

Changes of orofacial somatosensory attenuation during speech production

Takayuki Ito, Hiroki Ohashi, Vincent L Gracco

▶ To cite this version:

Takayuki Ito, Hiroki Ohashi, Vincent L Gracco. Changes of orofacial somatosensory attenuation during speech production. Neuroscience Letters, 2020, 730, 10.1016/j.neulet.2020.135045 . hal-02573957

HAL Id: hal-02573957 https://hal.science/hal-02573957

Submitted on 7 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1	Title: Changes of orofacial somatosensory attenuation during speech production
2	
3	Authors: Takayuki Ito ^{1,2*} , Hiroki Ohashi ² , Vincent L. Gracco ^{2,3}
4	
5	Author's affiliations:
6	1: Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 11 rue des Mathématiques,
7	Grenoble Campus BP46, F-38402, Saint Martin D'heres Cedex, France
8	2: Haskins Laboratories, 300 George Street, New Haven, CT, 06511, USA
9	3: McGill University, 2001 Avenue McGill College, Montréal, QC, H3A 1G1, Canada
10	
11	
12	
13	
14	

16 Abstract

17 Modulation of auditory activity occurs before and during voluntary speech movement. 18 However, it is unknown whether orofacial somatosensory input is modulated in the same 19 manner. The current study examined whether or not the somatosensory event-related 20 potentials (ERPs) in response to facial skin stretch are changed during speech and 21 nonspeech production tasks. Specifically, we compared ERP changes to somatosensory 22 stimulation for different orofacial postures and speech utterances. Participants produced 23 three different vowel sounds (voicing) or non-speech oral tasks in which participants 24 maintained a similar posture without voicing. ERP's were recorded from 64 scalp sites in 25 response to the somatosensory stimulation under six task conditions (three vowels \times 26 voicing/posture) and compared to a resting baseline condition. The first negative peak for 27 the vowel /u/ was reliably reduced from the baseline in both the voicing and posturing 28 tasks, but the other conditions did not differ. The second positive peak was reduced for 29 all voicing tasks compared to the posturing tasks. The results suggest that the sensitivity 30 of somatosensory ERP to facial skin deformation is modulated by the task and that 31 somatosensory processing during speaking may be modulated differently relative to 32 phonetic identity.

33

Keywords: Electroencephalography; event-related potentials; orofacial somatosensory
 processing; speech production; sensory suppression.

36

37 1. Introduction

Gating of sensory information around the onset of movement is a characteristic of many voluntary motor behaviors. Somatosensory attenuation has been reported during selfinitiated limb movement [1] and psychophysical detection threshold on the skin is increased during arm movement [2,3]. Somatosensory-evoked potentials from cortical and subcortical sites are attenuated during voluntary activity apparently to filter irrelevant signals during motor behavior with the magnitude of the gating dependent of the nature of the task [4–6].

45 In the speech motor system, the majority of studies investigating the modulation 46 of sensory processing during speech have focused on an auditory-mediated response. The 47 amplitude of auditory cortical potentials is reduced during speech production, the so-48 called speech induced suppression consistent with a down-regulation of reafferent 49 feedback [7–9]. Prior to the onset of speech, the amplitude of the first negative-going 50 potential (100 ms) and the first positive going potential (200 ms) are reduced compared to 51 silent reading with the effects on the first component stronger than the second component 52 [10,11]. For the somatosensory system, the behavioral data are consistent with a down-53 regulation of reafferent input. In the absence of auditory change, mechanical loads to the 54 facial skin during speech induce rapid compensatory movement for the production of 55 bilabial utterances with a significant cortical component [12,13]. The time-varying 56 reduction in compensation to mechanical loads to the lips during speech production and 57 increased vibrotactile detection thresholds during lip force generation [12,14] suggests 58 attenuation of somatosensory input during voluntary speech and nonspeech oral behavior.

However, the apparent gating of somatosensory input during speech and non-speechorofacial behavior has not been examined at the cortical level.

The current study was designed to assess the change in somatosenosory cortical processing during speech and nonspeech production and whether the modulation is dependent on different vowels and related orofacial postures. Somatosensory eventrelated potentials (ERPs) were induced by stretching the facial skin using a computercontrolled robotic device [15]. We observed somatosensory ERP changes from a resting baseline condition, and compared those changes during the different tasks and conditions.

