Current Biology 16, 1918–1923, October 10, 2006 DOI 10.1016/j.cub.2006.07.069

Report

Somatosensory Precision in Speech Production

Sazzad M. Nasir[1] and David J. Ostry[1,2,*]
[1] McGill University
Montreal, Québec
Canada
[2] Haskins Laboratories
New Haven, Connecticut

Summary

Speech production is dependent on both auditory and somatosensory feedback [1–3]. Although audition may appear to be the dominant sensory modality in speech production, somatosensory information plays a role that extends from brainstem responses to cortical control [4–6]. Accordingly, the motor commands that underlie speech movements may have somatosensory as well as auditory goals [7]. Here we provide evidence that, independent of the acoustics, somatosensory information is central to achieving the precision requirements of speech movements. We were able to dissociate auditory and somatosensory feedback by using a robotic device that altered the jaw's motion path, and hence proprioception, without affecting speech acoustics. The loads were designed to target either the consonant- or vowel-related portion of an utterance because these are the major sound categories in speech. We found that, even in the absence of any effect on the acoustics, with learning subjects corrected to an equal extent for both kinds of loads. This finding suggests that there are comparable somatosensory precision requirements for both kinds of speech sounds. We provide experimental evidence that the neural control of stiffness or impedance—the resistance to displacement—provides for somatosensory precision in speech production [8–10].

Results and Discussion

The subject's task was to repeatedly produce a test word (either *row* or *straw*) while a robotic device applied a lateral load to the jaw. The loads were applied to coincide with vowel or consonant production and to thus alter somatosensory feedback during these phases of movement (Figures 1A and 1B). The loads were designed to have a destabilizing effect on the movement end points and were greatest at the two extremes. In this way, we were able to affect positioning accuracy in speech movement. Sensorimotor learning was evaluated over the course of a training period that involved several hundred utterances. Adaptation was quantified with a measure of movement curvature.

Similar Adaptation for Vowels and Consonants

Figure 1C shows a frontal plane view of jaw movement. Movements are initially straight (null field, blue); the path is deflected laterally at the beginning of training (initial exposure, red); curvature decreases with training (end-training, black); there is no after-effect following unexpected removal of the load (after-effect, green). Subjects differ in their degree of adaptation. Figure 1C shows an example of complete adaptation. Figure 1D is more typical; there is a significant decrease of curvature relative to the beginning of training, but performance never returns to the baseline level.

Adaptation was observed for both vowel- and consonant-related loads (Figure 2). For vowel-related loads, six out of seven subjects showed adaptation with the test word *straw* (Figure 2A), as indicated by a significant decrease in curvature over the course of training ($p < 0.01$). For *row*, all five subjects showed adaptation (Figure 2D). For consonant-related loads, four out of five subjects showed adaptation for *row* (Figure 2E). For *straw*, only two out of six subjects adapted to a 3 N maximum load (Figure 2B), however when the load was increased to 4.5 N maximum, four new subjects all showed significant adaptation (Figure 2C).

We assessed the amount of adaptation on a per-subject basis by computing the reduction in curvature over the course of training as a proportion of the curvature due to the introduction of a load. A value of 1.0 indicates complete adaptation. For vowel-related loads, the amount of adaptation averaged across subjects and test words was 0.46 ± 0.09 (mean $\pm$ 1 SEM). For consonants, the mean adaptation was 0.35 ± 0.05. Thus, there was comparable adaptation when loads coincided with both vowel and consonant production ($p > 0.33$). This suggests that somatosensory precision requirements are similar for both kinds of movements.

Adaptation Is Achieved through Impedance Control

We assessed the neural control strategy employed by subjects in achieving adaptation to these destabilizing force fields that had maximum effect at movement ends. We will provide three lines of evidence to suggest that subjects used impedance control to achieve adaptation.

