# Maintaining Somatosensory Precision in Speech

**Sazzad M. Nasir**[1] **and David J. Ostry**[1,2]

[1]McGill University, Montreal, QC, Canada

[2]Haskins Laboratories, New Haven, CT, USA

nasir@motion.psych.mcgill.ca, ostry@motion.psych.mcgill.ca

***Abstract.*** *As we talk the nervous system receives both auditory and somatosensory feedback. Accordingly, the motor commands that underlie speech movements may have somatosensory as well as auditory goals. Although audition may appear to be the central player in speech production, somatosensory information plays a role that extends from brainstem responses to cortical control. Here we provide evidence that, independent of the acoustics, somatosensory information is critical 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 since these are the major sound categories in speech. We found that, even in the absence of any effect on the acoustics, with learning subjects adapted to an equal extent for both kinds of loads suggesting that somatosensory precision requirements are comparable for both kinds of speech sounds. We provide experimental evidence that the neural control of stiffness or impedance provides for somatosensory precision in speech production.*

## 1. Introduction

Speech production is dependent on both auditory and somatosensory feedback (Houde and Jordan, 1998; Tremblay et al., 2003; Jones and Munhall, 2005). Mechanical perturbations during speech production result in bulbar (Abbs and Gracco, 1983; Weber and Smith, 1987) and transcortical reflexes (Ito et al., 2005). Speech motor learning is also dependent on somatosensory information (Tremblay et al., 2003). Thus the underlying motor commands of speech movements may have somatosensory in addition to auditory goals (Guenther et al., 2006). In the present paper, we have been able to assess the role of somatosensory information in achieving the precision requirements of different speech sounds. We show that somatosensory precision in its own right is important in speech production. We have developed a technique in which mechanical loads to the jaw selectively alter somatosensory input without measurable acoustical effects. By applying mechanical loads to the jaw that alter the jaw's motion path in either vowel or consonant related phases of a speech utterance, we show that there are comparable somatosensory precision requirements for both kinds of speech sounds.

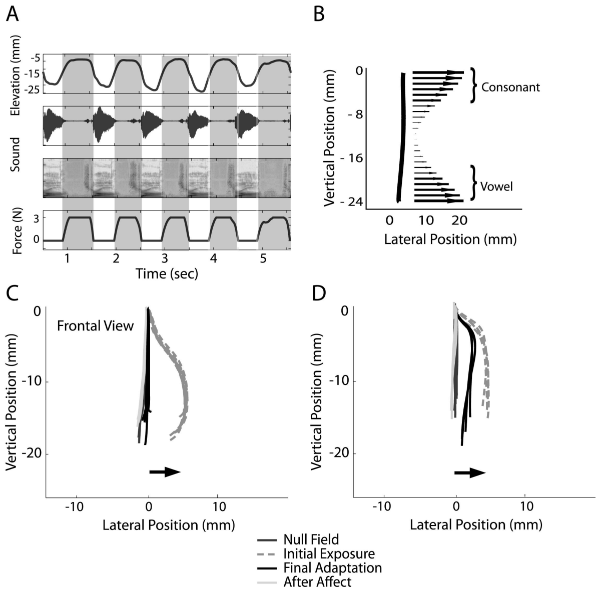

**Figure 1.** Forces applied to the jaw and typical patterns of adaptation. A. An example of force application during consonant production. The top panel shows the vertical position of the jaw during repetitions of the utterance *straw*. The second and third panels show the raw speech waveform and the 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 maximum 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 (solid gray). When the load is introduced, the jaw path deviates to the right (dashed light-gray). With training, adaptation is achieved (solid black). When the load is switched off unexpectedly at the end of training, the movement paths do not show an after-effect (solid light-gray). D. An example of imperfect adaptation. Black arrows indicate the direction of the applied load.

The present study also contributes to a growing body of evidence that points to impedance or stiffness control, the neural regulation of resistance to displacement, as contributing to movement production in speech motor control (Hogan, 1985). Previous studies have indicated a relationship between stiffness and variability in speech (Shiller et al, 2002). Variability is low in directions where stiffness is high and vice versa. It has also been shown that humans can voluntarily modify the directional specificity of jaw stiffness (Shiller et al, 2005). Both of these findings raise the possibility that speakers can purposefully exploit impedance in the context of speech tasks.

