Speech Motor Adaptation Using Facial Skin Stretch

Takayuki Ito[1], David J. Ostry[2,1]
[1] Haskins Laboratories, [2] McGill University
E-mail: taka@haskins.yale.edu

Abstract

Our goal was to document somatosensory effects in speech motor adaptation that are related to facial skin deformation. The study assessed the role of somatosensory information in speech learning by focusing on the deformation of the facial skin and the motion of the lip. We found that facial skin deformation applied over the course of a series of training trials affected lip movements in a progressive and adaptive manner. The results suggest that the motor plan for the target task was modified in an adaptive manner on the basis of the skin stretch perturbation. This is consistent with the idea that somatosensory afferent input associated with skin deformation may modify the plan for articulatory motion in speech motor learning.

1 Introduction

In the process of speech acquisition, humans acquire a variety of regulatory mechanisms as well as the articulatory motions needed to produce specific vocal tract shapes. Recent studies [1, 2] have demonstrated that somatosensory inputs play a role in speech learning that is independent of the acoustic signal. However it is still unclear how somatosensory information is utilized in the speech learning process. It is also unclear which mechanoreceptors provide the needed kinesthetic information, since structures such as the perioral system lack muscle proprioceptors. Since the facial skin is deformed in different ways in conjunction with articulatory motions for different speech sounds, patterns of facial skin deformation may offer a new way of understanding sensory function in speech [3, 4].

Recently the involvement of somatosensory information associated with skin deformation in speech has been demonstrated [5]. However its functional role is still unclear not only in the on-going control of articulatory motion, but also in the acquisition process of speech articulation.

In this study, we focused on the role of somatosensory inputs associated with facial skin deformation in speech motor adaptation, and examined whether facial skin deformation over the course of a series of training trials affects articulatory motion in a progressive and adaptive manner.

2 Method

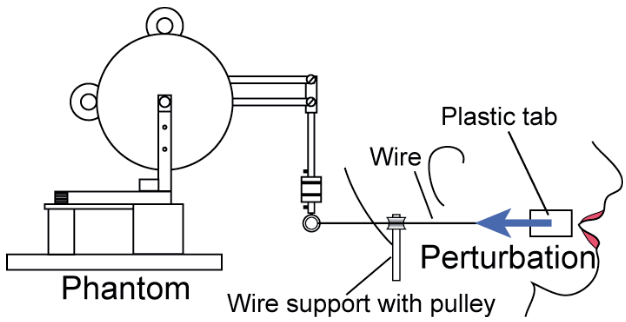

Figure 1: *Experimental setup of facial skin stretch perturbation*

Ten native speakers of American English participated in this test. The subjects were all young adults, had no neurological deficits. All subjects signed the approved Yale University Human Investigation Committee informed consent form.

Subjects were asked to repeat the utterance "see wood". We assessed the vertical aperture between the upper and lower lip at the point of maximum upper lip protrusion for the production of the /w/ sound. Upper lip, lower lip and head motion were sampled at 250 Hz by using the OPTOTRAK system (Northern digital Inc.). In off-line processing after the experiment, head motion was removed from articulatory position data using standard coordinate transform procedures.

We applied a perturbation load in conjunction with the speech utterance — a constant force load of 1.0 N that pulled the facial skin lateral to the oral angle in backward direction by using robotic device (Phantom 1.0, SensAble technorogies) as shown in Figure 1. The perturbation was timed so as to modulate somatosensory inputs during the period immediately preceding the target articulatory motion (see Figure 2a). The load was turned on at the start of each trial (in conjunction with the presentation of visual and auditory cue signals) and released at the speech onset for the /s/ sound. Hence, the load had no direct mechanical effect on the production of lip motion for the target sound.

We first recorded the normal trajectory over the course of 20 repetitions in the absence of load. The skin stretch perturbation was then applied over the next 200 repetitions as an adaptation phase. After the adaptation phase, articulatory motion without the skin stretch perturbation was then recorded (20 repetitions) as an aftereffect phase (see Figure 2b). To examine motor learning effects, we evaluated the upper lip trajectories for the production of /w/ over the course of training. We especially focused on differences in displacement in the null condition before and after learning and in the perturbed condition between beginning and end of learning.

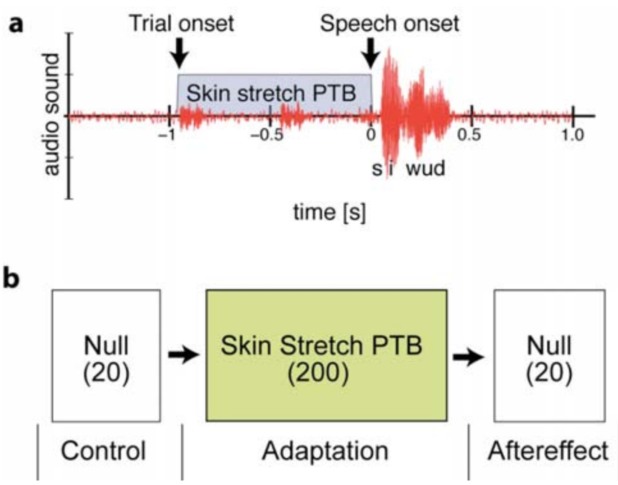

Figure 2: *(a) Temporal sequence of one perturbed trial in the adaptation phase (b) Block diagram showing sequence of motor adaptation session*

