an equal number of subjects ($n = 12$; numbers of male subjects were 7 in AS_B, 6 in AS_F, 8 in A, and 8 in CTL) across the four groups for the subsequent analysis (see *Data analysis*). Table 1 shows the total number of subjects we tested and the number of subjects who adapted to auditory perturbation in each group.

The follow-up experiment examined whether or not the skin stretch perturbation itself directly changes acoustical properties of the produced fricative sounds. Potential changes to speech sounds were assessed during the production of /aʃɛd/ in two conditions: forward and backward skin stretch, as in the main experiment. Twelve subjects were tested in each of the skin stretch conditions.

Speech motor adaptation task. Subjects were instructed to produce the phrase “a shed” at a comfortable volume level and to maintain a consistent sound volume across trials by monitoring a volume level presented on a display. Produced speech sounds were recorded using a unidirectional microphone (Sennheiser MKH 416). An intertrial interval was varied between 1 and 2 s to prevent anticipation. A total of 160 utterances were produced over the course of the speech training. Produced speech sounds were played back to subjects through headphones (Sennheiser HD 25-SP). Sound volume of auditory feedback was set large enough to hear played-back sounds predominantly over airborne and bone-conducted sounds. Although masking noise is often used to prevent subjects from hearing their own voices, we did not use masking noise because of its potential impact on the fricative sounds that were of primary interest (Shiller et al. 2009) and because perception of bone-conducted sounds would be expected to be minimal for a voiceless fricative such as /ʃ/.

For altered auditory feedback manipulation, the spectral centroid of the target fricative sounds /ʃ/ was shifted up to make the feedback sound more like /s/ (Fig. 1*C*) using the pitch shift function of the Audapter software (Cai et al. 2011). Over the course of the training, the shift was gradually introduced as has been done previously (Houde and Jordan 1998; Purcell and Munhall 2006; Schuerman et al. 2017; Shiller et al. 2009; Villacorta et al. 2007). The initial 30 trials served as a baseline phase in which no auditory feedback alteration was applied. Over the next 30 trials, a magnitude of auditory feedback alteration was gradually increased (ramp phase). The final magnitude at the end of the ramp phase was then maintained over the following 100 trials (hold phase). Given previous findings (Shiller et al. 2009), adaptation to an upward shift of the spectral centroid is expected to result in the spectral centroids of produced fricative sounds being lower than that in the baseline.

In the altered auditory feedback system, audio signals were digitally sampled at 48 kHz and then downsampled at 24 kHz to reduce processing time. Fricative sounds were extracted on the basis of high zero-crossing rate by our custom function implemented in the Audapter software, and then an entire of spectral envelope of the extracted sounds was skewed (Fig. 1*C*). This manipulation results in a shift of the spectral centroid value in the fricative sound. A magnitude of this shift was varied in each subject because of individual differences in the spectral envelope of speech sounds. The maximum shift was set at three semitones, which corresponds to a 7.21% (SE: ±0.69) change of the spectral centroid value of fricative sounds. The resultant sounds at the hold phase of the shift were more similar to the /s/ sound but still in the category of an /ʃ/ sound. Note that the acoustical characteristics of speech sounds other than the fricative part remained unchanged. The onset of the vowel /a/ in “a shed” was detected as an increase in the root mean square of speech signals to trigger the somatosensory perturbation. A feedback delay due to signal processing was ~20 ms.

Table 1. *Population of subjects who adapted or did not adapt to altered auditory feedback in each experimental condition*

	Adapted	Nonadapted	Total
AS_B	12 (52%)	11	23
AS_F	12 (75%)	4	16
A	12 (86%)	2	14

Values are the number (%) of adapted, nonadapted, and total subjects in each condition. AS_B, altered auditory feedback with backward skin stretch; AS_F, altered auditory feedback with forward skin stretch; A, altered auditory feedback alone.

For the somatosensory perturbation, the facial skin stretch was applied using a robotic device (Phantom 1.0; SensAble Technologies). Two plastic tabs (2 cm × 3 cm) were attached to the skin on both sides, lateral to the oral angle (see Fig. 1*A*). These tabs were connected to a robotic device using thin wires. The wires were supported by wire supports with pulleys to avoid contact with the facial skin and the other body parts. The skin was stretched when the robotic device applied a force to the wires. The temporal profile of the applied force was a single cycle of a 4-Hz sinusoid (250-ms duration) that approximates a duration of the lip protrusion in /ʃ/ production (see Fig. 1*C*). The peak force was 4 N, resulting in 10–15 mm of skin stretch. The perturbation began at the onset of the first vowel in “a shed” on the basis of our preliminary observation that lip protrusion during /ʃ/ production begins approximately at the onset of a preceding vowel sound. The somatosensory perturbation was presented on each trial over the course of the training (160 trials). Note that whereas the somatosensory stimulation was applied in all trials, the auditory feedback perturbation was introduced in gradual manner to avoid subjects’ being aware of an auditory change. This is because we eliminated a possibility that simultaneous introduction of the somatosensory and auditory perturbations could give subjects a clue to be aware of a change of auditory conditions and lead unexpected responses by attention or cognitive efforts.

Perceptual test. Identification tests were carried out using a synthesized continuum between the fricative sounds /ʃ/ and /s/ embedded within the carrier phrase. The stimuli were presented through headphones at a comfortable volume. On each trial, subjects were asked to identify whether the sound was “a said” or “a shed” by pressing keys on a keyboard.

We applied an adaptive method, the maximum-likelihood procedure, to select test stimulus on each trial (Shen and Richards 2012). The benefit of this procedure is that it collects subjects’ responses efficiently and thereby is able to estimate the psychometric function with a relatively small number of responses compared with other conventional methods such as the method of constant stimuli. In this procedure, test stimulus on each trial was determined in an adaptive fashion based on the stimulus that provides the most information about the shape of the estimated psychometric function. Each of the perceptual tests consisted of four 17-trial blocks (Fig. 1*B*).