67

68 **2. Methods**

69 2.1 Participants

Fifteen native speakers of American English (ten for the main experiment and five for a secondary experiment) participated in the study. The participants were all healthy young adults with normal hearing. There was no report of history of neurological problem. All participants signed informed consent forms approved by the Yale University Human Investigation Committee.

75

76 **2.2 EEG acquisition with speech production task**

The details of the somatosensory stimulation procedure with EEG are described in our previous studies [15]. Briefly, electroencephalography (EEG) was recorded using a 64electrode Biosemi ActiveTwo system (256 Hz sampling rate). Eye blinks and motion was recorded using electro-oculography. For somatosensory stimulation, a computercontrolled small robotic device applied skin stretch to the facial skin to evoke

82 somatosensory ERPs [15]. The skin stretch was produced by attaching two small plastic 83 tabs bilaterally with tape to the skin at the corners of the mouth. We applied a single 84 cycle of a 4-Hz sinusoid with 4 N maximum force in a rearward direction (relative to the 85 participant). We selected this pattern of stimulation based on our previous work 86 demonstrating changes in speech perceptual and physiological processing [16–18]. The 87 direction of stimulation was set by focusing on the horizontal articulatory contrast in the 88 task vowels, /i/ and /u/ described below. For each participant, we recorded 560 89 somatosensory ERPs associated with facial skin stretch and the speech produced by each 90 participant was recorded with a microphone (Sennheiser, ME66) with the acosutic signal 91 digitized at 22.05 kHz.

92

93 The participants were instructed to produce the vowels, /a/, /i/ and /u/ and their 94 related postures under two different conditions (silently and with voicing). The vowels 95 contrasted in terms of articulatory movement with the vowels /i/ and /u/ produced 96 primarily in a horizontal direction with /a/ produced primarily in a vertical direction. In 97 the voicing condition, the participants were instructed to sustain the vowels for 98 approximately 2 seconds. In the posture condition, the participants were instructed to 99 hold the posture for the utterance without producing any voice or airflow. In addition to 100 these six conditions (three tasks \times two conditions), a control condition was included in 101 which the skin stretch was applied while the participants were at rest. Each of the seven 102 conditions were repeated 80 times. The sequence of trials is presented in Figure 1. A 103 visual cue was used for the instructions. Somatosensory stimuli were applied 1.5 sec after 104 the trial onset. All trials were presented in pseudo-random order with the constraint that all seven experiment conditions were tested every seven trials in order to minimize anadaptation.

107

108 **2.3 Data analysis**

109 EEG signals were filtered using a 1-30 Hz band-pass filter and re-referenced to the 110 average across all electrodes. Continuous EEG signals were segmented into epochs 111 between -500 and 1000 ms relative to the stimulus onset. A bias level of each epoch was 112 adjusted to the average amplitude in the pre-stimulus interval (-200 to -100 ms). 113 Independent component analysis [19] was applied and the component related to large 114 signal noise and artifacts including eye-blink and movement were excluded by manual 115 inspection. Finally, the processed ERPs were averaged across trials in each condition on a 116 per-participant basis.

As shown in previous studies, the largest amplitude of somatosensory ERPs from facial skin stretch were obtained at the electrodes around Fz in the 10-20 system. In the current analysis, we took a spatial average over Fz and four surrounding electrodes (F1, F2, AFz, and FCz) in order to improve the signal-to-noise ratio. The peak amplitude in the first negative peak and the following positive peak were captured using a 40 ms time window.

In order to examine the change from the resting baseline condition, amplitudes of the positive and negative peaks in the task conditions were normalized to those in the resting baseline condition by subtraction. Repeated-measures ANOVA was applied to assess changes in the task condition (vowel utterances and speaking manner). The overall

analyses were followed by Bonferroni corrected multiple comparisons and t-test toevaluate any reliable differences from the resting baseline condition.

129

130 **2.4 Estimating the displacement of the skin stretch under the different conditions**

131 We carried out a follow-up test to assess the actual displacements of the skin due to facial 132 skin stretch in five native speakers of American English using electromagnetic 133 articulometry (Northern Digital Inc., Wave). The sensor was place on the plastic tabs for 134 the skin stretch stimulation. The same tasks in the main experiment (three vowels \times two 135 speaking manners) were carried out 10 times each in random order. The same 136 somatosensory stimulation was applied (4N) and the amplitude of facial skin 137 displacement was derived by comparing maximum displacement to that during the 138 baseline period prior to the stimulation. Repeated measures ANOVA was was used to 139 compare the change in disaplacement by conditions.

