

# Sensory Preference in Speech Production Revealed by Simultaneous Alteration of Auditory and Somatosensory Feedback

Daniel R. Lametti,<sup>1</sup> Sazzad M. Nasir,<sup>2</sup> and David J. Ostry<sup>1,3</sup>

<sup>1</sup>Department of Psychology, McGill University, Montreal, Quebec H3A 1B1, Canada, <sup>2</sup>The Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208, and <sup>3</sup>Haskins Laboratories, New Haven, Connecticut 06511

The idea that humans learn and maintain accurate speech by carefully monitoring auditory feedback is widely held. But this view neglects the fact that auditory feedback is highly correlated with somatosensory feedback during speech production. Somatosensory feedback from speech movements could be a primary means by which cortical speech areas monitor the accuracy of produced speech. We tested this idea by placing the somatosensory and auditory systems in competition during speech motor learning. To do this, we combined two speech-learning paradigms to simultaneously alter somatosensory and auditory feedback in real time as subjects spoke. Somatosensory feedback was manipulated by using a robotic device that altered the motion path of the jaw. Auditory feedback was manipulated by changing the frequency of the first formant of the vowel sound and playing back the modified utterance to the subject through headphones. The amount of compensation for each perturbation was used as a measure of sensory reliance. All subjects were observed to correct for at least one of the perturbations, but auditory feedback was not dominant. Indeed, some subjects showed a stable preference for either somatosensory or auditory feedback during speech.

## Introduction

When we speak, how do we know we are saying our words correctly? The answer seems simple: we listen to the sound of our own voice. This idea—that accurate speech production is maintained by carefully monitoring one's own auditory feedback—is widely held (Lombard, 1911; Lane and Tranel, 1971; Brainard and Doupe, 2000; Perkell et al., 2000). But this explanation ignores the possible role of somatosensory feedback from the movement of the articulators (Tremblay et al., 2003; Ito and Ostry, 2010). From the first words that a child utters, speech sounds are correlated with the movements that produce them (Gracco and Löfqvist, 1994). Somatosensory feedback from orofacial movement could play an important role in monitoring the accuracy of produced speech. In adults who retain intelligible speech after total hearing loss, this seems essential (Lane and Wozniak-Webster, 1991; Nasir and Ostry, 2008). But does somatosensory feedback play a significant role in the speech of healthy adults?

The idea that accurate speech is maintained by auditory feedback is supported by the observation that subjects change the

sound of their voice to compensate for auditory perturbations that alter their speech sounds (Houde and Jordan, 1998; Jones and Munhall, 2005; Purcell and Munhall, 2006a,b; Villacorta et al., 2007; Feng et al., 2011). However, unlike in studies of sensorimotor adaptation and motor learning in limb movement (Shadmehr and Mussa-Ivaldi, 1994; Krakauer et al., 2000), a significant percentage of subjects in auditory studies fail to compensate for auditory perturbations. One possible reason is that, in contrast to the nearly uniform way in which people are observed to use sensory feedback to control limb movements (van Beers et al., 2002), the integration of sensory feedback during speech might differ significantly among individuals. Some individuals may even rely more heavily on somatosensory feedback during the production of some speech sounds (Yates, 1965).

We tested this idea by simultaneously altering auditory and somatosensory feedback during speech production. We placed the two sensory systems in competition to determine the relative reliance on each. To do this, using two experimental paradigms adapted from studies of speech motor learning, both somatosensory and auditory feedback were altered in real time, alone or in combination, as subjects repeated a consonant-vowel-consonant word. A robotic device that caused subtle changes in the movement of the lower jaw altered somatosensory feedback; an acoustical effects processor that lowered the first formant frequency of the vowel sound altered auditory feedback. The amount of compensation for each perturbation was used as a measure of sensory reliance.

We found that all subjects compensated for at least one form of altered sensory feedback. In contrast to the idea that accurate

Received Jan. 27, 2012; revised May 15, 2012; accepted May 19, 2012.

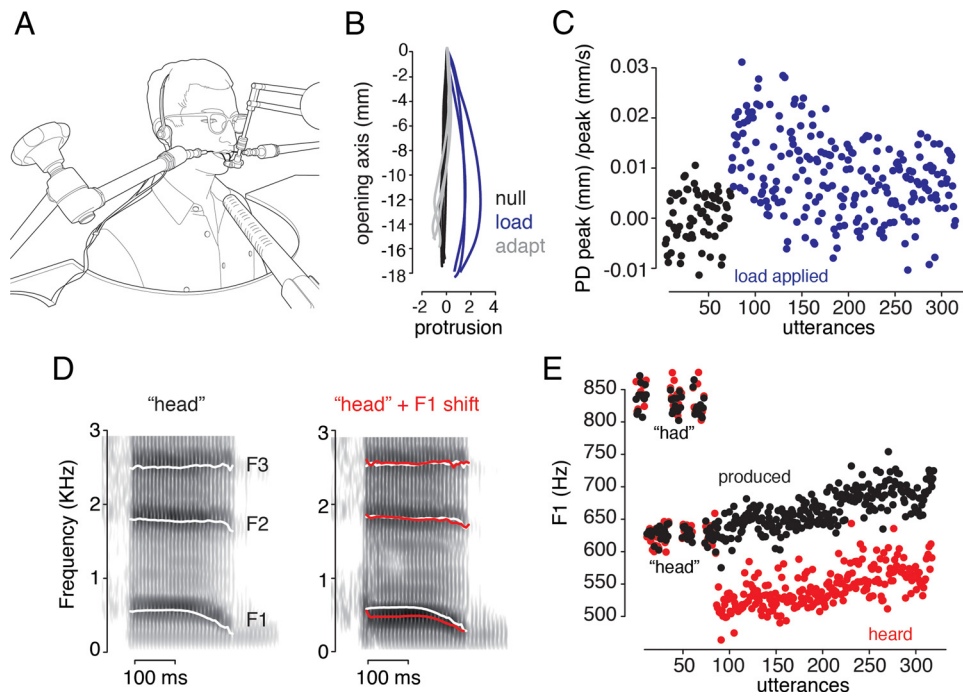
Author contributions: D.R.L., S.M.N., and D.J.O. designed research; D.R.L. and S.M.N. performed research; D.R.L. analyzed data; D.R.L. and D.J.O. wrote the paper.