One signature of impedance control is the absence of after-effects when the load is switched off unexpectedly. Figures 1C and 1D show examples of after-effect trials recorded at the end of training. In neither case is the movement path different from that observed under null-field conditions. A quantitative examination of after-effects shows that movement curvature during after-effect trials does not differ significantly from that observed during null-field trials ($p > 0.05$ for each of the conditions shown in Figure 3A). The curvature of after-effect trials did, however, differ from that obtained for initial-exposure trials ($p < 0.01$ in all cases). Had adaptation involved a precise remapping of neural commands to offset the external load, one would have expected a negative after-effect with a curvature comparable to that of initial-exposure trials, as is typically observed in studies of arm movement [11].

*Correspondence: ostry@motion.psych.mcgill.ca

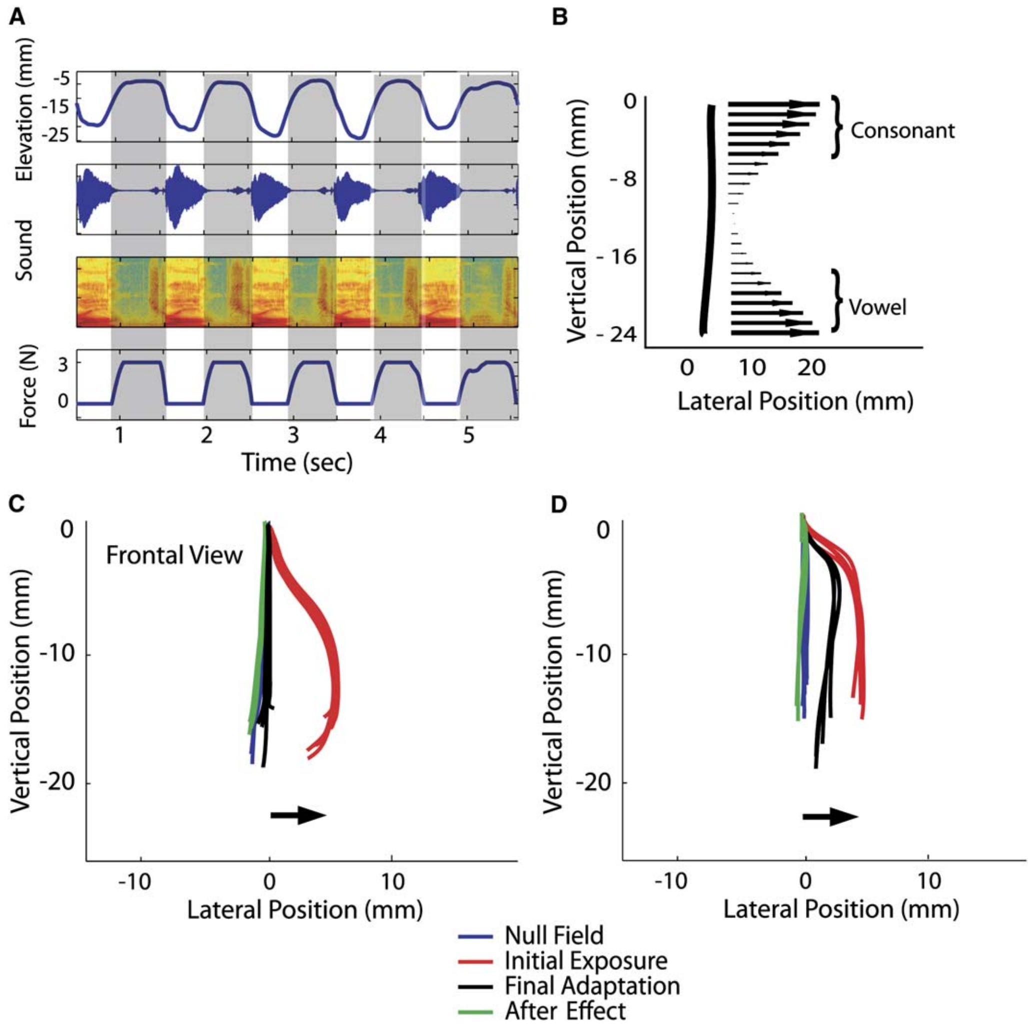

Figure 1. Forces Applied to the Jaw and Typical Patterns of Adaptation

(A) An example of force application during consonant production. The top panel shows the vertical position of the jaw during repetitions of the utterance *straw*. The second and third panels show the raw speech waveform and the corresponding sound spectrogram. The shaded area in the bottom panel shows the commanded force to the jaw. The load scales linearly with vertical jaw position and reaches a maximum when the jaw is fully closed.