The subject's task in these experiments was to repeatedly produce a test word (either *row* or *straw*) while a robotic device applied a lateral load to the jaw. Lateral loads were applied to the jaw either during the closing or 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 and vice-versa for opening related loads. 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). Sensorimotor learning was evaluated over the course of a training period that involved several hundred utterances. Adaptation was quantified for both vowel and consonant related perturbations taking the lowering phase of the movement and using a measure of movement curvature - maximum perpendicular distance from the path to a straight line from movement start to end - as an index of learning. Adaptation was assessed by computing the mean curvature for the first 35% and the last 35% of the force-field training trials. A measure of after-effect was based on the first five trials following unexpected removal of load.

## 2. Adaptation patterns and load's effects on acoustics

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

Adaptation was observed for both vowel- and consonant-related loads (Figure 2). For vowel-related loads, 7 of 8 subjects showed adaptation with the test word *straw* (Figure 2A), as indicated by a significant decrease in curvature over the course of training ($p < 0.01$). All 4 subjects adapted for consonant related loads (Figure 2B). For *row*, all 5 subjects adapted for vowel relateds load and 4 of 5 subjects for consonant related loads. The amount of adaptation was assessed 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 load. A value of 1.0 indicates complete adaptation. For vowel-related loads, the amount of adaptation averaged across subjects and test words was $0.46 \pm 0.09$ (mean ± 1 sem). For consonants, the mean adaptation was $0.35 \pm 0.05$. 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.

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 changes in the acoustics 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 computing 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* (Figures 2C and 2D). There were no differences in acoustical measures due to the introduction of load (p >0.05). Nor were there any differences in the acoustics from the start to 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.

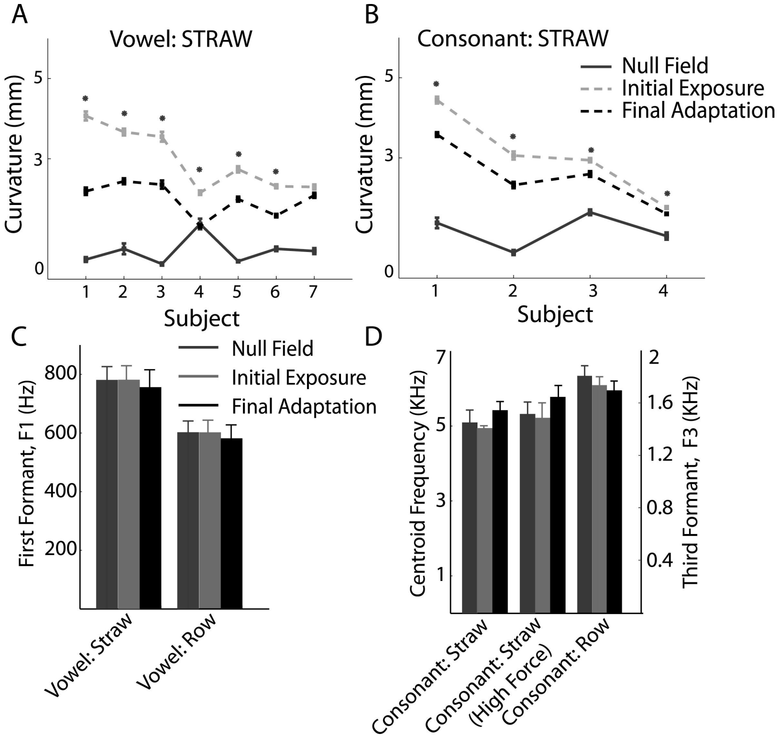

**Figure 2.** Commparable adaptation patterns for consonants and vowels and absence of load's effect on acoustics. Curvature increases with the introduction of load (dashed light-gray) relative to no load conditions (solid gray). Adaptation is observed following training (dashed black). Stars designate significant adaptation (p < 0.01). A. 6 out of 7 subjects showed adaptation when the test utterance was *straw* and the load was applied during vowel production. B. All subjects adapted when load was applied during the consonant in *straw*. C. First formant frequencies for vowel production in *straw* and *row*. D. Consonant production: Centroid frequencies are shown for *straw* and third formant frequencies for *row*.