This work was supported by grants from the National Institute on Deafness and Other Communication Disorders (DC-04669), the Natural Sciences and Engineering Research Council of Canada, and Le Fonds Québécois de la Recherche sur la Nature et les Technologies. We thank M. Darainy, G. Houle, A. Mattar, L. Richer, and A. Rochet-Capellan for helpful comments and suggestions about the experiments detailed here.

Correspondence should be addressed to David J. Ostry, 1205 Dr. Penfield Avenue, Department of Psychology, McGill University, Montreal, Quebec H3A 1B1, Canada. E-mail: david.ostry@mcgill.ca.

DOI:10.1523/JNEUROSCI.0404-12.2012

Copyright © 2012 the authors 0270-6474/12/329351-08\$15.00/0



**Figure 1.** Experimental methods. **A**, Illustration of a subject in the experimental setup. Custom-made dental appliances attached the lower jaw to a small robot and the upper jaw to two articulating arms that held the head motionless. Subjects wore headphones and spoke into a microphone. **B**, The robot applied a load that displaced the jaw outward (blue vs black lines) during the downward movement associated with the word “head.” With practice, subjects compensated for this somatosensory perturbation (gray lines). **C**, Adaptation to the somatosensory perturbation was quantified using the perpendicular deviation (PD) of the jaw at peak velocity (peak) divided by peak velocity. The robot caused a significant increase in deviation (blue vs black points) that was decreased over the course of training. **D**, First formant (F1), second formant (F2), and third formant (F3) frequencies for a male speaker saying the utterance “head.” The first formant frequency of the vowel sound was shifted downward in real time (right, red lines vs white) altering auditory feedback. **E**, Subjects compensated for a decrease in the F1 frequency of what they heard (red points) by increasing the F1 frequency of what they produced (black points).

speech production is largely dependent upon auditory feedback, we show that there is an inverse relationship between reliance on auditory versus somatosensory feedback: the more subjects compensate for one perturbation the less they compensate for the other. By applying the two perturbations alone and then in combination we show that this inverse relationship is the result of a preferential reliance on either auditory or somatosensory feedback during speech production.

## Materials and Methods

**Subjects, apparatus, task.** Seventy-five native English speakers (23 males) between the ages of 18 and 40 participated in the experiments. The McGill University Faculty of Medicine Institution Review Board approved the experimental protocol. Test subjects reported normal speech and hearing and gave informed consent before participating. All subjects were naive to the experimental manipulation upon initial recruitment.

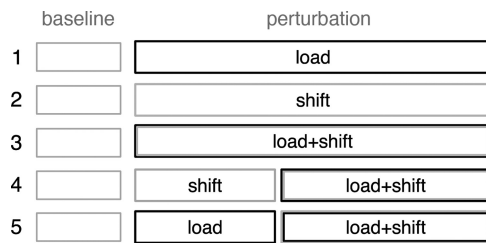
Subjects were seated during testing. Custom-made acrylic and metal dental appliances were individually constructed to fit on the upper and lower teeth of each subject (Tremblay et al., 2003). The lower appliance was attached to a small robotic device (Phantom 1.0, SensAble Technologies) via a rotary connector fixed to a force torque sensor (ATI Industrial Automation). The robot tracked the movement of the jaw and could also apply forces. The upper appliance connected the upper jaw to two articulated arms that held the head motionless during the experiment. Subjects also wore headphones (Stax SR001-MK2 electrostatic) and spoke into a unidirectional microphone (Sennheiser). Figure 1A illustrates the experimental setup.

During the experiment, the word “had” or “head” was displayed on a computer monitor. Subjects were instructed to repeatedly speak the displayed word at a comfortable pace until it was removed from the computer screen. They were also instructed to bring their mouth to a complete close between the individual utterances. On average, the displayed word was repeated 11 times (SD, 1) before the experimenter re-

moved the word from the display. These 11 utterances were considered one “block” of trials.

**Somatosensory and auditory perturbations.** We perturbed somatosensory feedback during speech production by using the robot to alter the movement path of the lower jaw. To do this, the robot applied a load that pulled the jaw outward (Fig. 1B) in a direction perpendicular to the movement path. The applied force depended on the equation  $F = k|v|$ , where  $F$  is the applied force in newtons,  $k$  is a scaling factor, and  $v$  is the instantaneous velocity of the jaw in millimeters per second. The scaling factor was set to 0.02. For the 61 subjects who received a somatosensory perturbation during speech, the average peak force applied to the jaw was 2 N (SD, 0.7 N). Males, who were larger and thus made bigger, faster movements, received an average peak force of 2.25 N (SD, 0.75 N); females received an average peak force of 1.89 N (SD, 0.66 N).

We perturbed auditory feedback during speech by altering the sound of the voice in near real-time. Vocal tract resonances are generated during the production of vowel sounds. These resonances, called formants, are seen as peaks in the spectral frequency distribution of vowels (Fig. 1D). Each vowel sound has a unique set of formants. The first formant, or F1, contains the most acoustical energy and, along with the second formant, F2, is critical in distinguishing vowels. But by altering F1 alone, one vowel can be made to sound like another (Delattre et al., 1952). As in Rochet-Capellan and Ostry (2011), an acoustical effects processor (VoiceOne, TC-Helicon Vocal Technologies) and filters were used to shift F1 downward, while leaving the other formants and the fundamental frequency (F0) unchanged. The resulting signal was then mixed with 70 dB speech-shaped masking noise and played back to subjects through the headphones. The F1 shift was applied during repetitions of the word “head.” The effects processor was set to produce an average downward F1 shift of  $\sim 125$  Hz in the vowel sound in “head” (Fig. 1D), although the amount of shift delivered by the processor scaled with subjects’ baseline F1 frequency. For the 61 subjects that received an auditory perturbation during speech, F1 was shifted down by an average of 125.36 Hz (SD, 28



**Figure 2.** There were five experimental conditions involving 75 subjects. Fourteen subjects experienced the somatosensory perturbation (“load”) alone (Experiment 1). Fourteen subjects experienced the auditory perturbation (“shift”) alone (Experiment 2). Fourteen subjects experienced both the auditory and somatosensory perturbations (“load + shift”) at the same time (Experiment 3). Seventeen subjects experienced the auditory perturbation first, followed by both perturbations at the same time (Experiment 4). Sixteen subjects experienced the somatosensory perturbation first, followed by both perturbations at the same time (Experiment 5).