(B) Frontal-plane schematic showing position dependence of the load. The load is greatest during either consonant or vowel production.

(C) Frontal view of the movement path of the jaw during the utterance *straw*. The force was applied to the jaw during vowel production. In the no-load condition, movements are straight (blue). When the load is introduced, the jaw path deviates to the right (red). With training, adaptation is achieved (black). When the load is switched off unexpectedly at the end of training, the movement paths do not show an after-effect (green).

(D) An example of imperfect adaptation. Black arrows indicate the direction of the applied load.

We directly tested the idea that subjects use impedance control to achieve adaptation. We tested four new subjects for whom, after adaptation, the direction of the force field was reversed unexpectedly rather than switched off completely. We reasoned that if an impedance based control strategy was being employed to achieve adaptation, then subjects' performance after force-field reversal would not differ from that observed at the end of training. Figure 3B shows a frontal view of performance under these conditions. The test word was *straw*, and the load was applied during the vowel. Null-field conditions are in blue. A large lateral deflection is observed with the introduction of load (red); substantial adaptation occurs after training (black). When the direction of the load is unexpectedly reversed, the movement path is a mirror image of that observed at the end of training (cyan).

Performance in this reversal test was assessed with ANOVA. Figure 3C shows significant adaptation to load by all but one subject ($p < 0.01$). Consistent with the idea that adaptation under these conditions is based on impedance control, movement curvature during the force-field reversal trials did not differ from that observed at the end of training ($p > 0.05$ for all subjects).

We quantified impedance over the course of learning for each of our subjects and for both test words (see Experimental Procedures). Figure 3D shows patterns of impedance and associated movement curvature pooled over subjects, test words, and vowel- versus consonant-related loads. Movement curvature is low under

