Auditory plasticity and speech motor learning

Sazzad M. Nasir[a] and David J. Ostry[a,b,1]

[a]Department of Psychology, McGill University, Montreal, QC, Canada H3A 1B1; and [b]Haskins Laboratories, New Haven, CT 06511

Is plasticity in sensory and motor systems linked? Here, in the context of speech motor learning and perception, we test the idea sensory function is modified by motor learning and, in particular, that speech motor learning affects a speaker's auditory map. We assessed speech motor learning by using a robotic device that displaced the jaw and selectively altered somatosensory feedback during speech. We found that with practice speakers progressively corrected for the mechanical perturbation and after motor learning they also showed systematic changes in their perceptual classification of speech sounds. The perceptual shift was tied to motor learning. Individuals that displayed greater amounts of learning also showed greater perceptual change. Perceptual change was not observed in control subjects that produced the same movements, but in the absence of a force field, nor in subjects that experienced the force field but failed to adapt to the mechanical load. The perceptual effects observed here indicate the involvement of the somatosensory system in the neural processing of speech sounds and suggest that speech motor learning results in changes to auditory perceptual function.

sensorimotor adaptation | speech perception | speech production

Fig. 1. Experimental set-up, protocol, and auditory test stimuli. (A) A robotic device delivered a velocity-dependent load to the jaw. (B) Experimental subjects completed an auditory identification task before and after force-field training. Control subjects repeated the same set of utterances but were not attached to the robot. (C) Subjects had to indicate whether an auditory test stimulus sounded more like *head* or *had*.

As a child learns to talk, or as an adult learns a new language, a growing mastery of oral fluency is matched by an increase in the ability to distinguish different speech sounds (1–5). Although these abilities may develop in isolation, it is also possible that speech motor learning alters a speaker's auditory map. This study offers a direct test of this hypothesis, that speech motor learning, and, in particular, somatosensory inputs associated with learning, affect the auditory classification of speech sounds (6–8). We assessed speech learning by using a robotic device that displaced the jaw and modified somatosensory input without altering speech acoustics (9–11). We found that even though auditory feedback was unchanged over the course of learning, subjects classify the same speech sounds differently after motor learning than before. Moreover, the perceptual shift was observed only in subjects that displayed motor learning. Subjects that failed to adapt to the mechanical load showed no perceptual shift even though they experienced the same force field as subjects that showed learning. Our findings are consistent with the idea that speech learning affects not only the motor system but also involves changes to sensory areas of the brain.

Results

To explore the idea that speech motor learning affects auditory perception, we trained healthy adults in a force-field learning task (12, 13) in which a robotic device applied a mechanical load to the jaw as subjects repeated aloud test utterances that were chosen randomly from a set of four possibilities (*bad*, *had*, *mad*, *sad*) (Fig. 1). The test utterances were displayed on a computer monitor that was placed in front of the subjects. The mechanical load was velocity-dependent and acted to displace the jaw in a protrusion direction, altering somatosensory but not auditory feedback. Subjects were trained over the course of ≈275 utterances. Subjects also participated in auditory perceptual tests before and after the force-field training. In the perceptual tests, the subject had to identify whether an auditory stimulus chosen at random from a synthesized eight-step spectral continuum sounded more like the word *head* or *had* (Fig. 1). A psychometric

function was fitted to the data and gave the probability of identifying the word as *had*. We focused on whether motor learning led to changes to perceptual performance.