Hz). Males, who typically had a lower baseline F1 than females, got a downward F1 shift of 104.56 Hz (SD, 12 Hz); females got a downward F1 shift of 134.77 Hz (SD, 27 Hz).

**Experimental procedures.** Before starting the experiment, subjects were asked to produce the words “had” and “head” 10 times each to familiarize themselves with speaking while attached to the robot and hearing their voice through the headphones. Subjects then produced six “baseline” blocks, switching from “had” to “head” between blocks. Twenty-five training blocks followed baseline blocks in which somatosensory and auditory perturbations were applied alone or in combination as subjects repeated just the word “head.” Although subjects only said the word “head” when the perturbations were applied, production of the word “had” was incorporated into the baseline blocks to give subjects a range of sound and movement experience before application of the perturbations.

Test subjects were divided into five groups (Fig. 2). Following baseline trials, the first group of subjects ( $n = 14$ ; four males) received only a somatosensory perturbation (or load) during the 25 training blocks following baseline. The second group ( $n = 14$ ; four males) received only an auditory perturbation (or shift) during training. The third group ( $n = 14$ ; four males) received both a somatosensory and an auditory perturbation (load plus shift) during training. The fourth group ( $n = 17$ ; five males) received an auditory perturbation in the first 10 blocks following baseline, and then received both an auditory and a somatosensory perturbation for the next 15 blocks. The fifth group of subjects ( $n = 16$ ; six males) received a somatosensory perturbation for the first 10 blocks following baseline, and then both an auditory and a somatosensory perturbation for the remaining 15 training blocks.

**Kinematic analysis.** The robot sampled jaw position at 1 kHz with a resolution of 0.03 mm. Jaw velocity was computed using numerical differentiation. As with previous studies of speech motor learning performed in our laboratory (Tremblay et al., 2003; Nasir and Ostry, 2008), only the opening movement of the jaw was analyzed. Movement start and end were scored at the point where jaw velocity exceeded or fell below 10% of peak movement velocity.

To quantify the way in which somatosensory perturbations altered movements, we examined how the robot altered the motion path of the jaw. At peak velocity, we computed the perpendicular deviation from a straight-line path joining the start and the end of movement. Because the amount of force applied by the robot depended on the velocity of the jaw, and because, unlike in studies of limb movement, movement velocity could not be tightly controlled (subjects were simply instructed to speak normally), we divided perpendicular deviation at peak velocity by peak velocity (Fig. 1C). This gave a measure of movement deviation that looked qualitatively similar to standard measures but accounted for differences in movement velocity and hence applied force.

Subjects were classified as having adapted to the somatosensory perturbation if there was a significant decrease in movement deviation over the course of trials in which the load was applied;  $t$  tests were used to see whether the mean deviation of the last 45 perturbed movements was

significantly less than the mean deviation of perturbed movements 5–49. The first four perturbed movements were excluded from this analysis because there was a transient reduction in jaw deflection upon initial load application, presumably due to an increase in jaw stiffness. Specifically, load-induced movement deviation in the first four perturbed trials averaged 0.54 mm, while load-induced movement deviation averaged 1.05 mm for perturbed trials 5–9, and 0.93 mm for perturbed trials 10–50.

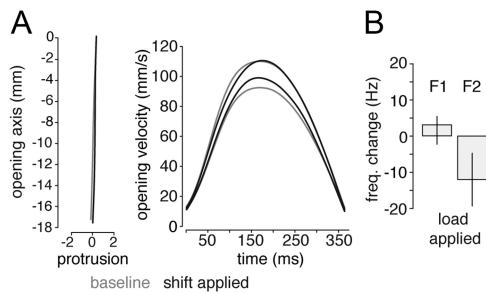
When examining changes in movement deviation, the mean deviation of baseline trials was subtracted from measures of movement deviation on a per-subject basis. This normalization procedure removed between-subject differences in baseline performance. For statistical tests of kinematic performance in different experimental conditions, movement deviation was calculated as the mean deviation of 30 movements at the points of interest—mainly, before the introduction of a perturbation, after the introduction of perturbation (without the first 4 trials), and at the end of training—and averaged over subjects. Split-plot ANOVAs with Bonferroni corrected *post hoc* tests were used to examine differences between these points of interest.

**Acoustical analysis.** Three channels of acoustical data were digitally recorded at 10 kHz. The first channel contained what subjects produced—what subjects said into the microphone. The second channel contained the F1-shifted audio that came out of the acoustical effects processor. The third channel contained what subjects heard—F1 shifted audio mixed with speech-shaped masking noise. The first formants of both the produced and heard vowels were calculated using the software program Praat. Praat automatically detected vowel boundaries and calculated F1 based on a 30 ms window at the center of the vowel (Rochet-Capellan and Ostry, 2011).

Subjects were classified as having adapted to the auditory perturbation if there was a significant increase in their F1 production frequency while the F1 they heard was shifted down (Fig. 1E);  $t$  tests were used to test whether the mean value of the produced F1 frequency for the last 45 acoustically shifted utterances was significantly greater than the mean F1 of baseline “head” trials. When comparing differences in vocal production in different experimental conditions, the mean F1 frequency of baseline “head” utterances was subtracted from individual F1 values on a per-subject basis. This normalization procedure removed between-subject differences in baseline measures of F1 and in particular corrected for well known differences in F1 between males and females. For statistical tests of performance in different experimental conditions, F1 was calculated as the mean value of F1 over 30 utterances at points of interest—mainly before the introduction of a perturbation, after the introduction of perturbation (without the first four trials), and at the end of training—and averaged over subjects. Split-plot ANOVAs with Bonferroni-corrected *post hoc* tests were used to examine individual differences.