The experimental continuum for the perceptual test was synthesized as in a previous study (Lane et al. 2007). The reference sounds “a said” and “a shed” were recorded from a male native speaker of American English. We extracted fricative sounds by visual inspection of the spectrum and estimated the lowest five spectral peaks for each of /s/ and /ʃ/ by linear predictive coding (Andersen 1974). By interpolating the amplitudes of the spectral peaks between /s/ and /ʃ/, we synthesized a 300-step continuum using the Klatt formant synthesizer (Klatt 1980). Fifty stimuli with linear steps between spectral centroids (mean step size: 87.89 Hz) were picked from the 300-step continuum. Finally, the fricative part /ʃ/ of the reference “a shed” sound was replaced with the synthesized fricatives to generate a 50-step continuum. The resultant spectral centroids at the end points of continuum were 3,835.81 Hz for /ʃ/ and 8,142.32 Hz for /s/.

Data analysis. Performance with altered auditory feedback was evaluated in terms of changes to the spectral centroid of produced fricative consonant /ʃ/ over the course of the training. On each trial, the fricative sound was extracted using Audapter (see *Speech motor adaptation task*). Because of temporal fluctuations of fricative sounds within a single production, we computed spectral centroids at three points (25, 50, and 75% of the total duration of the fricative) using a 20-ms Hamming window and then averaged these three values (Maniwa et al. 2009). The obtained spectral centroids were normalized by dividing by the mean value over the last 20 trials of the baseline phase (11–30th trials).

We included subjects who adapted to auditory perturbation in the following group-level analyses of speech production and perception. Subjects were classified as adapted when the mean spectral centroid in the hold phase (141–160th trials) was significantly lower than the baseline value (11–30th trials) (uncorrected $P < 0.005$; one-tailed two-sample t test) based on the expectation that the spectral centroid value would be reduced as an adaptive behavior to auditory perturbation. We applied a relatively strict level of significance by considering a matter of multiple comparisons.

In the group-level analyses of the speech training, we evaluated three measurements: *1*) the amount of adaptation at the end of the training, *2*) the adaptation rate in the ramp phase, and *3*) the proportion of the total number of subjects who were found to adapt. The amount of adaptation was quantified as the average of the normalized spectral centroids over the last 20 trials at the end of the hold phase (141–160th trials). One-way ANOVA was applied across the four conditions (AS_B, AS_F, A, and CTL). Pairwise comparisons were followed. We applied nonparametric bootstrap tests with Holm–Bonferroni correction to avoid a potential impact on statistical results caused by a bias of data distribution with extreme values (see actual data distribution in Fig. 2*B*). For the adaptation rate in the ramp phase, we tested differences in the spectral centroid in the ramp phase (31–60th trials) across the three perturbation conditions (AS_B, AS_F, or A) using a linear mixed-effect model with Holm–Bonferroni correction. The conditions and the trial numbers were the fixed effects, and subjects was the random effect. A pattern of changes in the spectral centroid value was tested as the random slope effect of the trial numbers (Barr et al. 2013). For the proportion of the subjects who were found to adapt, we tested differences in the proportion between the altered auditory feedback alone and backward skin stretch conditions (A vs. AS_B) and between the altered auditory feedback alone and forward skin stretch conditions (A vs. AS_F) using χ^2 tests with Holm–Bonferroni correction.

In the analysis of the perceptual performance, psychometric functions were estimated by fitting a logistic function to all responses for each of the two perceptual tests in each condition. Residual deviation, which is index of the goodness of the fit of psychometric function, was 56.7 ± 1.67, 51.7 ± 1.85, 53.6 ± 1.82, and 52.9 ± 1.26 (means ± SE) for CTL, A, AS_B, and AS_F, respectively. There was no reliable difference in the goodness of the fit among the conditions [$F_{(3,44)} = 1.64$, $P = 0.193$; one-way ANOVA]. The perceptual threshold between /s/ and /ʃ/ was obtained as the 50% point of the estimated psychometric function in the perceptual identification test (see Fig. 1*B* for an example). Perceptual acuity was obtained as half of the distance between 25% and 75% points of the psychometric function.

In the group-level analyses of the perceptual tests, we evaluated the perceptual change due to speech training using two measurements: *1*) changes in the perceptual threshold following adaptation and *2*) changes in the perceptual acuity following adaptation. These changes were quantified as a difference in each measurement between before and following training. We also evaluated differences in baseline perceptual threshold and baseline perceptual acuity across groups. These baseline threshold and acuity were obtained in the perceptual test before the training. One-way ANOVA was applied to each of the four measures. Post hoc tests were followed. We applied nonparametric bootstrap tests with Holm–Bonferroni correction based on the assumption that the data may not be completely in normal distribution due to extreme values (see actual data distribution in Fig. 3*B*).

A correlation analysis was conducted to test for a linear relationship between adaptation related motor effects (the amount of adaptation) and the perceptual effect following the speech adaptation (the amplitude of the perceptual shift).

Follow-up experiment. Subjects in a follow-up experiment produced the phrase "a shed" 60 times as in the main experiment. The skin stretch perturbation was applied in 15 randomly selected trials. The temporal profile of skin stretch was the same as in the main experiment (see *Speech motor adaptation task*). Note that there was no altered auditory feedback in the follow-up experiment. Two skin stretch directions were tested (backward and forward). Subjects were tested in either the backward or forward condition. We compared the spectral centroid of produced fricative sounds between the conditions with and without the skin stretch perturbation. We also assessed whether there were systematic acoustical effects depending on the direction of skin stretch.