Sensorimotor learning was evaluated by using a composite measure of movement curvature that reflected both the deviation of the jaw path from a straight line and curvature in three dimensions. Kinematic tests of adaptation were conducted quantitatively on a per-subject basis by using repeated measures ANOVA followed by Tukey's HSD posthoc tests. The dependent measure in the test of motor learning was movement curvature. Curvature was assessed on a per-subject basis in null condition trials and at the start and the end of learning. Statistically reliable adaptation was observed in 17 of the 23 subjects ($P < 0.01$; see
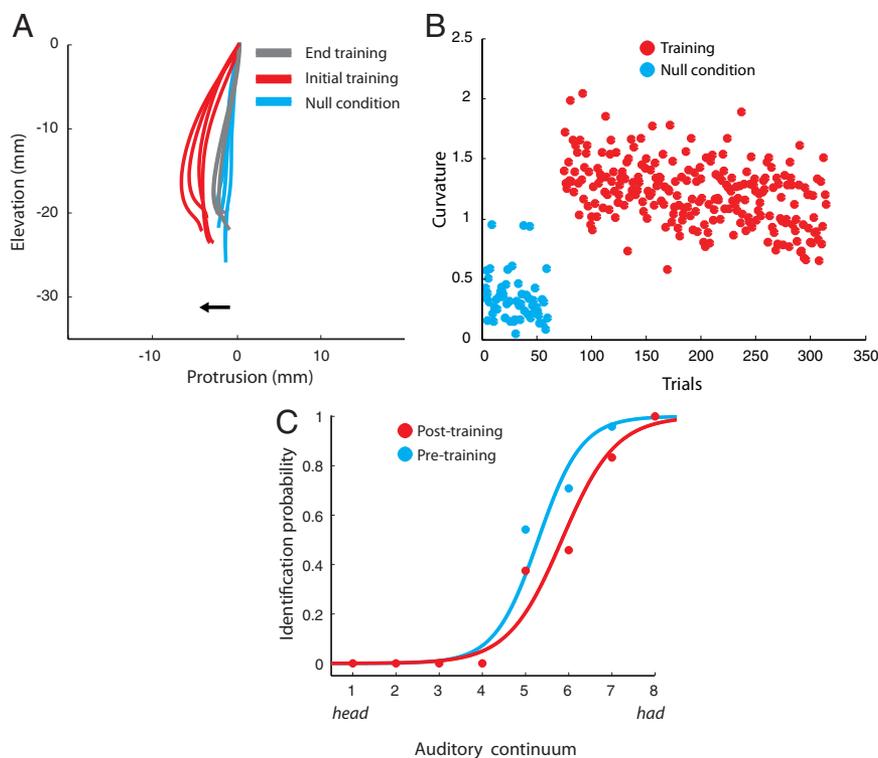
Fig. 2. Auditory remapping after force-field learning. (*A*) Sagittal plane jaw movement paths for a representative subject showing adaptation. Jaw paths were straight in the absence of load (cyan). The jaw was deflected in the protrusion direction when the load was introduced (red). Movement curvature decreased with training (gray). (*B*) Scatter plot showing movement curvature over the course of training for the same subject. The ordinate depicts movement curvature; the abscissa gives trial number. Curvature is low during null condition trials (cyan), increases with the introduction of the load, and then decreases over the course of training (red). (*C*) The psychometric function depicting identification probability for *had* before (cyan) and after (red) force-field learning. A perceptual shift toward *head* was observed after learning.

Materials and Methods for details). This result is typical of studies of speech motor learning in which approximately one-third of all subjects fail to adapt (9–11, 14). Fig. 2*A* shows a representative sagittal plane view of jaw trajectories during speech for a subject that adapted to the load. Movements are straight in the absence of load (null condition: cyan); the jaw is displaced in a protrusion direction when the load is first applied (initial exposure: red); curvature decreases with training (end training: gray). Fig. 2*B* shows movement curvature measures for the same subject, for individual trials, over the entire course of the experiment. As in the movements shown in Fig. 2*A*, values of curvature were low in the null condition, increased with the introduction of load and then progressively decreased with training. The auditory psychometric function for this subject shifted to the right after training (Fig. 2*C*), which indicates that words sounded more like *head* after learning.

Fig. 3*A* shows perceptual psychometric functions for adapted subjects before (cyan) and after (red) force-field training. A rightward shift after training is evident. A measure of probability, which was used to assess perceptual change, was obtained by summing each subjects' response probabilities for individual stimulus items and dividing the total by a baseline measure that was obtained before learning. The change in identification probability from before to after training was used to gauge the perceptual shift. We found that the majority of subjects that adapted to the mechanical load showed a consistent rightward shift in the psychometric function after training (15 of 17). The amount of the rightward shift averaged across adapted subjects was $7.9 \pm 1.4\%$ (mean $\pm$ SE) that was significantly different from zero ($t_{16} = 5.36, P < 0.001$). This rightward perceptual shift means that after force-field learning the auditory stimuli are

more likely to be classified as *head*. In effect, the perceptual space assigned to *head* increased with motor learning. The remaining six subjects, which failed to adapt, did not show any consistent pattern in their perceptual shifts. The mean shift was $0.2 \pm 3.9\%$ (mean $\pm$ SE) that was not reliably different from zero ($t_5 = 0.05, P > 0.95$).

We evaluated the possibility that the perceptual shift might be caused by factors other than motor learning by testing a group of 21 control subjects who repeated the entire experiment without force-field training, including the repetition of the entire sequence of several hundred stimulus words. Under these conditions, the perceptual shift for the control subjects, computed in the same manner as for the experimental subjects, was $1.3 \pm 1.7\%$ (mean $\pm$ SE), which was not different from zero ($t_{20} = 0.77, P > 0.4$) (Fig. 3*C*). Moreover, we found no difference in the shifts obtained for the nonadapted and the control subject groups ($t_{25} = 0.29, P > 0.75$). The perceptual shift for nonadapted and control subjects combined was $1.1 \pm 1.7\%$ (mean $\pm$ SE), which was also not different from zero ($t_{26} = 0.69, P > 0.45$). Fig. 3*B* shows the psychometric functions averaged over the 27 nonadapted and control subjects combined, before (cyan) and after (red) word repetition (or force-field training for the nonadapted subjects). No difference can be seen in the psychometric functions of the subjects that did not experience motor learning.