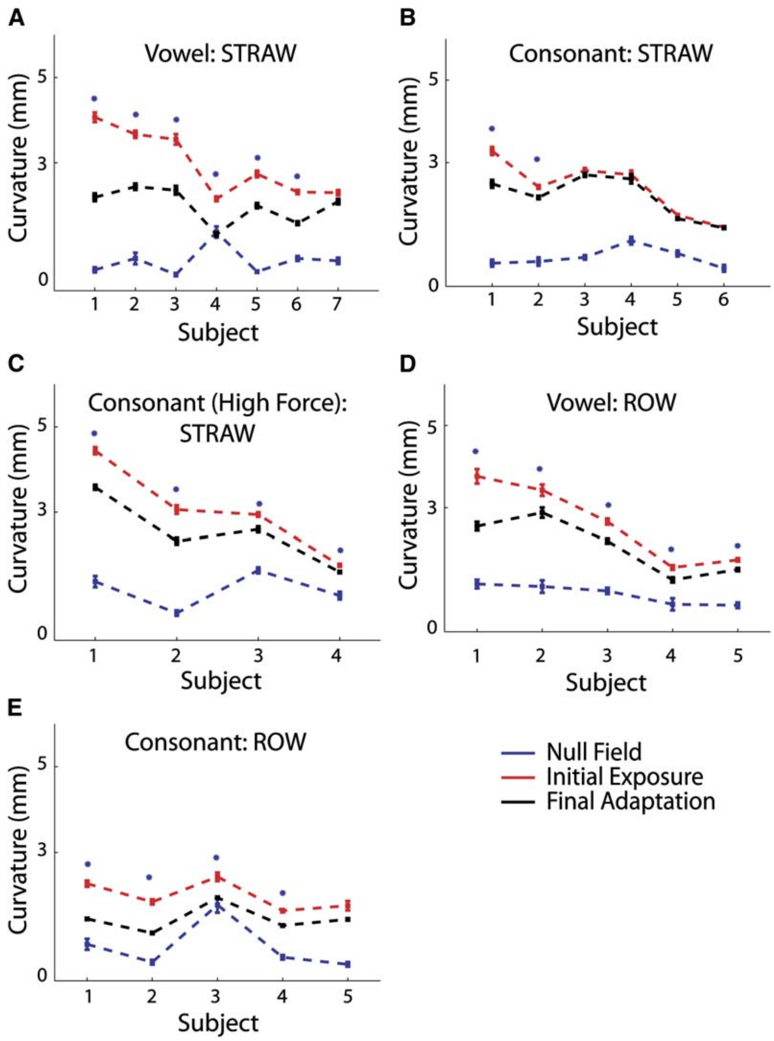

Figure 2. Comparable Adaptation Patterns Are Observed for Consonants and Vowels

Curvature increases with the introduction of a load (red) relative to no-load conditions (blue). Adaptation is observed after training (black). Stars designate significant adaptation ($p < 0.01$). Error bars show ± 1 standard error.

(A) Six out of seven subjects showed adaptation when the test utterance was *straw* and the load was applied during vowel production.

(B) Two out of six subjects showed adaptation for *straw* with a consonant load.

(C) When the load was increased by 50% (load application during consonant in *straw*), four out of four subjects adapted.

(D) All five subjects adapted when the load was applied during the vowel in *row*.

(E) Four out of five subjects adapted when consonant-related loads were used with *row*.

null-field conditions, increases after the introduction of a load, and decreases significantly with adaptation ($p < 0.01$). In contrast, jaw impedance shows a steadily increasing pattern such that impedance is low initially and progressively increases with learning to result in significantly higher impedance ($p < 0.01$). This suggests that subjects achieved adaptation by increasing impedance in order to reduce movement curvature.

The relationship between impedance and movement curvature was assessed quantitatively by computation of impedance change and curvature change on a per-subject basis over the course of learning. Figure 3E shows data for all participants. The abscissa shows the ratio between curvature during initial force-field exposure and curvature at the end of training. Values greater than 1.0 indicate adaptation, whereas values less than 1.0 denote lack of adaptation. Larger values indicate greater curvature reduction. The ordinate of the plot shows the ratio of the impedance coefficient at the end of training to that observed with the initial introduction of a load. As can be seen, the impedance ratio correlates well with the amount of adaptation ($r = 0.8$). The 99% confidence interval for the slope of the linear-regression line is 0.63–1.25, with a mean value of 0.94. Thus, subjects that had greater impedance at the end of training showed greater adaptation.

Loads Have Limited Effects on the Acoustics

The observed adaptation patterns could have been mediated by somatosensory or auditory feedback or the two in combination. The load affects the movement path of the jaw and thus directly alters somatosensory input. The load may also affect the acoustics by changing the shape of the vocal tract. Any acoustical changes due to the load would suggest a role for auditory input in mediating the observed adaptation.

The effect of the load on the acoustics was assessed by computation of the first and second formant frequencies for vowels, the centroid frequency (first spectral moment) for the consonant *s* in *straw*, and the third formant frequency for the consonant *r* in the utterance *row* (Figure 4). Acoustical measures showed no differences due to the introduction of a load ($p > 0.05$). Moreover, there were no differences in the acoustics from the start to the end of training ($p > 0.05$). The absence of any measurable acoustical effect is suggestive of the primary role of somatosensory input in mediating the adaptation observed in these experiments. (For additional