**Quantifying adaptation.** For subjects who received both somatosensory and auditory perturbations, percentage measures of adaptation were computed for each perturbation on a per-subject basis. In the case of the somatosensory perturbation, the mean deviation of baseline movements was subtracted from the mean deviation of perturbed movements 5–49, giving a measure of how much the robot perturbed the jaw at the start of training. The mean deviation of the last 45 perturbed movements was subtracted from the mean deviation of perturbed movements 5–49, giving a measure of how much a subject compensated for the load. The measure of load compensation was then divided by the initial measure of how much the robot perturbed the jaw at the start of training and multiplied by 100 to give a percentage measure of how much a subject compensated for the somatosensory perturbation.

For the auditory perturbation, the amount of acoustical shift was determined by subtracting shifted F1 values from produced F1 values. In this case, the mean-shifted F1 and the mean-produced F1 were calculated from shifted utterances 5–49; the difference between these measures gave the amount of acoustical shift at the start of training. The amount of compensation for the shift was determined by subtracting the produced F1 for the baseline “head” utterances from the produced F1 for the last 45 shifted utterances. This value was then divided by the amount of the shift and multiplied by 100 to give a percentage measure analogous to that used for the somatosensory perturbation.



**Figure 3.** The auditory and somatosensory perturbations were independent. **A**, The introduction of the auditory perturbation did not alter the movement path of the jaw (left) or the velocity of the jaw (right). **B**, The introduction of the somatosensory perturbation did not alter the first formant frequency (F1) or the second formant frequency (F2) of heard speech.

## Results

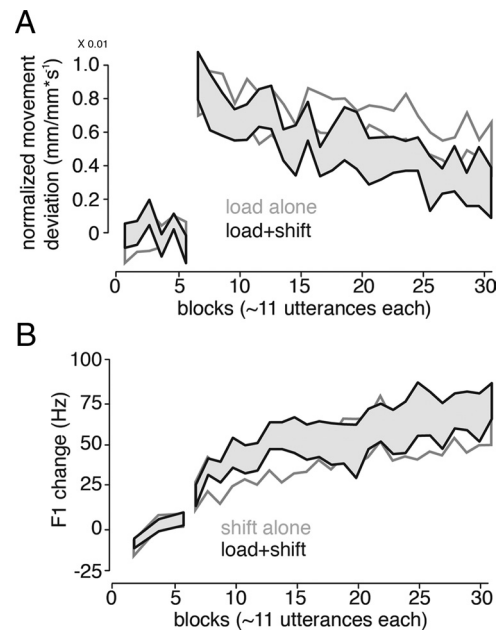
Subjects were divided into five experimental conditions (Fig. 2) in which somatosensory feedback and auditory feedback were altered in real time, either alone or in combination, as the consonant–vowel–consonant utterance “head” was repeated. Auditory feedback was altered by decreasing the F1 frequency of the vowel sound in “head” (Fig. 1*D*); somatosensory feedback was altered by displacing the lower jaw outward during movements associated with production of “head” (Fig. 1*B*). An increase in F1 frequency was used as a measure of compensation for the auditory perturbation (Fig. 1*E*); a decrease in robot-induced movement deviation was used as measure of compensation for the somatosensory perturbation (Fig. 1*C*).

### The effects of the perturbations were independent

Fourteen subjects experienced the somatosensory perturbation alone and 14 different subjects experienced the auditory perturbation alone (Fig. 2, Experiments 1 and 2). The effects of the perturbations were independent of each other—that is, the somatosensory perturbation did not alter the sound of the voice and the auditory perturbation did not alter the movement path of the jaw. Figure 3*A* shows that jaw movement amplitude, curvature, and peak velocity were similar before and after the introduction of the auditory perturbation ( $p > 0.05$  in each case, two-tailed  $t$  test). Figure 3*B* shows that the introduction of the somatosensory perturbation had no effect on F1 and F2 frequencies ( $p > 0.05$  in each case, two-tailed  $t$  test), a finding consistent with previous studies (Tremblay et al., 2003; Nasir and Ostry, 2006, 2008).

### Applying the perturbations at the same time did not affect the amount of compensation

The presence of the acoustical shift did not affect adaptation to the mechanical load nor did the presence of the mechanical load affect adaptation to the acoustical shift. Figure 4*A* shows subjects who adapted to the load. The curves outlined in gray show changes in movement deviation over the course of training for subjects who only received the mechanical load (Fig. 2, Experiment 1); the curves outlined in black show changes in deviation over the course of training for subjects who simultaneously received both the load and the auditory perturbation (Fig. 2, Experiment 3). In each case, seven of 14 subjects met the criterion for somatosensory adaptation, defined as a significant reduction ( $p < 0.05$ ) in load-induced movement deviation over the course of training. Both groups also showed a reduction in movement deviation with training ( $p < 0.01$ , in each case). The presence of



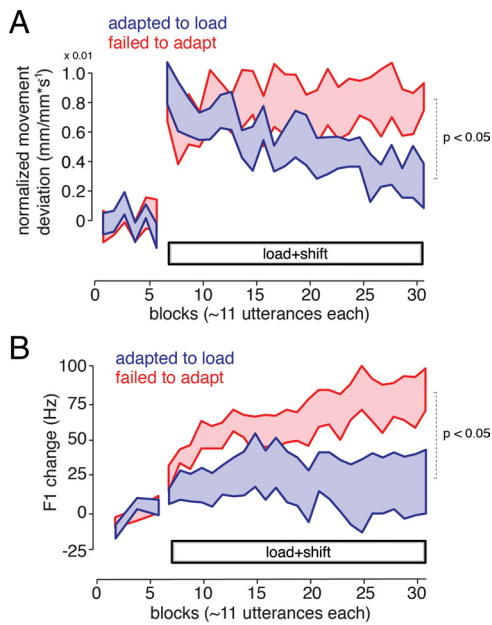
**Figure 4.** Applying the perturbations at the same time did not change the amount of compensation for each perturbation. **A**, Jaw movement deviation over the course of training for the subjects who adapted to the somatosensory perturbation. Curves outlined in gray represent subjects that received just the somatosensory perturbation (Experiment 1). Curves outlined in black represent subjects who received both the somatosensory and auditory perturbations at the same time (Experiment 3). In each case, seven of 14 subjects compensated for the perturbation. Shaded and enclosed regions represent  $\pm 1$  SE. **B**, Change in F1 frequency over the course of training for subjects who adapted to the auditory perturbation. Curves outlined in gray show subjects that received just the auditory perturbation (Experiment 2). Curves outlined in black show subjects that received both the auditory and somatosensory perturbations at the same time (Experiment 3). In each case, 11 of 14 subjects compensated for the perturbation.

the acoustical shift did not increase or decrease the amount of compensation for the load ( $p > 0.05$ ).