Repeated-measures ANOVA followed by Bonferroni-corrected comparisons was conducted on the perceptual probability scores before and after training. The analysis compared the scores of adapted subjects with those of control subjects and nonadapted subjects combined. The test thus compares the perceptual performance of subjects that successfully learned the motor task with those that did not. We obtained a reliable
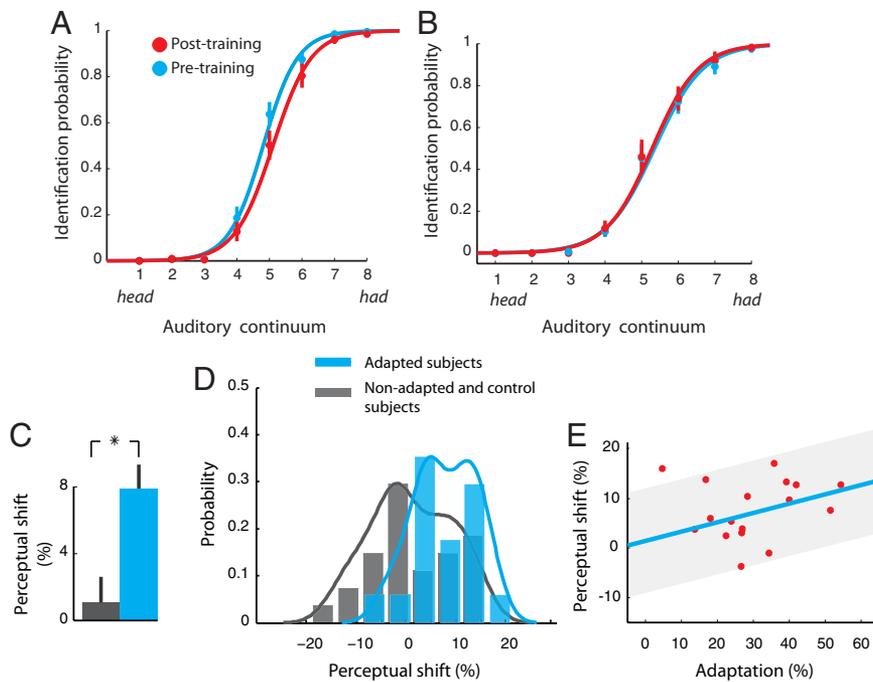
Fig. 3. Auditory perceptual change is observed after speech motor learning. (*A*) Psychometric function averaged over adapted subjects shows an auditory perceptual shift to the right after speech learning (cyan: pretraining, red: posttraining). (*B*) There is no perceptual shift for nonadapted and control subjects. (*C*) The perceptual shift for adapted subjects (cyan) was reliably greater than the shift observed in nonadapted and control subjects (gray); *, $P < 0.01$. In the latter case, the perceptual change was not different from zero. (*D*) Histograms showing the distribution of perceptual change for the adapted (cyan) and the nonadapted/control subjects (gray). (*E*) The perceptual shift was correlated with the amount of adaptation. Subjects that showed greater adaptation also had greater perceptual shifts.

statistical interaction ($F_{(1,42)} = 8.84$, $P < 0.01$), indicating that perceptual performance differed depending on whether or not subjects experienced motor learning. Posthoc comparisons indicated that for adapted subjects, identifications scores were significantly different after training than before ($P < 0.001$). For subjects that did not show motor learning, the difference in the two perceptual tests was nonsignificant ($P > 0.45$). Thus, speech motor learning in a force-field environment modifies auditory performance. Word repetition alone cannot explain the observed perceptual effects.

To further characterize the pattern of perceptual shifts, we obtained histograms giving the distribution of shifts for both the adapted (Fig. 3*D*, cyan) and the combined nonadapted and control (Fig. 3*D*, gray) groups. The histogram for the adapted group is to the right of the histogram for the others in Fig. 3*D*. Overall, 88% of adapted subjects (15 of 17) showed a rightward shift, whereas 48% of the subjects in the combined group shifted to the right. As noted above, the mean shift for the combined group was small and not reliably different from zero. A rank-sum test indicated a significant difference in the perceptual shifts between the adapted and the combined groups (Kruskal-Wallis, $\chi^2_{(1)} = 7.09$, $P < 0.007$). A further comparison between the adapted and combined groups was made by using the bootstrap sampling technique. Using this procedure, the perceptual difference between the adapted and combined groups was 6.8% (see *SI Text*).

To test the idea that speech learning affects auditory perception, we also examined the possibility that subjects that showed greater learning would also show a greater perceptual shift. We calculated an index of learning for each adapted subject by computing the reduction in curvature over the course of training divided by the curvature caused by the introduction of load. A value of 1.0 indicates complete adaptation. Computed in this fashion, adaptation ranged from 0.05 to 0.55 and when averaged

across subjects and test words it was 0.29 ± 0.03 (mean $\pm$ SE). Fig. 3*E* shows the relationship between the amount of adaptation and the associated perceptual shift. We found that adapted subjects showed a small, but significant, correlation of 0.53 ($P < 0.05$) between the extent of adaptation and the measured perceptual shift.