The spectral centroid of /ʃ/ sounds was computed for each trial as described in *Data analysis*. To avoid a possible aftereffect of the perturbed trials, we excluded trials immediately after the perturbed trials from a subsequent analysis. The spectral centroid values were normalized with respect to the mean value of the unperturbed trials to remove intersubject differences in baseline centroid value. The mean value of the unperturbed trials was estimated using a nonparametric bootstrap on a per-subject basis; we computed the mean value of 15 trials randomly selected from the unperturbed trials 10,000 times and then obtained the representative mean value as the grand mean value.

A linear mixed-effect model was applied to the normalized spectral centroid values of the 15 perturbed trials to test differences in the

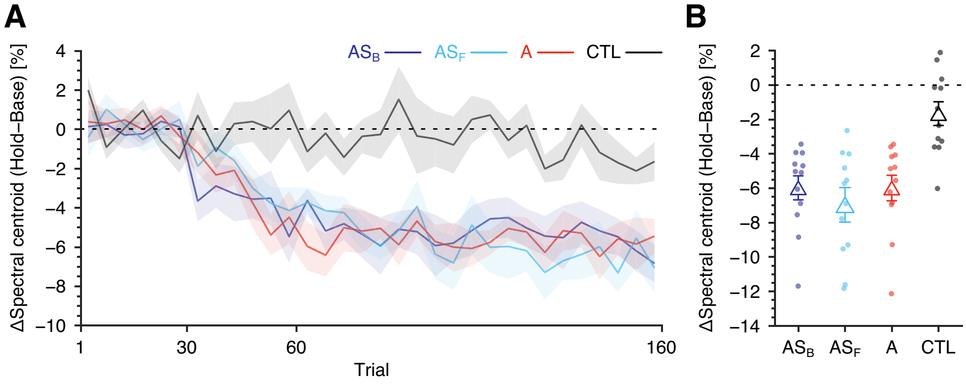

Fig. 2. Changes in the spectral centroid value relative to baseline over the course of speech training (Hold–Base). *A*: each line shows changes in the spectral centroid value (ΔSpectral centroid) averaged over subjects for each condition. Shaded areas represent ±SE. Each time point was obtained as the mean value over 5 trials. *B*: amount of adaptation at last 20 trials. Mean values and SE are represented by triangles and error bars. Each dot indicates an individual value. Blue, cyan, red, and black represent altered auditory feedback with backward skin stretch (AS_B), altered auditory feedback with forward skin stretch (AS_F), altered auditory feedback alone (A), and the control condition (CTL), respectively.

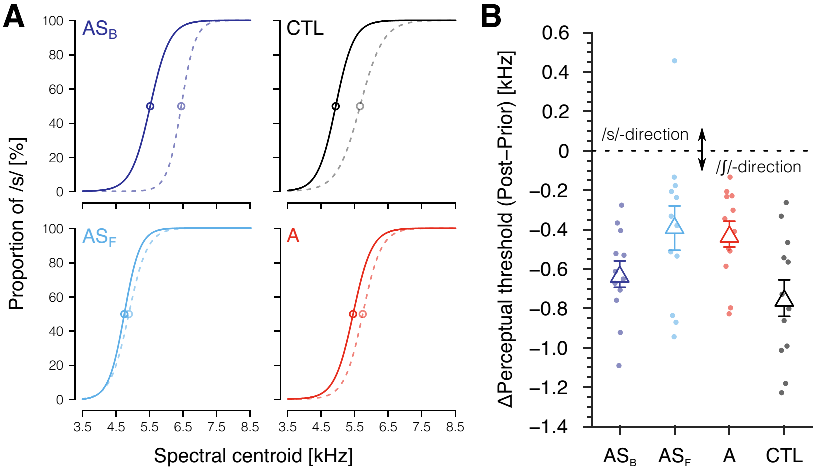

Fig. 3. Perceptual shift following speech motor adaptation. *A*: representative psychometric functions obtained from perceptual tests before (dashed lines) and after (solid lines) speech training. Circles represent the perceptual thresholds between /s/ and /ʃ/ (50% points of the psychometric functions). Blue, cyan, red, and black represent altered auditory feedback with backward skin stretch (AS_B), altered auditory feedback with forward skin stretch (AS_F), altered auditory feedback alone (A), and control condition (CTL), respectively. *B*: mean differences in the threshold (ΔPerceptual threshold) between the two perceptual tests (Post–Prior). Mean values and SE are represented by triangles and error bars. Each dot indicates an individual value. The negative value means that the /ʃ/–/s/ perceptual boundary shifted in the /ʃ/-direction (subjects perceived sounds as more /s/-like).

spectral centroid between the perturbed trials and the unperturbed trials and between the two directions of skin stretch. The direction of the skin stretch (forward or backward) was the fixed effect, and subjects was the random effect. A difference between the perturbed or unperturbed trials was tested as the intercept.

RESULTS

Main experiment. We assessed perceptual thresholds before and after speech training to examine how somatosensory inputs affect the perceptual changes that occur in conjunction with speech motor adaptation. In the main manipulation, we applied somatosensory perturbations that involved facial skin deformation during speech training.