Similarly, the presence of the mechanical load did not affect adaptation to the acoustical shift. Figure 4*B* shows subjects who adapted to the auditory perturbation. The curves outlined in gray show changes in F1 frequency over the course of training for subjects who only received the auditory perturbation (Fig. 2, Experiment 2); the curves outlined in black show changes in F1 over the course of training for subjects who simultaneously received both the auditory perturbation and the mechanical load (Fig. 2, Experiment 3). In each case, 11 of 14 subjects met the criterion for adaptation to the auditory perturbation, defined as a significant increase ( $p < 0.05$ ) in produced F1 frequency over the course of training. Both groups also showed an average increase in measures of F1 to compensate for the downward frequency shift ( $p < 0.01$ , in each case). The presence of the mechanical load did not affect how much subjects changed their speech acoustics to compensate for the auditory perturbation ( $p > 0.05$ ).

### Subjects who compensated for the somatosensory perturbation compensated less or not at all to the auditory perturbation

All 14 subjects who simultaneously experienced both the mechanical load and the acoustical shift (Experiment 3) met the criterion for adaptation to at least one of the two perturbations. Did the subjects who compensated for the somatosensory perturbation compensate less for the auditory perturbation? Figure 5*A* shows changes in movement deviation for subjects who adapted to the load (blue curves) and for subjects who did not adapt (red

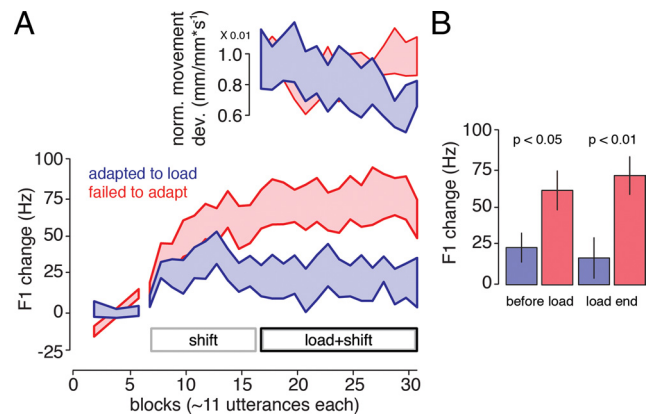


**Figure 5.** Adaptation to the auditory perturbation differed, depending on whether subjects adapted to the somatosensory perturbation. **A**, Jaw movement deviation for subjects who compensated for the somatosensory perturbation (blue curves,  $N = 7$ ) and subjects who did not (red curves,  $N = 7$ ) when both the somatosensory and auditory perturbations were applied simultaneously (Experiment 3). Shaded regions represent  $\pm 1$  SE. **B**, Change in F1 frequency for the same subjects. Subjects who compensated for the somatosensory perturbation failed to compensate for the auditory perturbation.

curves). Figure 5B shows changes in F1 frequency for the same groups of subjects. Subjects who adapted to the somatosensory perturbation did not increase their F1 frequency in response to the auditory shift (blue) as much as subjects who failed to adapt (red) to the somatosensory perturbation ( $p < 0.05$ ). By the end of training, subjects who adapted to the somatosensory perturbation showed no change in F1 frequency ( $p > 0.05$ ). On the other hand, subjects who did not adapt to the somatosensory perturbation increased their F1 frequency to compensate for the acoustical shift ( $p < 0.01$ ).

The results of Experiment 3 suggest that subjects who compensate for the somatosensory perturbation, compared with those who do not, compensate less or not at all for the auditory perturbation. Would these subjects have adapted more to the auditory perturbation if the load had never been applied? In other words, was the failure to adapt to the auditory perturbation caused by sensory competition between auditory and somatosensory feedback? To answer this question, 17 new subjects experienced the auditory perturbation alone before receiving both the somatosensory and auditory perturbations at the same time (Fig. 2, Experiment 4). As in Experiment 3, all subjects met the criterion for adaptation to at least one of the two perturbations.

Figure 6A shows changes in F1 frequency and jaw movement deviation over the course of training. After several baseline blocks, the auditory shift was applied alone and then both the mechanical load and auditory shift were applied at the same time. The bottom panel shows changes in F1 frequency in response to the acoustical shift; the top panel shows movement deviation in response to the mechanical load, starting from the point at which the mechanical load was applied. Again, subjects who adapted to the somatosensory perturbation (blue curves) were compared with subjects who did not (red curves). As in Experiment 3, Figure 6A shows that those who compensated for the mechanical load compensated



**Figure 6.** Adaptation to the auditory perturbation was not affected by the introduction of the somatosensory perturbation. **A**, Change in F1 frequency (bottom) and movement deviation (top) for subjects who compensated for the somatosensory perturbation (blue curves,  $N = 8$ ) and subjects who did not (red curves,  $N = 9$ ). The auditory perturbation was applied before the simultaneous application of the auditory and somatosensory perturbations (Experiment 4). Shaded regions represent  $\pm 1$  SE. The bottom panel shows change in F1 frequency. **B**, Subjects who compensated for the somatosensory perturbation (blue bars) compensated less for the auditory perturbation even before the somatosensory perturbation was applied. Error bars represent  $\pm 1$  SE.

less for the auditory perturbation ( $p < 0.01$ ). Crucially, this difference in F1 frequency was present before the load was applied (Fig. 6B,  $p < 0.05$ ). Subjects who would later adapt to the mechanical load were already adapting less or not at all to the auditory perturbation before the mechanical load was turned on. These subjects responded more to changes in somatosensory feedback during the task than to changes in auditory feedback.