The speech motor learning that is observed in the present study and in others that have used similar techniques primarily depends on somatosensory, rather than auditory, information. This conclusion is supported by the observation that adaptation to mechanical load occurs when subjects perform the speech production task silently, indicating that it does not depend on explicit acoustical feedback (9). It is also supported by the finding that adaptation to mechanical loads during speech is observed in profoundly deaf adults who are tested with their assistive hearing devices turned off (11).

In the present study, we assessed the possibility that there are changes in auditory input over the course of force-field training that might contribute to motor learning to the observed perceptual shift. Acoustical effects related to the application of load and learning were evaluated by computing the first and second formant frequencies of the vowel/æ/ immediately after the initial consonant in each of the test utterances. Fig. 4*A* shows an example of the raw acoustical signal for the test utterance *had* and the associated first and second formants of the speech spectrogram. The acoustical data included for analysis were for only those subjects who adapted to load. Null condition blocks are shown in cyan. The red and black bars give acoustical results during the initial and final training blocks with the force field on. In all cases, the first formant is displayed in solid colors and the second formant is in pale colors (Fig. 4*B*).

Acoustical effects were assessed quantitatively on a between-subjects basis. We focused on potential effects of the load's introduction and possible changes with learning. The dependent

A

B
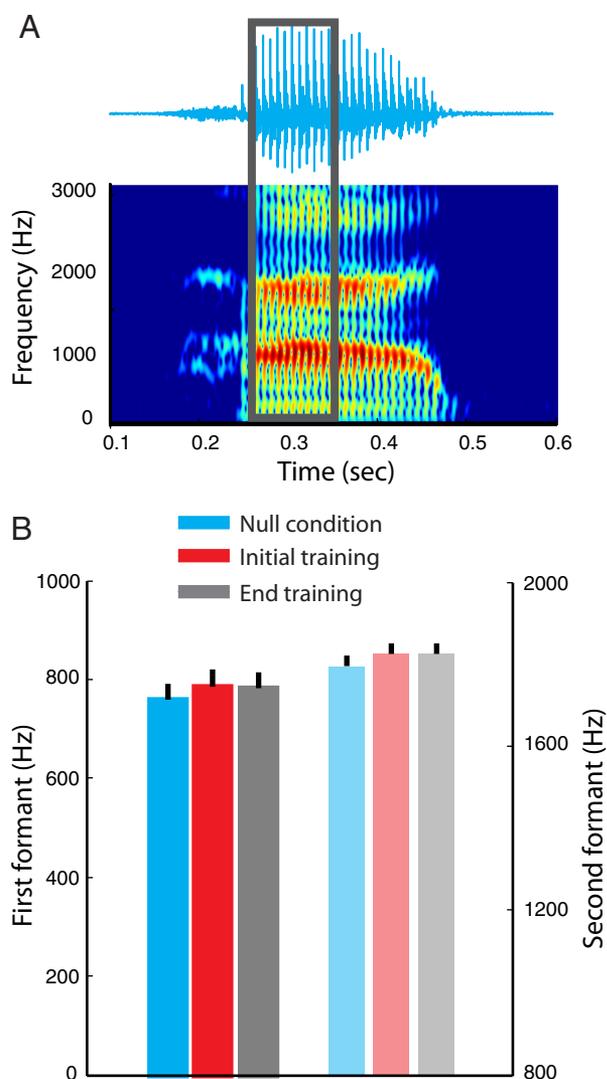
Fig. 4. There were no systematic acoustical effects associated with force-field learning. (*A*) (*Upper*) The raw acoustical waveform for the word *had*. (*Lower*) The formants of the corresponding spectrogram are shown in yellow and red bands. (*B*) The load had little effect on first and second formant frequencies for adapted subjects. Formants values for the steady-state portion of the vowel /æ/ were computed in the absence of load (cyan), at the introduction of the load (red) and at the end of training (black). The first formant values are shown at the left; the second formant values are shown in pale colors at the right. In all cases ± 1 SE is shown.

measures for the acoustical tests were the first and second formant frequencies. We calculated formant frequencies for each subject separately in null condition trials and at the start and end of learning (see *Materials and Methods*). A repeated-measures ANOVA with three levels (null trials, introduction of load, end of training) produced no statistically significant acoustical effects over the course of learning for subjects that displayed reliable adaptation to load. Specifically, we found no reliable differences in either the first ($F_{(2,32)} = 0.2, P > 0.8$) or the second formant ($F_{(2,32)} = 1.0, P > 0.2$) frequencies.