140

141 **3. Results**

142 The effect of facial skin stretch was quantified using somatosensory ERPs during the 143 speech and orofacial postures. We measured the amplitude of the first negative peak and 144 the following positive peak of the somatosensory ERPs across the conditions and tasks 145 and compared the ERPs to the base-line (rest) condition. Figure 2A presents the 146 somatosensory ERPs at the mid-sagittal frontal area (around Fz) for the different tasks. 147 Each color in the figure represents a different vowel for the voicing or a different posture; 148 the average ERP for the rest condition is superimposed (black-dashed line). Overall, the 149 ERP response exhibits a negative peak around 140 ms followed by a positive peak around

150 250 ms for all tasks and conditions. For the first negative peak in the voicing condition 151 (top panel in Fig. 2A), the ERP amplitude for i/i is similar to the control (rest) condition 152 while the ERP for /a/ and /u/ are reduced. For the posturing task (bottom panel in Fig. 153 2A), the ERP amplitude for /u/ is slightly reduced from rest, with the other two postures 154 are similar to the control condition. For the following positive peak, the amplitudes for all 155 vowels were reduced relative to the rest condition. For the posturing task, the peak 156 amplitude of /u/ was similar to the rest condition, but the other two the amplitudes were reduced. 157

158 The peak amplitudes relative to the resting baseline condition are summarized in 159 Figures 2B and C. For the negative peak (Fig. 2B), a two-way ANOVA resulted in significantly different peak amplitudes for the vowel condition [F(2,45) = 6.30, p < 0.01,160 $\eta_p^2 = 0.219$], but not for the different tasks [F(1,45) = 2.456, p > 0.1] with no reliable 161 interaction [F(2,45) = 0.08, p > 0.9]. Pairwise t-test with Bonferroni correction 162 demonstrated that the amplitude in /u/ was different from /a/ (p < 0.05) and /i/ (p < 0.001) 163 164 with no difference between $\frac{a}{and}$ and $\frac{i}{(p > 0.9)}$. In addition, the amplitude relative to the 165 resting response was significantly different from zero in /u/(p < 0.05), but not in /i/(p > 0.05)0.9) and in $\frac{a}{p} > 0.8$. The results indicate that the somatosensory ERP was reduced in 166 167 the production for /u/, but not for /a/ and /i/ in both conditions (voicing and posturing) 168 and suggest that the somatosensory attenuation can be induced depending on the vowel 169 production.

For the positive peaks (Fig. 2C), a two-way ANOVA resulted in a significant difference for the conditions [F(1,45) = 4.583, p < 0.05, $\eta_p^2 = 0.092$], but not in the tasks [F(2,45) = 0.062, p > 0.9] with no reliable interaction [F(2,45) = 0.95, p > 0.3]. Post-hoc testing found that the amplitude relative to the resting response was significantly different from zero in voicing task (p < 0.02), but not in posturing task (p > 0.1). This result suggests that the generation of voicing can suppress the somatosensory processing related to the generation of positive peak of ERPs when the motor system was actively involved by producing the vowel sound.

The results of the follow-up examination of the skin displacement for the differnt conditions and tasks are presented in Figure 3. The averaged value represented the skin was stretched around 6.5 mm in all tasks including speaking manner and vowels. Oneway ANOVA indicated no reliable difference across all tasks [F(5,20) = 1.116, p > 0.3].

182

183

184 **4. Discussion**

We found that somatosensory ERPs elicited with the facial skin deformation were 185 186 attenuated during speech production. This attenuation depended on the specific vowel 187 and production manner. The amplitude of the first negative peak of the somatosensory 188 ERP for the vowel /u/ was reliably attenuated from the resting condition, but not for the 189 vowels /a/ and /i/. This attenuation was consistent across the production conditions, 190 suggesting that the somatosensory attenuation of the orofacial system may be specific to 191 the vocal tract configuration, at least in terms of the early feedback processing. On the 192 other hand, the effect of voicing was seen on the following positive peaks. The peak 193 amplitude was attenuated in the voicing condition, but not in the posturing condition. 194 This attenuation was consistent across the vowels compared to the first negative peak. 195 This suggests that overt speech can modulate the orofacial system in different ways

- depending on the time window. The different patterns of somatosensory ERP response
- 197 may reflect different sources or mechanisms of orofacial somatosensory processing.
- 198