To further examine the idea that subjects show a sensory preference during speech production, 16 new subjects were tested in the opposite order: these subjects experienced the somatosensory perturbation before receiving both the auditory perturbation and the somatosensory perturbation at the same time (Fig. 2, Experiment 5). The goal was to see whether subjects who failed to adapt to the mechanical load in the presence of the acoustical shift would have adapted to the load on its own. As in Experiments 3 and 4, all subjects met the criterion for adaptation to at least one of the two perturbations. Figure 7A compares changes in movement deviation and F1 frequency over the course of training. The bottom panel shows changes in movement deviation; the top panel shows changes in F1 frequency starting at the point at which the acoustical perturbation was applied. Again, subjects who adapted to the somatosensory perturbation (blue curves) were compared with subjects who did not (red curves). Figure 7A shows that there was already a difference between subjects who adapted to the load and those who did not before the acoustical shift was applied. This difference is quantified in Figure 7B, which shows the response to the load before the introduction of the acoustical shift and at the end of training, in the presence of both perturbations. Even before the acoustical shift is applied, there is a difference in the amount of compensation for the mechanical load ( $p < 0.01$ ). This suggests that introduction of the acoustical shift did not alter the response to the mechanical load. But when the acoustical shift was applied, subjects who failed to adapt to the mechanical load adapted to the auditory perturbation to a greater extent than the subjects who had adapted to the load (Fig. 7A, top;  $p < 0.01$ ). This result, in combination with Experiments 3 and 4, provides evidence that subjects show a stable preference for either somatosensory feedback or auditory feedback during speech production.

### A negative correlation was observed between the amount of compensation for each perturbation

In total, 47 subjects in Experiments 3, 4, and 5 had both their auditory and somatosensory feedback simultaneously perturbed during speech. Every subject met the criterion for adaptation to at least one of the two perturbations (Fig. 8A): 53% of subjects ( $n = 25$ ) adapted only to the auditory perturbation; 26% of subjects ( $n = 12$ ) adapted to both the somatosensory and auditory perturbations; and 21% of subjects ( $n = 10$ ) adapted to only the somatosensory perturbation. Figure 8B shows that subjects who adapted more to the somatosensory perturbation adapted less to the auditory perturbation and vice versa. The correlation between the percentage of adaptation to the somatosensory perturbation and the percentage of adaptation to the auditory perturbation was  $-0.54$ , statistically significant at  $p < 0.001$ . For each of the three groups that received both perturbations at the same time, the correlations between the percentages of adaptation to each perturbation were  $r = -0.47$  ( $p = 0.09$ ),  $r = -0.50$  ( $p < 0.05$ ), and  $r = -0.64$  ( $p < 0.01$ ), respectively.

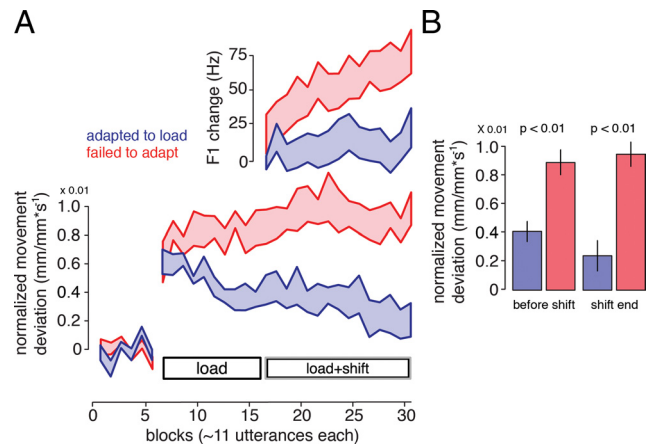
This pattern held for both males and females. Fifteen of the 47 subjects who received both perturbations at the same time were male. Of these, nine (60%) adapted to the somatosensory perturbation and 11 (73%) adapted to the auditory perturbation. The correlation between the percentage of adaptation to the somatosensory perturbation and the percentage of adaptation to the auditory perturbation for males was  $-0.48$  ( $p = 0.07$ ). The remaining 32 subjects who received the two perturbations simultaneously were female. Thirteen (41%) adapted to the somatosensory perturbation and 26 (81%) adapted to the auditory perturbation. The correlation for females between the percentage of adaptation to the somatosensory and auditory perturbations was  $-0.55$  ( $p < 0.01$ ).

As compared with females (see Materials and Methods), males received more of a somatosensory perturbation and less of an auditory perturbation. Even so, correlations across subjects, between the average amount of force delivered upon initial load application and the percentage of somatosensory compensation ( $r = 0.13$ ), and the average initial change in perceived F1 frequency and the percentage of auditory compensation ( $r = 0.12$ ), were not significant ( $p > 0.05$ ). This suggests that differences in the magnitudes of the two perturbations did not play a significant role in how individuals responded.

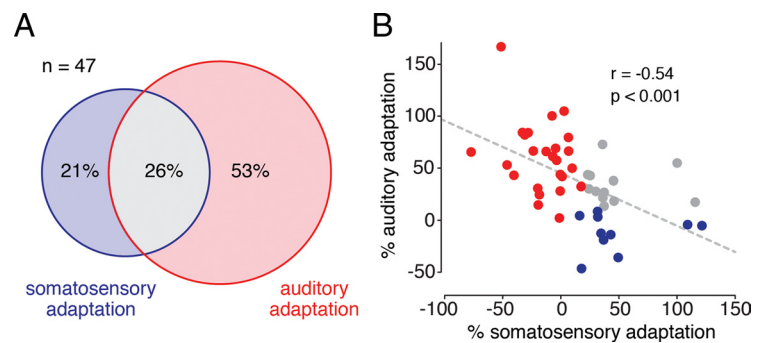
Last, we tried to predict the percentage of adaptation for each of the perturbations based on a number of measures—mainly, baseline F1 frequency and variance in F1 frequency, and baseline jaw opening amplitude and variance in this measure. In each case, we found no significant correlations. One exception was that baseline perpendicular deviation was a weak predictor of both somatosensory adaptation ( $r = 0.3$ ,  $p = 0.6$ ) and auditory adaptation ( $r = -0.32$ ,  $p < 0.05$ ).