Discussion

In summary, we found that speech motor learning results in changes to the perceptual classification of speech sounds. When subjects were required to indicate whether an auditory stimulus sounded more like *head* or *had*, there was a systematic change after motor learning such that stimuli were more frequently

classified as *head*. Moreover, the perceptual shift varied with learning. Subjects that showed greater adaptation also displayed greater amounts of perceptual change. The perceptual shift was observed only in subjects who adapted. Subjects who did not adapt and subjects in a control group who repeated the same words but did not undergo force-field training showed no evidence of perceptual change, which suggests that perceptual change and motor learning are linked. The findings thus indicate that speech learning not only alters the motor system but also results in changes to auditory function related to speech.

We assessed the sensory basis of the auditory perceptual effect and concluded that it was somatosensory in nature. The loads to the jaw during learning resulted in changes to the motion path of the jaw and hence to somatosensory feedback. However, there were no accompanying changes to the acoustical patterns of speech at any point in the learning process. Hence, there was no change in auditory information that might result in perceptual modification. The observed motor learning and auditory reca-libration presumably both rely on sensory inputs that are somatosensory in origin.

The perceptual shift that we have observed is in the same direction as that reported in studies of perceptual adaptation (6, 15). Cooper and colleagues (6, 15) observed that listening to repetitions of a particular consonant–vowel stimulus reduced the probability that they would report hearing this same stimulus in subsequent perceptual testing. Although the effect reported in the present study is similar to that observed by Cooper and colleagues, there are important differences, which suggest that the effects are different in origin. The absence of any systematic perceptual shift in nonadapted subjects shows that simply re-peating or listening to a given test stimulus will not produce the observed perceptual change. Moreover, our control subjects who also repeated and listened to the same set of utterances did not show a reliable perceptual change. Both of these facts are consistent with the idea that the perceptual change observed in the present study is specifically tied to motor learning and is not caused by the effects of sensory adaptation.

Influences of somatosensory input on auditory perception have been documented previously. The reported instances extend from somatosensory inputs to the cochlear nucleus through to bidirec-tional interactions between auditory and somatosensory cortex (16–22). There is also evidence that somatosensory inputs affect auditory perceptual function including cases involving speech (8, 23, 24). The present example of somatosensory–auditory interaction is intriguing because subjects do not normally receive somatosensory inputs in conjunction with the perception of speech sounds but rather with their production. Indeed, the involvement of somato-sensory input in speech perceptual processing would be consistent with the idea that the perception of speech sounds is mediated by the mechanisms of speech production (25, 26). Other evidence for this view include the observation that evoked electromyographic responses to transcranial magnetic stimulation (TMS) to primary motor cortex in humans are facilitated by watching speech move-ments and listening to speech sounds (27, 28) and that repetitive TMS to premotor cortex affects speech perception (29). However, the perceptual effects that we have observed may well occur differently, as a result of the direct effects of somatosensory input on auditory cortex (30, 31).

Here, we have documented changes to speech perception that occur in conjunction with speech motor learning. It would also be reasonable to expect changes to motor function after per-ceptual learning that complement our findings. Preliminary evidence consistent with this possibility has been reported by Cooper and colleagues (6, 15). A bidirectional influence of movement on perception and vice versa might be contrasted with the linkage proposed in the motor theory of speech perception where a dependence of auditory perception on speech motor function is assumed.

We have observed a modest, but statistically significant, relationship between the amount of motor learning and the associated perceptual shift. This finding may indicate that perception is only weakly coupled to motor function and motor learning and may be yet another reflection of the variability that is found in both speech movements and speech acoustical signals. The weak link could also reflect inherent complexities in speech circuitry, in the organization of projections between motor and sensory areas of the brain. The moderate correlation may also arise because learning is incomplete and accordingly our measures sample only partial levels of adaptation that may be more prone to interference by noise. A further possibility is that the measures we have chosen to quantify adaptation and perceptual function may themselves poorly capture the motor and sensory variables that are actually modified by learning.

The present study was motivated by previous reports of the effects of somatosensory input on auditory function and by the known role of somatosensory input in speech motor learning. The specific nature of the mapping between learning and auditory perceptual change needs to be identified. The pattern presumably depends on the interaction of the speech training set, the force field (or sensorimotor transformation), and the region of the vowel space used for perceptual testing. To the extent that perceptual change depends on motor learning, it should be congruent with both the compensatory motor adjustments required for learning and the associated somatosensory change. The specificity of the relationship between sensory change and motor learning is presently unknown. Given the evidence that motor learning is primarily local or instance based (32–35), we would expect perceptual change after speech learning to be similarly restricted to speech sounds closely related to the utterances in the motor training set.

Materials and Methods

Subjects and Tasks. Forty-four paid subjects gave their informed consent to participate in this study. The subjects were young adult speakers of North American English and were free of any speech or speech-motor disorders. The Institutional Review Board of McGill University approved the experimental protocol.