199 Sensory attenuation during speech movement has been investigated mostly in 200 response to auditory input. Auditory evoked responses from self-produced speech are 201 suppressed, the so-called speech induce suppression [7–9]. The mechanism of speaking-202 induced suppression is considered as a partial neural cancellation of incoming sensory 203 feedback as it is matched to the motor prediction. It has been suggested the amplitude 204 reduction may reflect the amount of error between the motor prediction and incoming 205 sensory feedback, and hence the amplitude of attenuation is reduced when the feedback is 206 different from the predicted speech [8,9]. The current somatosensory attenuation at 100 207 ms after stimulus onset may reflect a similar mechanism although we only observed an 208 early reduction (the negative-going peak) for the vowel /u/. Our previous work on the 209 somatosensory stimulation during speech perception demonstrated directionally-sensitive 210 interactions dependent on the movement direction associated with the vowel [18]. In the 211 current paradigm the direction of stimulation (a horizontal stretch) was maximally in the 212 same direction as for the vowel /u/with no overlap for /a/and minimal overlap for /i/. We 213 suggest that early somatosensory processing may have access only to the motoneuron 214 pools that are active for the specific vowel or vocal tract configuration for the action 215 being produced.

216

217 In addition to the first negative peak, we also found attenuation of the second 218 (positive) peak with more attenuation in the voicing condition. In contrast to the early 219 negative response, all vowels showed a reduction with the largest reduction for the vowel 220 /u/. The attenuation of somatosensory ERP during speaking appears consistent with the 221 finding from Daliri and Max [10], in which auditory attenuation in P2 was induced by 222 speech but not by non-speech stimulation. In addition, the posturing condition showed 223 less attenuation compared to the voicing condition. This suggests that the somatosensory-224 motor network involved in voicing and posturing condition may be different. This 225 difference may be related to the difference in control strategy between speech and non-226 speech tasks. The previous behavioral study showed that compensatory response in 227 adaptation is specific in a speech task, but not in a non-speech task [20]. Similarly, when 228 the real-line compensatory response in the tongue to a mechanical perturbation was 229 examined for voicing and posturing tasks similar to the current experimental tasks, a 230 systematic change of compensation was found in the voicing condition only [21]. These 231 observations suggest that specific motor control mechanisms are involved for speech 232 compared with posturing even though the posturing condition contained related orofacial 233 movements. Hence the difference in attenuation of the second positive peak of the 234 somatosensory ERP for the speech task reflects the involvement of a speech-specific 235 control mechanism modulating somatosensory processing.

236

We carried out a follow-up test to examine whether the current ERP modulation can be caused by a difference in displacement by facial skin stretch among the task vowels. Facial configurations are different among the three vowels. For example, the jaw is opened more for the production of /a/ than /i/ and /u/, and the lip are spread more for the production of /i/ than the other two. There is hence a possibility that the same

amplitude of force (4 N at the peak) may induce a different displacement of the facial skin deformation according to the facial configuration. Since we found displacements of the facial skin stretch was similar across all tasks, we conclude that the current somatosensory stimulation was consistent across the tasks in terms of displacement of facial skin deformation.

247

248 Somatosensory attenuation has been examined mostly in the limb system 249 [1,22,23]. Although there was a behavioral evidence to support somatosensory 250 attenuation [12,14], it is still unknown how orofacial somatosensory ERPs are modified 251 during orofacial movement including speaking. The current results provide evidence that 252 orofacial somatosensory system can be attenuated during facial movement similar to the 253 limb system. The attenuation has been, however, observed mostly in short-latency ERPs 254 (less than 100 ms after stimulation) using brief electrical stimulation. In contrast, typical 255 pattern of orofacial ERP using facial skin deformation has a negative peak between 100 256 and 200 ms and positive peak between 200 and 300 ms on the mid-sagittal frontal area, 257 similar to an auditory ERP [15,17,16]. Considering the relatively long-latency of the 258 orofacial somatosensory attenuation together with the task-specific nature of the response, 259 the observed attenuation most likely reflects a cortical level of sensory modulation.

260

261

262 **5. Conclusion**

We found orofacial somatosensory ERPs are attenuated during speech production.
This attenuation was modulated depending on speech vowels and articulatory conditions.

- 265 The results suggest that the processing of orofacial somatosensory inputs can be regulated
- 266 differently relative to phonetic identity and speaking manner.

Acknowledgements

This work was supported by the National Institute on Deafness and Other Communication Disorders Grants R21DC012502, R21DC013915 and R01DC017439.