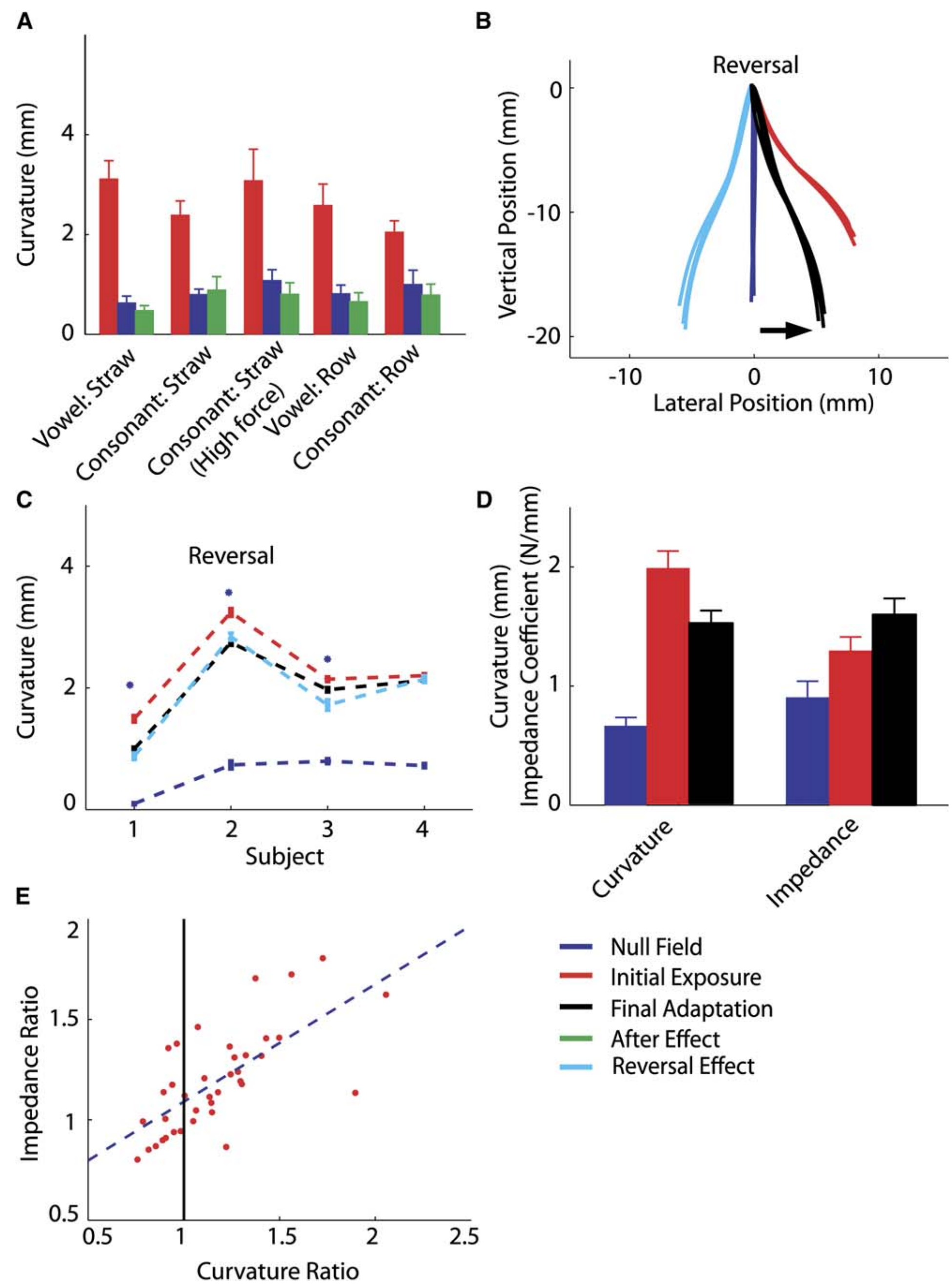

Figure 3. Adaptation Is Based upon the Neural Control of Jaw Impedance

(A) After-effects across utterances and load conditions. The curvature of after-effect trials (green) is comparable to that of null-field trials (blue). Error bars show ± 1 standard error.

(B) Frontal view of the movement path under conditions of force-field reversal. The movement is straight in the no-load condition (blue). There is a prominent deflection with the introduction of the load (red); after training, curvature is reduced (black). When the direction of the load is reversed unexpectedly after training, the movement path (cyan) is the mirror image of the path at the end of training. The black arrow indicates the direction of the training load.

(C) Analysis of curvature with force-field reversal. Three out of four subjects show adaptation. Red denotes curvature at the introduction of the load, and black denotes curvature after training. When the load is reversed, curvature (cyan) does not differ from that observed at the end of training. The baseline curvature is in blue.

(D) Mean impedance for null-field trials (blue), at the introduction of the load (red) and at the end of training (black). Note that impedance progressively increases over the course of training, whereas curvature decreases with adaptation.