3 Results

Figure 3 shows changes in distance between the upper and lower lips across all trials in all ten subjects. Each panel corresponds to a different subject. The blue dots are the data in the null conditions before and after the adaptation phase. The red dots are the data recorded in skin stretch trials during the learning phase. Larger displacement magnitudes indicate wider aperture of the lips. We found that lip aperture gradually increased with training as an overall tendency in almost all subjects. Since the perturbation force was released before the onset of the lip protrusion for /w/ sound, the upper lip movement always occurred without external forces, even during learning phase. When questioned, some subjects reported that they were aware that the force had been released before, or just after, they started to speak, but some were not aware. Thus regardless of subject's awareness that the load that was no longer present, their upper lip tended to protrude more than in control trials, suggesting that the somatosensory information associated with skin deformation affects motor planning for the target speech task.

Lip aperture tended to increase across trials as shown by the regression line (black dashed-line in Figure 3). In addition, lip aperture in the aftereffect phase remained greater than normal, although subjects often showed a decrease in amplitude in comparison to the end of learning. This further supports the idea that the motor plan for the target task was modified in an adaptive manner on the basis of the skin stretch perturbation.

We calculated mean behavior for each subject by averaging the data from 20 repetitions in each phase of the experiment. We normalized the data using the z-transformation and took an average across all subjects to summarize average displacement. We compared mean aperture in four conditions: null control, 1st and last block of the adaptation phase and null condition aftereffect phase. The obtained results are shown in Figure 4 with larger positive values indicating wider aperture. The error bars show standard errors across subjects. As can be seen in the figure, lip aperture differed reliably among conditions (ANOVA: $p < 0.01$). In post-hoc tests with Bonferroni correction, there was a reliable difference between the control and after effect conditions (p <

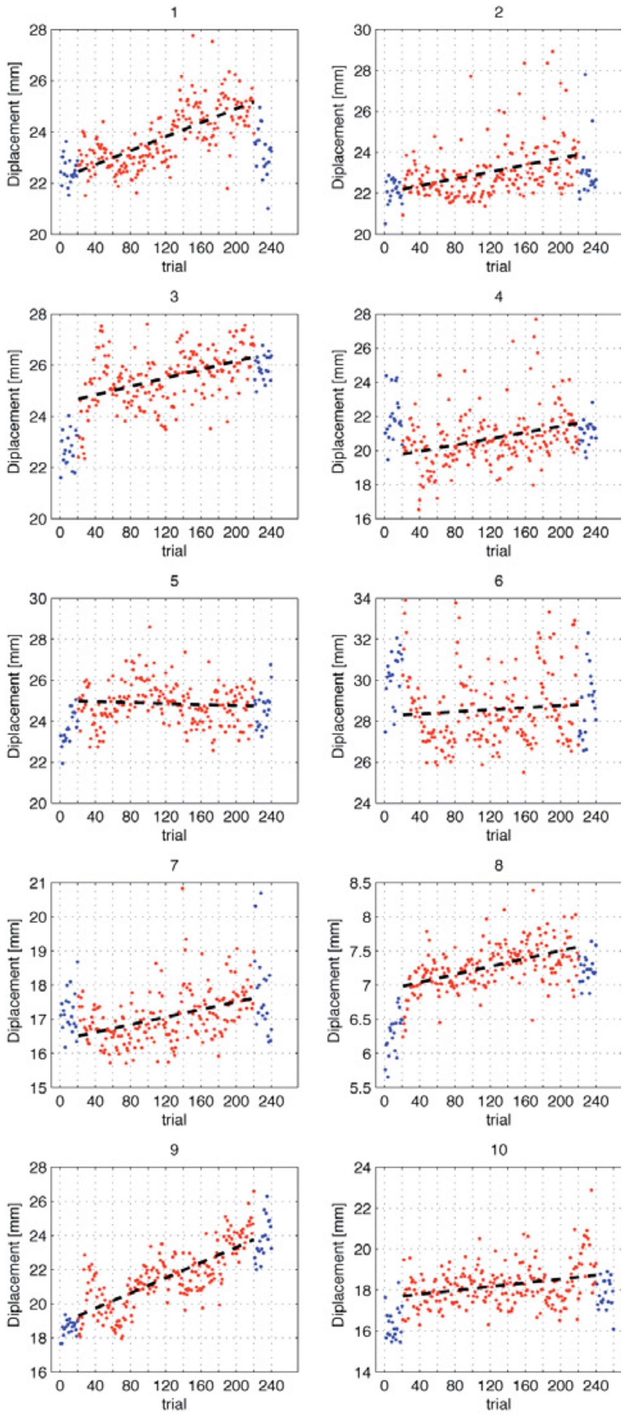

Figure 3: *Maximum vertical aperture between the upper and lower lips across the trials in all subjects.*