In the experimental condition, 23 subjects read a set of test utterances, one a time, from computer screen while a robotic device altered the jaw movement path by applying a velocity-dependent load that produced jaw protrusion. These subjects also took part in perceptual tests in which they were required to indicate whether an auditory stimulus drawn from an eight-step computer-generated continuum sounded more like *head* or *had*. This procedure allowed us to determine their perceptual threshold for identifying one word rather than the other. The identification task was carried out twice before the force-field training session and once immediately afterward.

A separate group of 21 subjects took part in a control study. The subjects in the control group did not participate in the force-field training phase; however, they read the same set of test utterances as the subjects in the experimental group, but without the robot connected to their jaw. The control group subjects also completed the three auditory identification tests in exactly the same sequence as the experimental group: two sets of identification tests before the word repetition session followed by the third identification task immediately afterward. Each subject was tested in only one of the two groups.

The force-field training task was carried out in blocks of 12 trials each. On a single trial, the test utterance was chosen randomly from a set of four words, *bad*, *had*, *mad*, and *sad*, and presented to the subject on a computer monitor. Within a block of 12 trials, each of the test utterances in the set was presented three times on average. These specific training utterances were used because they involve large amplitude jaw movements and, in the context of the experimental task, result in high force levels that promote adaptation. Note that the continuum from *head* to *had* was chosen for perceptual testing because it overlaps the vowel space used for speech motor training and thus should be sensitive to changes in perceptual classification caused by motor learning.

The presentation of the test words on the computer monitor was manually controlled by the experimenter. There was a delay of 1–2 s between the utterances. The first six blocks in the force-field training session were recorded under null or no-load conditions. The next 24 blocks, involving ≈275 repeti-

tions of the test utterances, were recorded with the load on and constituted the force-field training phase. The control subjects simply had to repeat the same test utterances for the same set of 30 blocks without the robot attached to the jaw. After-effect trials were not included in the experiment to avoid the possibility that subjects might deadapt before perceptual testing.

A set of eight auditory stimuli spanning the spectral continuum between *head* and *had* was presented one at a time through headphones in a two-alternative forced-choice identification task. Each test stimulus was presented 24 times and, thus in total, subjects had to listen to and identify 192 stimuli. A William's square (a variant of Latin square) was used to order the presentation of stimuli for the auditory identification test. In a William's square, the stimuli are presented equally often, and each stimulus occurs the same number of times before and after every other stimulus, thus balancing first-order carryover effects caused by stimulus order. The two sets of identification tests that preceded the training phase were separated by ≈10 min. Subjects completed the third and final identification test immediately after the training session. Subjects were told not to open their mouths or repeat the stimulus items during the perceptual tests.

Experimental Procedures. A robotic device applied a mechanical load to the jaw that resulted in jaw protrusion (see *SI Text*). The load varied with the absolute vertical velocity of the jaw and was governed by the following equation: $F = k|v|$, where F is the load in Newtons, k is a scaling coefficient, and v is jaw velocity in m/s. The scaling coefficient was chosen to have a value ranging from 60 to 80. A higher coefficient was used for subjects who spoke more slowly and vice versa. The maximum load was capped at 7.0 N.

Jaw movement was recorded in three dimensions at a rate of 1 KHz and the data were digitally low-pass filtered offline at 8 Hz. Jaw velocity estimates for purposes of load application were obtained in real time by numerically differentiating jaw position values obtained from the robot encoders. The computed velocity signal was low-pass filtered by using a first-order Butterworth filter with a cut-off frequency of 2 Hz. The smoothed velocity profile was used online to generate the protrusion load. The subject's voice was recorded with a unidirectional microphone (Sennheiser). The acoustical signal was low-pass analogue filtered at 22 KHz and digitally sampled at 44 KHz.

A graphical interface coded in Matlab was used to control the presentation of the auditory stimuli in the perceptual identification task. Stimulus presentation was self-paced; on each trial, subjects initiated stimulus presentation by clicking on a push button displayed on a computer monitor. The stimuli were presented by using Bose QuietComfort headphones, and subjects had to indicate by pressing a button on the display screen whether the stimulus sounded more like *head* or *had*.

The stimulus continuum for the auditory identification test was generated by using an iterative linear predictive coding (LPC)-based Burg algorithm for estimating spectral parameters (14). The procedure involved shifting the first (F1) and the second formant (F2) frequencies in equal steps from values observed for *head* to those associated with *had*. The stimuli were generated from speech tokens of *head* provided by a male native speaker of North American English. For this individual, the average values across five tokens of *head* were 537 and 1,640 Hz for F1 and F2 respectively. The corresponding values for *had* were 685 and 1,500 Hz, respectively. The synthesized continuum had 10 steps in total; we used the middle eight values for perceptual testing because in pilot testing the two extreme values were always identified with perfect accuracy.