The amount of adaptation to altered auditory feedback at the end of the training was assessed in terms of changes in the spectral centroid of produced fricative consonant /ʃ/. Figure 2 shows changes in the normalized spectral centroid in each experimental condition. In all three conditions with altered auditory feedback (AS_B, AS_F, and A), the spectral centroid of produced sounds gradually decreased in the ramp phase (31–60th trials) to compensate for the gradual increase of the spectral centroids in feedback sounds. The changes at the end of the ramp phase were maintained through the hold phase (61–160th trials). On the other hand, in the CTL condition, the spectral centroids of produced sounds did not change over the course of the training. The amount of the change at the end of the training (averaged over the last 20 trials) in each condition is shown in Fig. 2*B*. In all of the four conditions, values of the spectral centroids in the last 20 trials were significantly different from baseline values [$t_{(11)} = -8.753$, $P < 0.001$ for AS_B; $t_{(11)} = -7.909$, $P < 0.001$ for AS_F; $t_{(11)} = -7.725$, $P < 0.001$ for A; $t_{(11)} = -2.349$, $P < 0.04$ for CTL; P values are corrected; one-sample t test]. The change in the control condition may be due to a variability in a short trial period caused by repetition of speech production, because a fluctuation is seen across trials in CTL in Fig. 2*A*. Indeed, a significant change was not observed in a longer trial period (50 trials) in the control condition [$t_{(11)} = -1.435$, $P = 0.179$; one-sample t test]. One-way ANOVA showed a reliable difference among the conditions [$F_{(3,44)} = 9.567$, $P < 0.001$]. Post hoc tests showed that the control condition was significantly different from the other three conditions in which altered auditory feedback was applied ($P < 0.001$ for AS_B vs. CTL; $P < 0.001$ for AS_F vs. CTL; $P < 0.001$ for A vs. CTL; P values are corrected). In addition, there was no significant difference between each pair of the altered auditory feedback conditions ($P = 0.352$ for AS_B vs. AS_F; $P = 0.977$ for AS_B vs. A; $P = 0.395$ for AS_F vs. A; P values are uncorrected). The results indicate that the amount of adaptation at the end of the training was similar across the three perturbation conditions regardless of the difference in the somatosensory input during the training.

We further examined the adaptation rate in the ramp phase. As shown in Fig. 2*A*, the change of spectral centroid in the ramp phase (31–60th trials) was faster in the backward skin stretch condition than in the other two altered auditory feedback conditions. These patterns of changes in the ramp phase were tested using a linear mixed-effect model as the slope effect of the trial numbers. The results showed significant differences among the conditions [$F_{(2,32.923)} = 4.478$, $P < 0.02$] and the trial numbers [$F_{(1,32.615)} = 32.252$, $P < 0.001$]. The interaction between these two was significant [$F_{(2,32.611)} = 3.855$, $P < 0.04$]. Post hoc tests on the interaction effect showed that a pattern of changes across the trial numbers was significantly different between the backward and forward skin stretch conditions and between the backward skin stretch and the altered auditory feedback alone conditions ($z = 2.513$, $P < 0.04$ for AS_B vs. AS_F; $z = 2.281$, $P < 0.05$ for AS_B vs. A; P values are corrected). There was no reliable difference between forward skin stretch and altered auditory feedback alone ($z = 0.218$, corrected $P = 0.828$ for AS_F vs. A). These results suggest that the direction of the somatosensory perturbation modulated the rate of auditory feedback adaptation in the ramp phase.

The proportion of the subjects that adapted to altered auditory feedback depended on the direction of the skin stretch (Table 1). The proportion in the backward stretch condition (52%) was smaller than those in the altered auditory feedback alone and forward stretch conditions (75% for AS_F and 86% for A). The χ^2 tests with Holm–Bonferroni correction revealed a significant difference between backward stretch and altered auditory feedback alone [$\chi^2_{(1)} = 6.311$, corrected $P < 0.03$ for AS_B vs. A] and no reliable difference between forward stretch and altered auditory feedback alone [$\chi^2_{(1)} = 0.857$, corrected $P = 0.355$ for AS_F vs. A]. This indicates that the backward skin stretch interfered with adaptation to altered auditory feedback.

We assessed the perceptual shift due to speech motor adaptation by comparing perceptual thresholds before and after speech training. Figure 3*A* shows psychometric functions of representative subjects in each condition, before and after training. In the control condition (CTL in Fig. 3*A*), the perceptual threshold shifted in the /ʃ/ direction after training, indicating that the subjects perceived the sounds as more /s/-like. This may be because repeated exposure to a specific phoneme attenuated sensitivity to the phoneme, which is known as the selective adaptation effect (Eimas and Corbit 1973; Shiller et al. 2009). In contrast, the change in threshold following training was smaller in the altered auditory feedback condition (A and CTL in Fig. 3*A*). This effect of the altered auditory feedback training is consistent with previous studies (Lametti et al. 2014; Shiller et al. 2009) showing that the perception of speech sounds shifts in a direction opposite to the auditory perturbation.

When facial skin stretch was applied together with altered auditory feedback, the effect on auditory perception was modulated depending on the direction of the stretch. The perceptual shift in the backward skin stretch condition was different from that in the altered auditory feedback alone condition but similar to that in the control condition (AS_B in Fig. 3*A*). On the other hand, the shift in the forward skin stretch condition was similar to that in the altered auditory feedback alone condition (AS_F in Fig. 3*A*). The mean (±SE) perceptual shift in each condition is summarized in Fig. 3*B*. One-way ANOVA showed that the amount of the perceptual shift was significantly different across the four conditions [$F_{(3,44)} = 3.833$, $P < 0.02$]. Post hoc tests showed that the difference between the altered auditory feedback alone and the CTL condition was significant (corrected $P < 0.03$). Similar to the condition with the altered auditory feedback alone, the forward skin stretch condition was significantly different from the control (corrected $P < 0.05$). There was no significant difference between the altered auditory feedback alone condition and the forward skin stretch condition (corrected $P = 0.818$). No effect of the forward skin stretch may be due to a saturation of the perceptual change following adaptation. Because the difference in perceptual shift between the control and the altered auditory feedback alone condition (–325.21 ± 65.43 Hz) was comparable to the just noticeable difference (the perceptual acuity) between /s/ and /ʃ/ (270.95 ± 16.49 Hz), the perceptual change might be already enough large. In contrast to the clear differences between the CTL condition and each of AS_F and A, the backward skin stretch condition was not reliably different from the CTL or the forward skin stretch conditions ($P = 0.542$ for AS_B vs. CTL; $P = 0.167$ for AS_B vs. AS_F; P values are corrected). A difference between the backward skin stretch and the altered auditory feedback alone conditions was marginally significant (corrected $P = 0.097$ for AS_B vs. A). This suggests that backward skin stretch did not have the same effect on perception as the forward skin stretch condition, but rather the perceptual shift in the backward skin stretch condition was similar to that in the control condition.