(E) A linear relation is observed between impedance and curvature. The ordinate gives the ratio of impedance at the end of training to impedance at the beginning. Values are shown for all subjects, whether there was adaptation or not. The abscissa gives curvature ratios between the beginning and end of training. A curvature ratio greater than 1 indicates an adaptive trend, and a value less than 1 denotes lack of adaptation. The vertical line separates subjects that showed any adaptation from those that did not. Subjects with greater amounts of adaptation show greater impedance. The dotted blue line is the regression line.

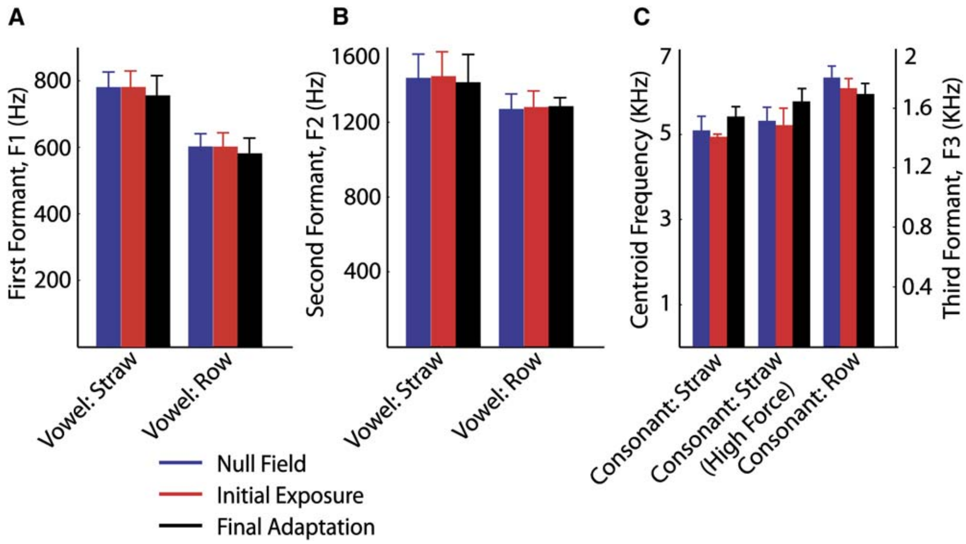

Figure 4. Application of the Load Has No Effect on the Acoustics

The first and second formants of vowels, the centroid frequency for *s*, and the third formant frequency for *r* were computed under no-load conditions (blue), at the introduction of the load (red), and at the end of training (black).

(A) First-formant frequencies for vowel production in *straw* and *row*.

(B) Second-formant frequencies.

(C) Consonant production. Centroid frequencies are shown for *straw*, and third-formant frequencies are shown for *row*. Error bars show ± 1 standard error.

acoustical analyses, see Supplemental Data, including Figure S2, available online.)

Comparable Kinematic Variability for Vowels and Consonants

A systematic examination of differences in variability in jaw position during the production of vowels and consonants is presented in the Supplemental Data. We assessed whether differences in variability were any greater than would be expected on the basis of differences in movement amplitude alone. Using the coefficient of variation, we found no reliable differences in variability between vowels and consonants for either of the test words, consistent with the idea that they have similar kinematic precision (Figure S1 and supplemental text).

In summary, we studied sensorimotor adaptation in speech production in response to mechanical loads that were applied during consonant or vowel production. The loads initially produced lateral deviation of the jaw, but with training the deviation was reduced. The extent of adaptation was comparable for vowel- and consonant-related loads, as was the coefficient of variation. This suggests that precision in jaw positioning is similar for the two kinds of speech sounds. The loads were small enough not to have measurable acoustical effects. Thus, the adaptation observed here seems to be based upon somatosensory change alone.

Subjects in the present study compensated for destabilizing loads at the extremes of speech movements by increasing impedance to reduce displacement of the jaw. These results are consistent with a growing body of evidence that impedance control plays a significant role in speech production [9, 10]. Moreover, they are consistent with observed changes in impedance in comparable situations in human arm movement [12–14]. The adaptations observed here should be distinguished from those that arise, in both speech and limb movements, when subjects learn dynamics that are stable and predictable [2, 11]. In this latter situation, subjects develop precise sensorimotor mappings, which are reflected in mirror-image after-effects when the load is removed. Movements under normal circumstances presumably involve a combination of sensorimotor adjustments that compensate for the dynamics of the limb or the vocal-tract apparatus and separately result in systematic impedance changes that provide stability and aid in achieving the precision requirements of movements.