The formant shifting procedure was applied to the voiced portion of the speech utterance. The speech signal was first decomposed into sound source and filter components. The first two poles of the filter provide estimates of the first and second formant frequencies in the original speech sample. These two formant values were manipulated to generate values for the shifted formants. A transfer function was formed consisting of one pair of zeros corresponding to the values of the original formants and one pair of poles at the shifted formant frequencies. This transfer function was used to filter the speech signal so as to eliminate the two poles at the original unshifted frequencies and create new poles at the shifted formant values.

Data Analyses. A composite measure of movement path curvature was computed for each repetition of the test utterance. As in previous studies, we have focused on the jaw opening movement to assess adaptation (9, 11). For each opening segment, we obtained two point estimates of path curvature that, in different ways, capture the effect of the load on jaw movement. We used the average of the two estimates as a composite measure of curvature that was used to assess adaptation. The first estimate of the curvature was the maximum perpendicular deviation of the movement path in 3D space relative to a straight line joining the two movement end points and normalized by movement amplitude. Normalization served to correct for differences in curvature

Nasir and Ostry

that were amplitude related. The second curvature estimate was the maximum 3D Frenet-Serret curvature computed at each point along the path. The Frenet-Serret curvature is an extension of the 2D formula in which curvature, K, is proportional to the product of acceleration and velocity. Specifically,

$$K = \frac{|\vec{a} \times \vec{v}|}{|\vec{v}|^3},$$

where

$$\vec{v} = \frac{d\vec{r}}{dt} \quad \text{and} \quad \vec{a} = \frac{d^2\vec{r}}{dt^2}$$

are, respectively, velocity and acceleration of the position vector for the curve, $\vec{r}$. The curvature measure was also normalized by movement amplitude. The two estimates were made unitless by dividing each measure by the average for that measure over all 30 blocks of trials. That allowed us to combine the two point estimates and take their average. The movement start and end were scored at 10% of peak vertical velocity.

We assessed the extent to which the two kinematic deviation measures, the maximum perpendicular deviation and the maximum Frenet-Serret curvature, used to characterize adaptation are correlated. For each subject we obtained estimates of the correlation between these measures, separately for null condition blocks and training blocks. The mean correlation coefficient across subjects during the null-field movement was 0.04 ± 0.04 (mean $\pm$ SE). During force-field training the correlation was -0.06 ± 0.04. The two measures are thus largely uncorrelated and capture different characteristics of the movement path.

Adaptation to the mechanical load was assessed by computing the mean curvature for the first 10% and the last 10% of the force-field training trials, which gave ≈ 25 movements in each case. A similarly computed measure of null condition curvature was obtained by taking the mean of the last 50% of trials in the null condition block, which also involved ≈ 30 movements. Statistical assessments of adaptation were conducted with ANOVA for each subject separately by using null blocks and initial and final training blocks. We classified subjects as having adapted to the mechanical load if an overall ANOVA indicated reliable differences in curvature over the course of training ($P < 0.01$) and subsequent Tukey HSD comparisons showed a reliable increase in movement curvature with the introduction of load ($P < 0.01$) and a

subsequent reliable decrease in curvature from the start to the end of training, also at $P < 0.01$.

The auditory identification test was used to generate a perceptual psychometric function. This function gave the probability of identifying a test stimulus as *had*. A normalized area of the psychometric function was used to assess perceptual changes after motor learning. First, a raw measure of area was computed by summing the probabilities for the response *had* after training for all eight test stimuli presented during the identification test. This raw area was divided out by a baseline measure of area obtained before training and thus provided an identification measure normalized to the probability of responding *had* before the training procedure. The baseline area was obtained by taking the average of the raw areas of the two pre-training identification tasks.

The psychometric functions were obtained by using the response probabilities for each of the eight auditory stimuli. The binomial distribution fitting method (Matlab glmfit) was used to generate a fit for the psychometric functions. A nonparametric distribution fitting method (Matlab ksdensity) was used to fit the histograms shown in Fig. 3D and Fig. S1.

Acoustical effects were quantified by computing the first and second formant frequencies of the vowels, using the same trials as were used to evaluate force-field learning. An interval of ≈ 100 ms that contained the steady-state portion of the vowel was selected manually on a per-trial basis. The formants within this interval were computed by using a formant-tracking algorithm that was based on standard LPC procedures implemented in Matlab. An analysis window of length 25 ms was used. The median of the formant estimates within the interval was used for subsequent analyses.

Statistical Analysis. The main statistical analyses were conducted by using ANOVA followed by posthoc tests. The relationship between perceptual changes and the amount of adaptation was assessed with Pearson's product-moment correlation coefficient. For purposes of this calculation, we excluded 2 of 17 observations that were >2 SDs from the regression line (the shaded region in Fig. 3E). Note that all 17 data points are shown in Fig. 3E. The two excluded observations are above or at the lower boundary of the shaded region.