## Discussion

In the experiments reported above, somatosensory feedback and auditory feedback were altered alone or in combination as subjects repeated a simple speech utterance. A negative correlation was observed in the amount of compensation for each perturba-



**Figure 7.** Adaptation to the somatosensory perturbation was not affected by the introduction of the auditory perturbation. **A**, Change in F1 frequency (top) and movement deviation (bottom) for subjects who compensated for the somatosensory perturbation (blue curves,  $N = 7$ ) and subjects who did not (red curves,  $N = 9$ ). The somatosensory perturbation was applied before the simultaneous application of the somatosensory and auditory perturbations (Experiment 5). Shaded regions represent  $\pm 1$  SE. **B**, Subjects who failed to compensate for the somatosensory perturbation (red bars) were failing to compensate for that perturbation before the auditory perturbation was applied. Error bars represent  $\pm 1$  SE.



**Figure 8.** Subjects who compensated more for the somatosensory perturbation compensated less for the auditory perturbation and vice versa. **A**, Twenty-one percent of subjects ( $N = 10$ ) compensated only for the somatosensory perturbation, 26% of subjects ( $N = 12$ ) compensated for both the somatosensory and auditory perturbations, and 53% of subjects ( $N = 25$ ) compensated only for the auditory perturbation. **B**, A negative correlation was observed between the amount of compensation for each perturbation: the more subjects compensated for the somatosensory perturbation, the less they compensated for the auditory perturbation. The red points represent subjects who compensated for the auditory perturbation. The blue points represent subjects who compensated for the somatosensory perturbation. The gray points represent subjects who compensated for both perturbations.

tion. By applying the perturbations alone and then in different combinations, the source of this negative correlation was found to be the result of a preferential reliance that individuals show for either somatosensory or acoustical feedback during speech production.

Over the past 15 years, several studies have altered either auditory feedback or somatosensory feedback to simulate speech motor learning (Houde and Jordan, 1998; Baum and McFarland, 2000; Tremblay et al., 2003; Jones and Munhall, 2005; Purcell and Munhall, 2006; Nasir and Ostry, 2006). In each case, adaptation was rarely observed in all subjects. Depending on the word or words used as test utterances and how the perturbations were applied, anywhere from 50 to 85% of subjects showed some amount of compensation, with higher rates typical of the auditory perturbation. This finding presented a puzzle because studies of motor learning in arm movements consistently find adaptation rates of almost 100% (Shadmehr and Mussa-Ivaldi, 1994; Brashers-Krug et al., 1996; Krakauer et al., 2000; Mattar and Ostry, 2007).

Here, as in previous speech studies, a significant percentage of subjects failed to adapt to each perturbation. The results from three experiments in which we applied the two perturbations at the same time provide an answer as to why. When we delivered the somatosensory and auditory perturbations simultaneously, every subject who failed to adapt to the auditory perturbation adapted to the somatosensory perturbation and vice versa. And those who adapted to the somatosensory perturbation largely ignored the auditory perturbation when it was applied on its own. Some individuals, it seems, show a greater reliance on either somatosensory or auditory feedback during speech motor learning.

We used the term “sensory preference” to describe the idea that subjects who adapted to the somatosensory perturbation adapted less or not at all to the auditory perturbation. Another way to characterize this finding is to say that some individuals are simply more sensitive to a particular type of sensory error signal during speech. Recent experiments have separately perturbed auditory and somatosensory feedback while imaging the brain (Tourville et al., 2008; Golfinoopoulos et al., 2011). Real-time perturbations of somatosensory feedback during speech resulted in an increased blood oxygen level-dependent (BOLD) response in parietal regions while real-time perturbations of auditory feedback saw an increased BOLD response in temporal regions. One prominent neural network model of speech production (Golfinoopoulos et al., 2010) suggests that, during ongoing speech, somatosensory error signals are used in combination with auditory error signals in frontal lobe motor areas. Here, motor commands are updated to compensate for discrepancies between expected sensory feedback of speech production and actual sensory feedback. Individual differences in the strength of somatosensory and auditory error signals that project to these motor regions or the importance placed on different sensory error signals within these motor regions could explain the behavioral phenomena observed here.

The idea that some people might be more sensitive to changes in somatosensory feedback during speech is not new. In experiments studying compensation for delayed auditory feedback, Yates (1965) hypothesized that differences in susceptibility to the perturbation might be “a function of the degree of dependence on auditory feedback for the normal monitoring of speech compared with dependence on kinaesthetic and sensory feedback.” Tests of this hypothesis using delayed auditory feedback have produced mixed results (Burke, 1975; Attanasio, 1987). As far as we know, the studies presented here are the first to alter somatosensory and auditory feedback during speech and find stable individual differences in how subjects respond to the two error signals. This finding contrasts with studies of limb movement in which individuals show a more uniform pattern of sensory integration (van Beers et al., 2002).

Increased sensitivity to a particular type of sensory error signal during speech could be shaped by sensory experience. When Nasir and Ostry (2008) perturbed somatosensory feedback during speech in postlingually deaf adults, every subject showed adaptation to the perturbation. Normal-hearing controls, on the other hand, showed more typical patterns of adaptation, with some compensating for the perturbation and others ignoring it. Hearing loss presumably drives changes in the reliance on somatosensory feedback observed during the speech of postlingually deaf adults. However, it is unknown how a similar reliance on somatosensory feedback might develop in healthy subjects, as observed here. As speech is necessarily tied to language, linguistic experience could play a role in determining whether individuals are more sensitive to auditory or somatosensory feedback during

speech motor learning. All tested subjects were native English speakers, but because our subjects were recruited in a bilingual city many also spoke French. Indeed, we feel that this is an avenue that merits further experimentation.