## 3. Role of impedance control in adaptation

We will provide several lines of evidence to suggest that impedance control was used to achieve adaptation. One signature of impedance control is the absence of after-effects when the load is switched off unexpectedly at the end of training (Figures 1C and 1D). 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 load conditions and test words), but they do differ from that obtained for initial-exposure trials ($p < 0.01$ in all cases). Had adaptation involved a precise re-mapping 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 (Shadmehr and Mussa-Ivaldi, 1994).

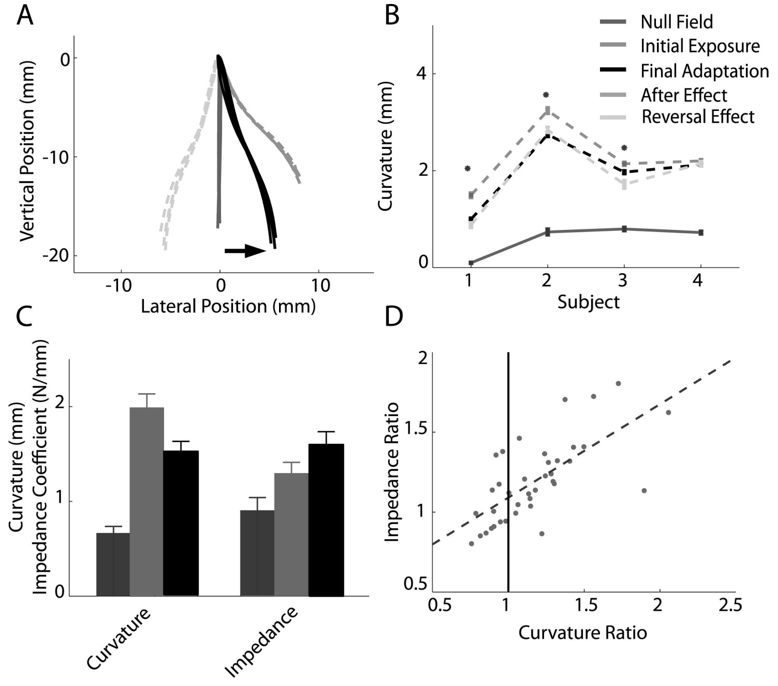

**Figure 3.** Adaptation is achieved by the neural control of jaw impedance. A. Frontal view of the movement path under conditions of force-field reversal. The movement is straight in the no load condition (solid gray). There is a prominent deflection with the introduction of the load (dashed gray); after training curvature is reduced (solid black). When the direction of the load was reversed unexpectedly after training, the movement path (dashed light-gray) was the mirror image of the path at the end of training. The black arrow indicates the direction of the training load. B. Analysis of curvature with force field reversal. 3 out of 4 subjects show adaptation. Solid light-gray denotes curvature at the introduction of load and dashed

black denotes curvature after training. When the load is reversed, curvature dashed light-gray does not differ from that observed at the end of training. The baseline curvature is in blue. C. Mean impedance for null field trials gray, at the introduction of load light-gray and at the end of training black. Note that impedance progressively increases over the course of training while curvature decreases with adaptation. D. 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.

We directly tested the idea that subjects use impedance control to achieve adaptation. We ran four new subjects for whom, following 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 following force-field reversal would not differ from that observed at the end of training. Figure 3A 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 solid gray. A large lateral deflection is observed with the introduction of load (dashed gray); substantial adaptation occurs following training (solid 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 (dashed light-gray). Figure 3B 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 field reversal trials did not differ from that observed at the end of training ($p > 0.05$ for all subjects).

Lastly, a coefficient of impedance was computed by subdividing each movement into two parts, one in which commanded force is zero and the other in which force arises 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. The ratio over force change to position change was taken to be impedance coefficient. Figure 3C shows patterns of impedance and associated movement curvature pooled over subjects, test words, and vowel-versus consonant-related loads. Movement curvature is low under null field conditions, increases following the introduction of 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 a 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 further assessed quantitatively by computing impedance change and curvature change on a per subject basis over the course of learning. Figure 3D 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 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 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.