1. Fox RA (1982) Individual variation in the perception of vowels: Implications for a perception–production link. *Phonetica* 39:1–22.
2. Rvachew S (1994) Speech perception training can facilitate sound production learning. *J Speech Lang Hear Res* 37:347–357.
3. Bradlow AR, Pisoni DB, Akahane-Yamada R, Tohkura Y (1997) Training Japanese listeners to identify English/r/ and/l/: IV. Some effects of perceptual learning on speech production. *J Acoust Soc Am* 101:2299–2310.
4. Perkell JS, et al. (2004) The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *J Acoust Soc Am* 116:2338–2344.
5. Perkell JS, et al. (2004) The distinctness of speakers'/s/-/S/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *J Speech Lang Hear Res* 47:1259–1269.
6. Cooper WE, Billings D, Cole RA (1976) Articulatory effects on speech perception: A second report. *J Phon* 4:219–232.
7. Shiller DM, Sato M, Gracco VL, Baum SR (2009) Perceptual recalibration of speech sounds following speech motor learning. *J Acoust Soc Am* 125:1103–1113.
8. Ito T, Tiede M, Ostry DJ (2009) Somatosensory function in speech perception. *Proc Natl Acad Sci USA* 106:1245–1248.
9. Tremblay S, Shiller DM, Ostry DJ (2003) Somatosensory basis of speech production. *Nature* 423:866–869.
10. Nasir SM, Ostry DJ (2006) Somatosensory precision in speech production. *Curr Biol* 16:1918–1923.
11. Nasir SM, Ostry DJ (2008) Speech motor learning in profoundly deaf adults. *Nat Neurosci* 11:1217–1222.
12. Lackner JR, Dizio P (1994) Rapid adaptation to coriolis force perturbations of arm trajectory. *J Neurophysiol* 72:299–313.
13. Shadmehr R, Mussa-Ivaldi FA (1994) Adaptive representation of dynamics during learning of a motor task. *J Neurosci* 14:3208–3224.
14. Purcell DW, Munhall KG (2006) Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *J Acoust Soc Am* 119:2288–2297.
15. Cooper WE, Lauritsen MR (1974) Feature processing in the perception and production of speech. *Nature* 252:121–123.
16. Jousmaki V, Hari R (1998) Parchment-skin illusion: Sound-biased touch. *Curr Biol* 8:190.
17. Foxe JJ, et al. (2002) Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study. *J Neurophysiol* 88:540–543.
18. Fu KM, et al. (2003) Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci* 23:7510–7515.
19. Murray MM, et al. (2005) Grabbing your ear: Rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 15:963–974.
20. Kayser C, Petkov CI, Augath M, Logothetis NK (2005) Integration of touch and sound in auditory cortex. *Neuron* 48:373–384.
21. Shore SE, Zhou J (2006) Somatosensory influence on the cochlear nucleus and beyond. *Hear Res* 216-217:90–99.
22. Schürmann M, Caetano G, Hlushchuk Y, Jousmäki V, Hari R (2006) Touch activates human auditory cortex. *NeuroImage* 30:1325–1331.
23. Schürmann M, Caetano G, Jousmaki V, Hari R (2004) Hands help hearing: Facilitatory audiotactile interaction at low sound-intensity levels. *J Acoust Soc Am* 30:830–832.
24. Gillmeister H, Eimer M (2007) Tactile enhancement of auditory detection and perceived loudness. *Brain Res* 1160:58–68.
25. Liebermann AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21:1–36.
26. Hickok G, Poeppel D (2000) Toward functional neuroanatomy of speech perception. *Trends Cognit Sci* 4:131–138.
27. Fadiga L, Craighero L, Buccino G, Rizzolati G (2002) Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *Eur J Neurosci* 15:399–402.
28. Watkins KE, Strafella AP, Paus T (2003) Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989–994.
29. Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M (2007) The essential role of premotor cortex in speech perception. *Curr Biol* 17:1692–1696.
30. Hackett TA (2007) Sources of somatosensory input to the caudal belt areas of auditory cortex. *Perception* 36:1419–1430.
31. Ghazanfar AA (2009) The multisensory roles for auditory cortex in primate vocal communication. *Hear Res*, 10.1016/j.heares.2009.04.003.
32. Gandolfo F, Mussa-Ivaldi FA, Bizzi E (1996) Motor learning by field approximation. *Proc Natl Acad Sci USA* 93:3843–3846.
33. Krakauer JW, Pine ZM, Ghilardi MF, Ghez C (2000) Learning of visuomotor transformations for vectorial planning of reaching movements. *J Neurosci* 20:8916–8924.
34. Thoroughman KA, Shadmehr R (2000) Learning of action through adaptive combination of motor primitives. *Nature* 407:742–747.
35. Mattar AA, Ostry DJ (2007) Modifiability of generalization in dynamics learning. *J Neurophysiol* 98:3321–3329.