In the experiments detailed above, we used a somatosensory perturbation that pulled the jaw outward with no measurable affect on F1, and an auditory perturbation that decreased the frequency of F1 without changing the motion path of the jaw (Fig. 3). In other words, the perturbations were independent. We believe this design was crucial because it left no ambiguity with regard to the reason for adaptation to each perturbation. Reductions in load-induced movement deviation could only have been driven by somatosensory feedback. Similarly, increases in produced F1 could only have been driven by changes in auditory feedback (although somatosensory feedback from the articulators would change over the course of learning as subjects adapted). If each of the perturbations had both somatosensory and auditory effects, the source of adaptation would be unclear, making it difficult to group subjects based on whether they responded to somatosensory feedback or auditory feedback or both.

Finally, one might wonder why some individuals would care to compensate at all for a somatosensory perturbation that has no measurable affect on the sound of the voice. Over the last decade, work from our group (Tremblay et al., 2003, 2008; Nasir and Ostry, 2006, 2008, 2009) has shown that individuals compensate for small jaw perturbations applied during speech. Compensation to similar perturbations is also observed during silent speech and during the speech of profoundly deaf individuals. We take this as evidence that the nervous system actively monitors somatosensory feedback during speech, and that speech has both acoustical goals and movement goals that can be experimentally dissociated.

## References

- Attanasio JS (1987) Relationships between oral sensory feedback skills and adaptation to delayed auditory feedback. *J Commun Disord* 20:391–402.
- Baum SR, McFarland DH (2000) Individual differences in speech adaptation to an artificial palate. *J Acoust Soc Am* 107:3572–3575.
- Brainard MS, Doupe AJ (2000) Auditory feedback in learning and maintenance of vocal behaviour. *Nat Rev Neurosci* 1:31–40.
- Brashers-Krug T, Shadmehr R, Bizzi E (1996) Consolidation in human motor memory. *Nature* 382:252–255.
- Burke BD (1975) Susceptibility to delayed auditory feedback and dependence on auditory or oral sensory feedback. *J Commun Disord* 8:75–96.
- Delattre P, Liberman AM, Cooper FS, Gerstman LJ (1952) An experimental study of the acoustic determinants of vowel color; synthesized from spectrographic patterns. *Word* 3:195–210.
- Feng Y, Gracco VL, Max L (2011) Integration of auditory and somatosensory error signals in the neural control of speech movements. *J Neurophysiol* 106:667–679.
- Golfinoopoulos E, Tourville JA, Guenther FH (2010) The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage* 52:862–874.
- Golfinoopoulos E, Tourville JA, Bohland JW, Ghosh SS, Nieto-Castanon A, Guenther FH (2011) fMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage* 55:1324–1338.
- Gracco VL, Löfqvist A (1994) Speech motor coordination and control: evidence from lip, jaw, and laryngeal movements. *J Neurosci* 14:6585–6597.
- Houde JF, Jordan MI (1998) Sensorimotor adaptation in speech production. *Science* 279:1213–1216.
- Ito T, Ostry DJ (2010) Somatosensory contribution to motor learning due to facial skin deformation. *J Neurophysiol* 104:1230–1238.
- Jones JA, Munhall KG (2005) Remapping auditory-motor representations in voice production. *Curr Biol* 15:1768–1772.
- Krakauer JW, Pine ZM, Ghilardi MF, Ghez C (2000) Learning of visuomotor

- tor transformations for vectorial planning of reaching trajectories. *J Neurosci* 20:8916–8924.
- Lane H, Tranel B (1971) The Lombard sign and the role of hearing in speech. *J Speech Hear Res* 14:677–709.
- Lane H, Webster JW (1991) Speech deterioration in postlingually deafened adults. *J Acoust Soc Am* 89:859–866.
- Lombard E (1911) Le signe de l'elevation de la voix. *Ann Mal Oreille Larynx* 37:101–119.
- Mattar AA, Ostry DJ (2007) Neural averaging in motor learning. *J Neurophysiol* 97:220–228.
- Nasir SM, Ostry DJ (2006) Somatosensory precision in speech production. *Curr Biol* 16:1918–1923.
- Nasir SM, Ostry DJ (2008) Speech motor learning in profoundly deaf adults. *Nat Neurosci* 11:1217–1222.
- Nasir SM, Ostry DJ (2009) Auditory plasticity and speech motor learning. *Proc Natl Acad Sci U S A* 106:20470–20475.
- Perkell JS, Guenther FH, Lane H, Perrier P, Vick J, Wilhelms-Tricarico R, Zandipour M (2000) A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *J Phonetics* 28:233–272.
- Purcell DW, Munhall KG (2006a) Compensation following real-time manipulation of formants in isolated vowels. *J Acoust Soc Am* 119:2288–2297.
- Purcell DW, Munhall KG (2006b) Adaptive control of vowel formant frequency: evidence from real-time formant manipulation. *J Acoust Soc Am* 120:966–977.
- Rochet-Capellan A, Ostry DJ (2011) Simultaneous acquisition of multiple auditory-motor transformations in speech. *J Neurosci* 31:2657–2662.
- Shadmehr R, Mussa-Ivaldi FA (1994) Adaptive representation of dynamics during learning of a motor task. *J Neurosci* 14:3208–3224.
- Tourville JA, Reilly KJ, Guenther FH (2008) Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39:1429–1443.
- Tremblay S, Shiller DM, Ostry DJ (2003) Somatosensory basis of speech production. *Nature* 423:866–869.
- Tremblay S, Houle G, Ostry DJ (2008) Specificity of speech motor learning. *J Neurosci* 28:2426–2434.
- van Beers RJ, Wolpert DM, Haggard P (2002) When feeling is more important than seeing in sensorimotor adaptation. *Curr Biol* 12:834–837.
- Villacorta VM, Perkell JS, Guenther FH (2007) Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *J Acoust Soc Am* 122:2306–2319.
- Yates AJ (1965) Delayed auditory feedback and shadowing. *Q J Exp Psych* 17:125–131.