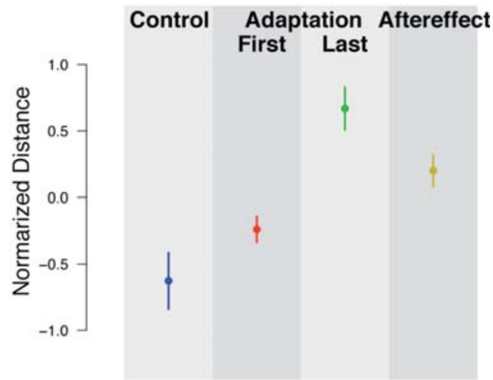

Figure 4: *Maximum vertical aperture between the upper and lower lips in initial null control phase, 1^{st} and the last blocks of the adaptation phase and in the null condition aftereffect phase. The error bars give standard error across the subjects.*

0.01) and between the first and last phase of adaptation phase ($p < 0.01$). On the other hand, there was no reliable difference between the control phase and the beginning of the adaptation phase ($p > 0.4$) nor between the end of adaptation phase and the aftereffect phase ($p > 0.2$). This indicates that lip motion was reliably modified in a course of training using facial skin stretch. These data support the idea that the altered somatosensory inputs modify the upper lip movement during speech motor leaning.

The result, that somatosensory inputs alter lip position in speech learning, is underscored by an acoustical analysis. We did not found any systematic modulation in first or second formant at 100 ms after peak displacement of the upper lip, although there was a small negative correlation between second formant frequency and peak vertical displacement of the upper lip over the trials. There were individual instances in which there were acoustical differences between conditions when the skin stretch perturbation was initially applied, when it is removed, indeed following the removal of all loads at the end of the experiment. However, these differences were inconsistent across the subjects and were not reliably corrected with lip motion. The results to date for this study thus suggest that somatosensory afferent input associated with skin deformation may modify the plan for articulatory motion in speech motor learning.

4 Discussion

The primary finding of this study was that modulating somatosensory inputs during a pre-motion phase modified the lip motion in adaptive manner on the basis of the skin stretch perturbation. The results are consistent with the idea that somatosensory afferent input associated with skin deformation may modify the plan for articulatory motion in speech motor learning.

Somatosensory function associated with facial skin deformation is a little recognized source of orofacial kinesthesia. Stretching the skin lateral to the oral angle, which is the area that we focused on in the present study, induced a cortical reflex that was associated with a modification of lip position in response to a sudden change in the position of the jaw [5]. The cutaneous mechanoreceptors lateral to the oral angle are activated during speech movements [3]. These findings suggest stretching the facial skin lateral to the oral angle could provide kinesthetic information concerning articulatory motion to the speech production system. The idea is also consistent with our result that the modulation of lip protrusion was induced in the direction opposite to facial skin stretch.

In general, when intended movements are not achieved for some reason, such as an external disturbance, motor commands are updated for the next motion by evaluating difference between intended and actual motion. In the current adaptation paradigm, target articulatory motions should be achieved in no force condition since the external load was removed just before the start of the lip protrusion movement. This means the subjects actually did not have to change their lip motion because there should have been no difference between intended and actual motion at any time, especially in somatosensory terms. Despite this, the subjects tended to change their lip motion in progressive manner and their lip motion was eventually achieved quite differently after adaptation. One possibility is that somatosensory modulation just before motion may affect sensory function during motion. In our case, facial skin stretch was applied in a direction opposite to the following movement. The resulting sensory input may have led the nervous system underestimate lip position. Consequently the actual motion may have always been evaluated as smaller than the intended one, and motor commands may have been updated to progressively yield a larger movement. This idea is consistent with limb studies using tendon vibration in which vibration just before the actual motion induced an underestimate of displacement [6]. We will further investigate this idea in future work.

5 Acknowledgements

This work was supported by the National Institute on Deafness and Other Communication Disorders Grant DC-04669. We thank Mark Tiede for advice and assistance.

References

[1] Tremblay, S., Shiller, D. M. & Ostry, D. J. Somatosensory basis of speech production. *Nature* 423, 866-869, 2003.

[2] Nasir, S. M. & Ostry, D. J. Somatosensory precision in speech production. *Curr Biol* 16: 1918-23, 2006.

[3] Johansson, R. S., Trulsson, M., Olsson, K. A. & Abbs, J. H. Mechanoreceptive afferent activity in the infraorbital nerve in man during speech and chewing movements. *Exp Brain Res* 72: 209-14, 1988.

[4] Connor, N. P. & Abbs, J. H. Movement-related skin strain associated with goal-oriented lip actions. *Exp Brain Res* 123: 235-41, 1998.

[5] Ito, T. & Gomi, H. Cutaneous mechanoreceptors contribute to the generation of a cortical reflex in speech. *Neuroreport* 18: 907-10 2007.

[6] Cordo P, Gurfinkel VS, Bevan L, and Kerr GK. Proprioceptive consequences of tendon vibration during movement. *J Neurophysiol* 74: 1675-1688, 1995.