We also assessed the changed in perceptual acuity following adaptation. One-way ANOVA revealed that there was no reliable changes of the acuity following adaptation across all conditions [$F_{(3,44)} = 0.63$, $P = 0.602$]. This suggests that speech audio-motor adaptation adjusts perceptual boundary of speech sounds rather than tunes sensitivity of the perception.

Baseline perception of the fricative sounds was tested to ensure no difference in abilities to perceive the fricative sounds across the conditions. One-way ANOVAs were applied to the perceptual threshold and acuity in the first perceptual test. There was no significant difference in each measurement across the conditions [$F_{(3,44)} = 0.07$, $P = 0.974$ for the perceptual threshold; $F_{(3,44)} = 1.61$, $P = 0.202$ for the perceptual acuity]. This suggests that the results in the speech training and perceptual tests cannot be accounted for by intercondition variabilities in baseline perception of the fricative sounds.

In correlation analyses, as observed in previous studies (Lametti et al. 2014), we did not find reliable correlations between the amounts of adaptation and the perceptual shift following adaptation in any of the four conditions [$r_{(11)} = 0.503$, $P = 0.099$ for AS_B; $r_{(11)} = -0.105$, $P = 0.749$ for AS_F; $r_{(11)} = 0.469$, $P = 0.127$ for A; $r_{(11)} = -0.273$, $P = 0.391$ for CTL; Spearman's correlation, P values are uncorrected]. This suggests that the perceptual shift following adaptation is not predictable by somatosensory-motor complex associated with adaptation in each condition. Note that somatosensory inflow is usually coupled with motor outflow, which was measured as the amount of adaptation in this study. However, because of the somatosensory modulations, the amount of adaptation does not necessarily represent total somatosensory inflow during speech production.

We further examined how the skin stretch manipulations affected a relationship between the amount of adaptation and the perceptual shift following adaptation by testing a correlation between these two measures for data pooled from all conditions. If a direction of the skin stretch systematically modulates both adaptation and the perceptual change following adaptation, then there is a relationship between the amount of adaptation and the perceptual change. A correlation analysis detected no significant relationship between the two measures [$r_{(47)} = -0.116$, $P = 0.430$; Spearman's correlation]. This supports our finding that the somatosensory stimulation modulated the perceptual shift following adaptation, but not the adaptation in production at the end of the training.

In sum, despite a nonsignificant effect of somatosensory modulation on the amount of adaptation in production at the end of the training, there was a significant effect of somatosensory modulation on the perceptual changes that accompany adaptation.

Follow-up experiment. The facial deformation due to skin stretch may change articulatory configurations physically and result in a change in the acoustics of produced sounds. This effect, which is irrelevant to speech motor adaptation, may be included in the acoustical data obtained in the main experiment. To assess this possibility, we carried out the separate follow-up experiment. We compared the spectral centroid of produced sound between conditions with and without the skin stretch and between the two stretch directions. Figure 4 shows relative differences in the spectral centroid between the perturbed and unperturbed trials in each skin stretch direction. A linear mixed-effect model showed no significant difference between the two skin stretch conditions [$F_{(1,22.067)} = 1.26$, $P = 0.274$]. The intercept in a linear mixed-effect model was not reliably different from zero [$F_{(1,22.044)} = 0.89$, $P = 0.357$], indicating no reliable difference in the spectral centroid value between the perturbed and unperturbed trials. We concluded

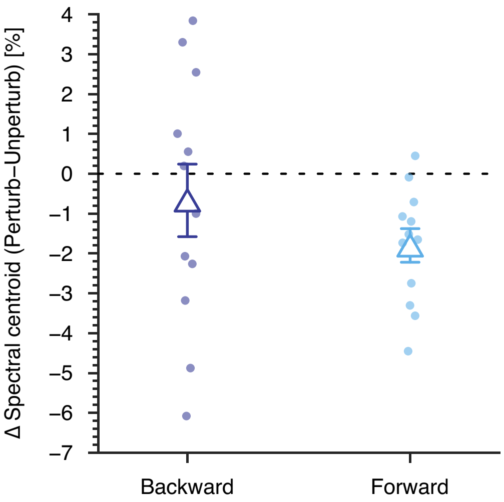

Fig. 4. Effects of unexpected skin stretch on the spectral centroid value of /ʃ/ sounds. The effect was measured as the proportionate change in the spectral centroid value (ΔSpectral centroid) of the perturbed trials relative to the unperturbed trials (Perturb–Unperturb). Mean values and SE are represented by triangles and error bars. Each dot represents an individual value.

that there is little or no effect of the skin stretch on the spectral centroid value.