It is unknown whether the phenomena reported here are restricted to speech movements or more generally characterize orofacial motor learning. The presence of compensation for loads in a lateral direction, which is thought to play a limited role in speech production, may indicate that impedance control applies broadly to orofacial movement. Nevertheless, these compensations are observed here in the presence of speech movements and suggest that the nervous system actively regulates the lateral position of the jaw in speech production. Note that for both vowels and consonants, the coefficient of variation is no different in the lateral direction than in the sagittal plane. This implies that even though movement amplitudes are small, they are no less tightly regulated along a lateral axis.

Approximately 30% of subjects fail to show any sign of adaptation, and indeed some get progressively worse over the course of training. Subjects tested with altered auditory feedback show a comparable trend where approximately 15% of subjects show no sign of adaptation [15]. It would be worthwhile to determine whether these failures are related. Indeed, individuals who fail to adapt to altered mechanical environments may be more sensitive to auditory than to somatosensory feedback, and those who fail with auditory perturbations may rely more heavily on proprioception.

Experimental Procedures

Subjects and Experiments

We tested a total of forty-seven subjects. The task was to repeat a test utterance while a robotic device delivered a lateral load to the jaw [2]. The experiment was carried out in blocks of 15 utterances each. The first three blocks were recorded under null or no-load conditions. In the next 20–25 blocks, equivalent to approximately 300 repetitions of the test word, the subject was trained with the load on. After training, the load was unexpectedly turned off, and one block of "catch" trials was recorded in the absence of a load.

Subjects repeated one of two test words: *straw* or *row*. The test words began with a consonant or consonant cluster and ended in a vowel. The consonant cluster in *straw* was chosen because its pronunciation requires a level of movement complexity sufficient to prompt adaptation. The consonant in *row* was used to restrict palatal contact to limit the possibility that subjects would respond to the perturbation by bracing the tongue against the palate. The vowel sounds *aw* and *ow* were used to produce large-amplitude movement of the jaw. Note that the consonant part of the test words corresponds to a jaw position near to closure. The vowel part of the words corresponds to maximum jaw opening.

Lateral loads were applied to the jaw during either the closing or the opening phase. For a jaw-closing-related load, the load came on at the mid-point of jaw raising and stayed on until the mid-point of jaw lowering. For jaw-opening-related loads, the load came on mid-way through jaw opening and stayed on until mid-way through jaw closing. The midpoints were determined on a trial-by-trial basis by use of the amplitude of movement from the preceding cycle. The load pushed the jaw laterally in proportion to jaw elevation such that the load was at its peak when the jaw was either fully closed or fully open (Figures 1A and 1B). When delivered in this fashion, a load was applied during consonant production 88% of the time and during vowel production 95% of the time.

Subjects participated in one of six experiments involving either vowel- or consonant-related loads. The vowel group experienced the load during jaw opening; the consonant group experienced the load during closing. The maximum lateral load was 3 N except in one consonant group where the maximum load was 4.5 N. In a control group the direction of load was unexpectedly altered by 180° after training.

Kinematic and Acoustical Analyses

A measure of path curvature—maximum perpendicular distance from the path to a straight line from movement start to end—was computed for each repetition of the test utterance. The raising segment, which began with the jaw fully open and ended with the jaw fully closed, was used for the analyses. The start and end of the movement were scored at 20% of peak vertical velocity (empirical patterns were similar when a 10% cut-off was used). The results were qualitatively similar for the jaw-lowering movement.

Subjects showed two trends in response to the application of a load. About 70% showed improvement (or at least no deterioration) in performance with training. The remainder got worse with training, indicating that there was no attempt to compensate for the load. These latter subjects were excluded from statistical analysis of adaptation.