## 4. Kinematic variability for vowels and consonant

We conducted tests of variability to determine the extent to which variation in jaw position during vowel and consonant production was related to the observed adaptation patterns (Perkell and Nelson, 1985). The extrema of jaw position during opening or closing were used to compute the variability of the vowel or consonant related utterance. Jaw position differs during vowel and consonant production. Figure 4A shows a representative sample of jaw positions in the frontal plane during repetitions of the word *straw*. As can be seen, jaw positions during the consonant phase are more tightly clustered than during vowel production. Figure 4B shows box-plots of variability for the same utterance. We also computed the coefficient of variation (CV), which is the standard deviation of jaw position divided by its mean (Figure 4C). The rationale for using this measure is that in many biological signals variability is proportional to amplitude (Fitts, 1954; Harris and Wolpert, 1998). That is, larger measures are naturally more variable. The coefficient of variation normalizes variability with respect to amplitude and hence enables one to test whether differences in variability in vowels and consonants are any greater than would be expected on the basis of differences in movement amplitude alone. Tukey tests in conjunction with

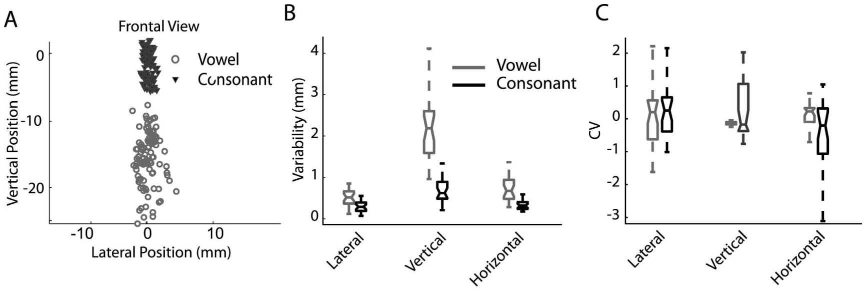

**Figure 4.** Variability of jaw position during vowel and consonant production. A. Frontal plane view of jaw position during consonants (filled triangles) and vowels (open circles) for the utterance *straw*. B. Variability in jaw position across subjects. C. Coefficient of variation (CV) during consonants and vowels. Once differences in variability due to differences in movement amplitude are accounted for, differences in kinematic variability are eliminated.

ANOVA, showed reliable differences in variability of vertical jaw position between vowels and consonants for both *straw* and *row* ($p < 0.01$). No differences were observed for lateral or horizontal jaw position. Using the coefficient of variation, we found no reliable differences in variability in any of the dimensions between vowels and consonant for either of the test words consistent with the idea that they have similar kinematic precision.

## References

Abbs, J.H., and Gracco, V.L. Sensorimotor actions in the control of multi-movement speech gestures. *Trends in Neuroscience,* 6: 391-395, 1983.

Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47: 381-391, 1954.

Guenther, F.H., Ghosh, S.S., and Tourville, J.A. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96: 280-301, 2006.

Harris, C.M., and Wolpert, D.M. Signal-dependent noise determines motor planning. *Nature,* 394: 780-784, 1998.

Hogan, N. The mechanics of multi-joint posture and movement control. *Biological Cybernetics*, 52: 315-351, 1985.

Houde, J.F., and Jordan, M.I. Sensorimotor adaptation in speech production. *Science*, 279: 1213-1216, 1998.

Ito, T., Kimura, T., and Gomi, H. The motor cortex is involved in reflexive compensatory adjustment of speech articulation. *Neuroreport*, 17: 1791-94, 2005.

Jones, J.A., and Munhall, K.G. Remapping auditory-motor representations in voice production. *Current Biology*, 15: 1768-72, 2005.

Perkell, J.S., and Nelson,W.L. Varibility in production of the vowels /i/ and /a/. *Journal of Acoustical Society of America,* 77: 1889-1895, 1985.

Shadmehr, R., and Mussa-Ivaldi, F.A. Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscence*, 14: 3208-3224, 1994.

Shiller, D.M., Ostry, D.J., and Laboissiere, R. The relationship between jaw stiffness and kinematic variability in speech. *Journal of Neurophysiology*, 88: 2329-2340, 2002.

Shiller, D.M., Houle, G., and Ostry, D.J. Voluntary control of human jaw stiffness. *Journal of Neurophysiology*, 94: 2207-17, 2005.

Tremblay, S., Shiller, D.M., and Ostry, D.J. Somatosensory basis of speech production. *Nature*, 423: 866-9, 2003.

Weber, C.M., and Smith, A. Reflex responses in human jaw, lip, and tongue muscles elicited by mechanical stimulation. *Journal of Speech and Hearing Research*, 30: 70–89, 1987.