DISCUSSION

Speech motor adaptation to altered auditory feedback changes the sensorimotor system and results in changes to the perception of speech sounds (Lametti et al. 2014; Schuerman et al. 2017; Shiller et al. 2009). This plastic change in perception is associated with adapted speech production rather than changes in auditory inflow (Lametti et al. 2014; Schuerman et al. 2017). This suggests that motor outflow and/or somatosensory inflow in production are key to the tuning of auditory perception of speech. Although somatosensory inflow is usually coupled with motor outflow, the somatosensory perturbation used in the present study modulates somatosensory inflow in the absence of a large change in motor outflow and enabled us to examine the somatosensory system separately from the motor system. In this study, we assessed a role of somatosensory inflow in the plastic changes in auditory perception of speech that are associated with audio-motor adaptation. The somatosensory perturbation produced by facial skin stretch was expected to enhance and attenuate somatosensory inflow in the production of a target fricative sound, depending on the direction of the perturbation. Our findings suggest that the attenuation of somatosensory inflow produced by backward stretch reduced the perceptual shift following adaptation, whereas the enhancement produced by forward stretch did not alter the shift. Given that motor outflow associated with adaptation was not different at the end of the training in the two skin stretch directions, it is somatosensory inflow rather than motor outflow that serves to tune the processing of auditory perception of speech.

Speech science has been extensively focused on the involvement of the motor system in auditory perception of speech (Liberman et al. 1967; Liberman and Mattingly 1985). In line with this perspective, the perceptual change following adaptation was thought to be associated with an audio-motor interaction (Lametti et al. 2014; Schuerman et al. 2017; Shiller et al. 2009). In the present study, the amount of final adaptation was not reliably different across the three somatosensory conditions, indicating that the motor flow acquired throughout the adaptation task was comparable across the three conditions. Given that speech perceptual measures differed according to the somatosensory condition, the adapted motor outflow may not be the source of the perceptual shift.

A potential role of somatosensory input in auditory perception of speech has been previously demonstrated (Gick and Derrick 2009; Ito et al. 2009). When the facial skin is deformed during the identification of a vowel contrast in a manner similar to that which occurs during speech articulation, perception is systematically biased toward the vowel related to the orofacial somatosensation (Ito et al. 2009). Considering this somatosensory-auditory interaction in auditory perception of speech, a similar perceptual bias could be caused by somatosensory inflow during speech, and its aftereffect would appear in the perceptual tests following adaptation. Based on this idea, our findings can be interpreted as follows. Speech motor adaptation was expected to lead to greater lip protrusion for /ʃ/ to compensate for the auditory feedback alteration, and thereby evoked somatosensation during the production of the more salient /ʃ/ than would be associated with unaltered auditory feedback. As a result, the change in somatosensation for the more salient /ʃ/ presumably biased subjects' perception toward the /ʃ/ category. The somatosensory perturbation produced by skin stretch was designed to attenuate somatosensory inflow during speech motor adaptation in one case, resulting in reduction of the perceptual shift observed in the altered auditory feedback alone condition. In the other case, the enhancement of the somatosensory inflow was expected to cause a larger perceptual shift following the adaptation. As shown in Fig. 3*B*, the attenuation of the perceptual shift induced by altered auditory feedback was observed when the skin was stretched in a backward direction. However, we did not observe the predicted enhancement effect presumably because the perceptual shift toward /ʃ/ following the adaptation was already quite large and perception could not be further biased toward /ʃ/.

The perceptual change by speech motor training would be local and mostly occur in the sounds that were in the speech training. The previous study showed that the perceptual change following speech audio-motor adaptation occurred only in the category that subjects produced (Lametti et al. 2014). As was observed in the control condition in the present study, the perception could be also biased selectively toward the sound that subjects repetitively heard, which is known as selective adaptation effect (Eimas and Corbit 1973). We thus expect that the perceptual change in the present study would not generalize into the sounds that were not in the speech training.

In contrast to the perceptual changes observed following adaptation, neither enhancement nor attenuation of somatosensory inflow affected the amount of adaptation measured at the end of the training. It should be noted that the absence of a significant effect of the skin stretch perturbation at the end of the training does not mean that somatosensory signals are not incorporated into the process of adaptation to auditory feedback. Rather, somatosensory effect on speech production was found in the adaptation rate. The backward skin stretch increased the adaptation rate compared with the other two conditions. With the assumption that the previously observed somatosensory effect on auditory perception of speech (Ito et al. 2009) occurred in auditory processing during speech production, auditory feedback might be perceived as less /ʃ/-like with backward skin stretch, which is expected to attenuate somatosensory inflow, and hence augment auditory feedback errors. This possible difference in error processing may be reflected in faster adaptation rates. Although this interpretation may also suggest that larger perceived errors that could be caused by forward skin stretch would induce larger amounts of adaptation, this was not the case, presumably because the amount of adaptation is not determined simply by the amount of auditory error, as seen in a partial compensation to auditory perturbations (Houde and Jordan 1998; Lametti et al. 2012, 2014; Purcell and Munhall 2006; Shiller et al. 2009; Villacorta et al. 2007). Speech production involves both auditory and somatosensory feedback, and there is a trade-off between errors in these two domains (Katseff et al. 2012; Lametti et al. 2012). A larger compensation for auditory perturbation leads to a deviation from usual articulations, resulting in a larger error in the somatosensory domain. This conflict can impede complete compensation in individual modality. In addition, an auditory target for the production of speech sound would not be specific formant frequencies but a region representing the phoneme in auditory perceptual domain (Niziolek et al. 2013; Perkell 2012). In this sense, there is no need to correct auditory errors as long as those errors are within a target category.

Proportions of the subjects that adapted to altered auditory feedback varied across the three perturbation conditions. In particular, the backward skin stretch interfered with adaptation to altered auditory feedback. This may be due to differences in sensory preference in sensorimotor adaptation as shown by Lametti et al. (2012). They demonstrated that when subjects received auditory and somatosensory perturbations together, a large proportion of subjects adapted to one perturbation or the other. Fifty-three percent of subjects adapted only to the auditory feedback, 21% adapted only to the somatosensory feedback, and 26% adapted to both perturbations. In the present study, possibly as a result of sensory preference, subjects who could potentially adapt to both audition and somatosensation might have failed to adapt properly to auditory feedback alteration when the somatosensory inflow was disturbed by application of perturbations that opposed the normal movement direction. The lower proportion may comprise only those subjects who have a preference to audition. This would be consistent with our finding that the proportion of adapted subject in the backward skin stretch condition (52%) is similar to the proportion of the subjects who adapted to auditory perturbation alone in the Lametti et al. study (53%).