We assessed adaptation by computing the mean curvature for the first 35% and the last 35% of the force-field training trials. This gave approximately 100 movements in each case. Null-field performance was based on curvature measured during the three familiarization blocks. A measure of after-effect was based on the first five trials after an unexpected removal of the load.

We computed a coefficient of impedance by subdividing each movement into two parts, one in which commanded force was zero and the other in which force arose as a result of displacement of the jaw. We recorded a 3D sensed-force vector for each of the two segments and took the magnitude of the vector difference as the measure of force change. The curvature of the movement path was taken as a measure of position change. Note that the impedance coefficient as defined above is a coarse measure of impedance. However, the measure provides ordinal impedance information that is suitable for the assessment of impedance change over learning.

Acoustical effects were assessed by computing spectral measures related to the application of the load. For the consonant-related loads, we manually selected a window that contained the *s* in *straw* or *r* in *row*. For the vowel calculation, the window contained the *aw* in *straw* or *ow* in *row.* The spectral measures were the first and second formant frequencies for vowels, the centroid frequency for *s*, and the third formant frequency for *r*.

Supplemental Data

Supplemental Data include additional Results and Experimental Procedures as well as two figures and are available online at http://www.current-biology.com/cgi/content/full/16/19/1918/DC1/.

Acknowledgments

We thank G. Houle for advice and assistance. This research was supported by National Institute of Deafness and Communicative Disorders grant DC-04669, the Natural Sciences and Engineering Research Council of Canada, and Fonds québécois de la recherche sur la nature et les technologies.

Received: May 26, 2006
Revised: July 14, 2006
Accepted: July 27, 2006
Published: October 9, 2006

References

1. Houde, J.F., and Jordan, M.I. (1998). Sensorimotor adaptation in speech production. Science *279*, 1213–1216.
2. Tremblay, S., Shiller, D.M., and Ostry, D.J. (2003). Somatosensory basis of speech production. Nature *423*, 866–869.
3. Jones, J.A., and Munhall, K.G. (2005). Remapping auditory-motor representations in voice production. Curr. Biol. *15*, 1768–1772.
4. Abbs, J.H., and Gracco, V.L. (1983). Sensorimotor actions in the control of multi-movement speech gestures. Trends Neurosci. *6*, 391–395.
5. Weber, C.M., and Smith, A. (1987). Reflex responses in human jaw, lip, and tongue muscles elicited by mechanical stimulation. J. Speech Hear. Res. *30*, 70–89.
6. Ito, T., Kimura, T., and Gomi, H. (2005). The motor cortex is involved in reflexive compensatory adjustment of speech articulation. Neuroreport *17*, 1791–1794.
7. Guenther, F.H., Ghosh, S.S., and Tourville, J.A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang. *96*, 280–301.
8. Hogan, N. (1985). The mechanics of multi-joint posture and movement control. Biol. Cybern. *52*, 315–351.
9. Shiller, D.M., Ostry, D.J., and Laboissiere, R. (2002). The relationship between jaw stiffness and kinematic variability in speech. J. Neurophysiol. *88*, 2329–2340.
10. Shiller, D.M., Houle, G., and Ostry, D.J. (2005). Voluntary control of human jaw stiffness. J. Neurophysiol. *94*, 2207–2217.
11. Shadmehr, R., and Mussa-Ivaldi, F.A. (1994). Adaptive representation of dynamics during learning of a motor task. J. Neurosci. *14*, 3208–3224.
12. Burdet, E., Osu, R., Franklin, D.W., Milner, T.E., and Kawato, M. (2001). The central nervous system stabilizes unstable dynamics by learning optimal impedance. Nature *414*, 446–449.
13. Franklin, D.W., Osu, R., Burdet, E., Kawato, M., and Milner, T.E. (2003). Adaptation to stable and unstable dynamics achieved by combined impedance control and inverse dynamics model. J. Neurophysiol. *90*, 3270–3282.
14. Osu, R., Burdet, E., Franklin, D.W., Milner, T.E., and Kawato, M. (2003). Different mechanisms involved in adaptation to stable and unstable dynamics. J. Neurophysiol. *90*, 3255–3269.
15. Purcell, D.W., and Munhall, K.G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. J. Acoust. Soc. Am. *120*, 966–977.