In the speech training, the skin stretch was applied throughout the training, whereas auditory feedback was altered from the ramp phase. A pattern of the amount of adaptation at the end of the training is unlikely to be accounted for by the different schedules of these two manipulations. The effect of the skin stretch on the amount of adaptation was observed in the ramp phase but disappeared in the hold phase. This implies that the somatosensory modulation would affect an initial phase of adaptation or adaptation to a small magnitude of auditory feedback perturbation.

The observed perceptual shift might be due to perceptual adaptation caused by repeated exposure to a specific auditory feedback (selective adaptation effect; Eimas and Corbit 1973) given that subjects produced the same phrase "a shed" throughout the speech training. Although the selective adaptation effect was observed in the control condition, it does not fully account for our results. In the speech training with altered auditory feedback, there was no significant difference in the amount of final adaptation among the three conditions (AS_B, AS_F, and A), indicating that subjects in these conditions received similar auditory feedback in the hold phase of the training. Nevertheless, a pattern of the perceptual shift relative to the shift in the control condition was different among these conditions. This suggests that somatosensory inputs during the speech motor training recalibrate speech perception.

A cross-modal interaction between auditory and somatosensory cortex has been found in previous studies other than those involving speech. Somatosensory signals arising from the hand propagate and coactivate auditory cortex (Foxe et al. 2002; Nordmark et al. 2012; Perez-Bellido et al. 2018; Schürmann et al. 2006), possibly through the anatomical connectivity between the caudal belt auditory area and somatosensory cortex (Fu et al. 2003; Ro et al. 2013). Although further investigations are required to assess the relationship between speech and nonspeech processing, this neural cross-modal interaction may provide the neuroanatomical substrate through which changes in auditory perception of speech are induced by somatosensory inflow in sensorimotor learning.

Cross-modal interactions are highly feature specific; for example, auditory and tactile signals can affect one another's perception in terms of frequency of tactile vibration and sound stimulation (Convento et al. 2018; Crommett et al. 2017; Yau et al. 2009). In speech, somatosensory inputs affect auditory perception of speech systematically according to somatosensory-auditory pairing (Gick and Derrick 2009; Ito et al. 2009). This feature-specific audio-somatosensory association can be built by co-occurrences of auditory and somatosensory signals during speech production and learning. Once this association has been established, auditory perception could be conditioned and affected by somatosensory signals without motor outflow. In the present study, this association can be modified by sensorimotor adaptation training. Due to the modified association, the same sound can be perceived in a different way even in the absence of any additional somatosensory inputs.

In conclusion, the present study investigated a role of the somatosensory system in the perceptual shift that follows adaptation to auditory feedback. We found that somatosensory inputs changed the perceptual shift following adaptation, whereas its influence was limited in the adaptive behavior finally acquired in the training. These findings suggest that the somatosensory inflow, rather than the motor outflow, may play an important role in tuning the processing of auditory perception of speech during speech motor adaptation.

GRANTS

This study was supported by National Institute on Deafness and Other Communication Disorders Grant R21-DC-013915.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

H.O. and T.I. conceived and designed research; H.O. performed experiments; H.O. and T.I. analyzed data; H.O. and T.I. interpreted results of experiments; H.O. and T.I. prepared figures; H.O. and T.I. drafted manuscript; H.O. and T.I. edited and revised manuscript; H.O. and T.I. approved final version of manuscript.

REFERENCES

Andersen N. On the calculation of filter coefficients for maximum entropy spectral analysis. *Geophysics* 39: 69–72, 1974. doi:10.1190/1.1440413.

Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang* 68: 255–278, 2013. doi:10.1016/j.jml.2012.11.001.

Cai S, Ghosh SS, Guenther FH, Perkell JS. 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: 16483–16490, 2011. doi:10.1523/JNEUROSCI.3653-11.2011.

Convento S, Rahman MS, Yau JM. Selective attention gates the interactive crossmodal coupling between perceptual systems. *Curr Biol* 28: 746–752.e5, 2018. doi:10.1016/j.cub.2018.01.021.

Crommett LE, Pérez-Bellido A, Yau JM. Auditory adaptation improves tactile frequency perception. *J Neurophysiol* 117: 1352–1362, 2017. doi:10.1152/jn.00783.2016.

Eimas PD, Corbit JD. Selective adaptation of linguistic feature detectors. *Cognit Psychol* 4: 99–109, 1973. doi:10.1016/0010-0285(73)90006-6.

Fadiga L, Craighero L, Buccino G, Rizzolatti G. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur J Neurosci* 15: 399–402, 2002. doi:10.1046/j.0953-816x.2001.01874.x.

Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM. Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol* 88: 540–543, 2002. doi:10.1152/jn.2002.88.1.540.

Fu KM, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty PE, Schroeder CE. Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci* 23: 7510–7515, 2003. doi:10.1523/JNEUROSCI.23-20-07510.2003.

Gick B, Derrick D. Aero-tactile integration in speech perception. *Nature* 462: 502–504, 2009. doi:10.1038/nature08572.

Hickok G, Houde J, Rong F. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69: 407–422, 2011. doi:10.1016/j.neuron.2011.01.019.

Houde JF, Jordan MI. Sensorimotor adaptation in speech production. *Science* 279: 1213–1216, 1998. doi:10.1126/science.279.5354.1213.

Ito T, Gomi H. Cutaneous mechanoreceptors contribute to the generation of a cortical reflex in speech. *Neuroreport* 18: 907–910, 2007. doi:10.1097/WNR.0b013e32810f2dfb.

Ito T, Ostry DJ. Somatosensory contribution to motor learning due to facial skin deformation. *J Neurophysiol* 104: 1230–1238, 2010. doi:10.1152/jn.00199.2010.

Ito T, Tiede M, Ostry DJ. Somatosensory function in speech perception. *Proc Natl Acad Sci USA* 106: 1245–1248, 2009. doi:10.1073/pnas.0810063106.

Katseff S, Houde J, Johnson K. Partial compensation for altered auditory feedback: a tradeoff with somatosensory feedback? *Lang Speech* 55: 295–308, 2012. doi:10.1177/0023830911417802.

Klatt DH. Software for a cascade-parallel formant synthesizer. *J Acoust Soc Am* 67: 971–995, 1980. doi:10.1121/1.383940.

Lametti DR, Nasir SM, Ostry DJ. Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *J Neurosci* 32: 9351–9358, 2012. doi:10.1523/JNEUROSCI.0404-12.2012.

Lametti DR, Rochet-Capellan A, Neufeld E, Shiller DM, Ostry DJ. Plasticity in the human speech motor system drives changes in speech perception. *J Neurosci* 34: 10339–10346, 2014. doi:10.1523/JNEUROSCI.0108-14.2014.

Lane H, Denny M, Guenther FH, Hanson HM, Marrone N, Matthies ML, Perkell JS, Stockmann E, Tiede M, Vick J, Zandipour M. On the structure of phoneme categories in listeners with cochlear implants. *J Speech Lang Hear Res* 50: 2–14, 2007. doi:10.1044/1092-4388(2007/001).

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. *Psychol Rev* 74: 431–461, 1967. doi:10.1037/h0020279.

Liberman AM, Mattingly IG. The motor theory of speech perception revised. *Cognition* 21: 1–36, 1985. doi:10.1016/0010-0277(85)90021-6.

Maniwa K, Jongman A, Wade T. Acoustic characteristics of clearly spoken English fricatives. *J Acoust Soc Am* 125: 3962–3973, 2009. doi:10.1121/1.2990715.

Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M. The essential role of premotor cortex in speech perception. *Curr Biol* 17: 1692–1696, 2007. doi:10.1016/j.cub.2007.08.064.

Möttönen R, Watkins KE. Motor representations of articulators contribute to categorical perception of speech sounds. *J Neurosci* 29: 9819–9825, 2009. doi:10.1523/JNEUROSCI.6018-08.2009.

Nasir SM, Ostry DJ. Auditory plasticity and speech motor learning. *Proc Natl Acad Sci USA* 106: 20470–20475, 2009. doi:10.1073/pnas.0907032106.

Niziolek CA, Nagarajan SS, Houde JF. What does motor efference copy represent? Evidence from speech production. *J Neurosci* 33: 16110–16116, 2013. doi:10.1523/JNEUROSCI.2137-13.2013.

Nordmark PF, Pruszynski JA, Johansson RS. BOLD responses to tactile stimuli in visual and auditory cortex depend on the frequency content of stimulation. *J Cogn Neurosci* 24: 2120–2134, 2012. doi:10.1162/jocn_a_00261.

Perez-Bellido A, Anne Barnes K, Crommett LE, Yau JM. Auditory frequency representations in human somatosensory cortex. *Cereb Cortex* 28: 3908–3921, 2018. doi:10.1093/cercor/bhx255.

Perkell JS. Movement goals and feedback and feedforward control mechanisms in speech production. *J Neurolinguist* 25: 382–407, 2012. doi:10.1016/j.jneuroling.2010.02.011.

Pulvermüller F, Huss M, Kherif F, Moscoso del Prado Martin F, Hauk O, Shtyrov Y. Motor cortex maps articulatory features of speech sounds. *Proc Natl Acad Sci USA* 103: 7865–7870, 2006. doi:10.1073/pnas.0509989103.

Purcell DW, Munhall KG. Adaptive control of vowel formant frequency: evidence from real-time formant manipulation. *J Acoust Soc Am* 120: 966–977, 2006. doi:10.1121/1.2217714.

Ro T, Ellmore TM, Beauchamp MS. A neural link between feeling and hearing. *Cereb Cortex* 23: 1724–1730, 2013. doi:10.1093/cercor/bhs166.

Schuerman WL, Nagarajan S, McQueen JM, Houde J. Sensorimotor adaptation affects perceptual compensation for coarticulation. *J Acoust Soc Am* 141: 2693–2704, 2017. doi:10.1121/1.4979791.

Schürmann M, Caetano G, Hlushchuk Y, Jousmäki V, Hari R. Touch activates human auditory cortex. *Neuroimage* 30: 1325–1331, 2006. doi:10.1016/j.neuroimage.2005.11.020.

Shen Y, Richards VM. A maximum-likelihood procedure for estimating psychometric functions: thresholds, slopes, and lapses of attention. *J Acoust Soc Am* 132: 957–967, 2012. doi:10.1121/1.4733540.

Shiller DM, Sato M, Gracco VL, Baum SR. Perceptual recalibration of speech sounds following speech motor learning. *J Acoust Soc Am* 125: 1103–1113, 2009. doi:10.1121/1.3058638.

Tremblay S, Shiller DM, Ostry DJ. Somatosensory basis of speech production. *Nature* 423: 866–869, 2003. doi:10.1038/nature01710.

Villacorta VM, Perkell JS, Guenther FH. Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *J Acoust Soc Am* 122: 2306–2319, 2007. doi:10.1121/1.2773966.

Yau JM, Olenczak JB, Dammann JF, Bensmaia SJ. Temporal frequency channels are linked across audition and touch. *Curr Biol* 19: 561–566, 2009. doi:10.1016/j.cub.2009.02.013.