References

- D.N. Rushton, J.C. Rothwell, M.D. Craggs, Gating of somatosensory evoked potentials during different kinds of movement in man, Brain. 104 (1981) 465–491.
- [2] R.W. Angel, R.C. Malenka, Velocity-dependent suppression of cutaneous sensitivity during movement, Experimental Neurology. 77 (1982) 266–274. https://doi.org/10.1016/0014-4886(82)90244-8.
- [3] C.E. Chapman, M.C. Bushnell, D. Miron, G.H. Duncan, J.P. Lund, Sensory perception during movement in man, Exp Brain Res. 68 (1987) 516–24.
- [4] T. Allison, G. McCarthy, C.C. Wood, S.J. Jones, Potentials evoked in human and monkey cerebral cortex by stimulation of the median nerve. A review of scalp and intracranial recordings, Brain. 114 (Pt 6) (1991) 2465–2503. https://doi.org/10.1093/brain/114.6.2465.
- [5] A. Starr, L.G. Cohen, "Gating" of somatosensory evoked potentials begins before the onset of voluntary movement in man, Brain Res. 348 (1985) 183–186. https://doi.org/10.1016/0006-8993(85)90377-4.
- Y. Lei, M.A. Perez, Cortical contributions to sensory gating in the ipsilateral somatosensory cortex during voluntary activity, The Journal of Physiology. 595 (2017) 6203–6217. https://doi.org/10.1113/JP274504.
- [7] G. Curio, G. Neuloh, J. Numminen, V. Jousmäki, R. Hari, Speaking modifies voiceevoked activity in the human auditory cortex, Hum Brain Mapp. 9 (2000) 183–191.
- [8] J.F. Houde, S.S. Nagarajan, K. Sekihara, M.M. Merzenich, Modulation of the auditory cortex during speech: an MEG study, J Cogn Neurosci. 14 (2002) 1125– 1138. https://doi.org/10.1162/089892902760807140.

- [9] C.A. Niziolek, S.S. Nagarajan, J.F. Houde, What does motor efference copy represent? Evidence from speech production, J Neurosci. 33 (2013) 16110–6. https://doi.org/10.1523/JNEUROSCI.2137-13.2013.
- [10] A. Daliri, L. Max, Modulation of Auditory Responses to Speech vs. Nonspeech Stimuli during Speech Movement Planning, Front Hum Neurosci. 10 (2016) 234. https://doi.org/10.3389/fnhum.2016.00234.
- [11] A. Daliri, L. Max, Modulation of auditory processing during speech movement planning is limited in adults who stutter, Brain Lang. 143 (2015) 59–68. https://doi.org/10.1016/j.bandl.2015.03.002.
- [12] V.L. Gracco, J.H. Abbs, Dynamic Control of perioral system during speech: kinematic analysis of autogenic and nonautogenic sensorimotor processes, J. of Neurophysiology. 54 (1985) 418–432.
- T. Ito, H. Gomi, Cutaneous mechanoreceptors contribute to the generation of a cortical reflex in speech, Neuroreport. 18 (2007) 907–10. https://doi.org/10.1097/WNR.0b013e32810f2dfb.
- [14] R.D. Andreatta, S.M. Barlow, Movement-related modulation of vibrotactile detection thresholds in the human orofacial system, Exp Brain Res. 149 (2003) 75– 82. https://doi.org/10.1007/s00221-002-1336-x.
- [15] T. Ito, D.J. Ostry, V.L. Gracco, Somatosensory event-related potentials from orofacial skin stretch stimulation, Journal of Visualized Experiments : JoVE. (2015) e53621. https://doi.org/10.3791/53621.

- [16] T. Ito, J.H. Coppola, D.J. Ostry, Speech motor learning changes the neural response to both auditory and somatosensory signals, Scientific Reports. 6 (2016) 25926. https://doi.org/10.1038/srep25926.
- [17] T. Ito, V.L. Gracco, D.J. Ostry, Temporal factors affecting somatosensory-auditory interactions in speech processing, Frontiers in Psychology. 5 (2014) 1198. https://doi.org/10.3389/fpsyg.2014.01198.
- [18] T. Ito, M. Tiede, D.J. Ostry, Somatosensory function in speech perception, Proc Natl Acad Sci U S A. 106 (2009) 1245–8. https://doi.org/10.1073/pnas.0810063106.
- [19] J. Onton, M. Westerfield, J. Townsend, S. Makeig, Imaging human EEG dynamics using independent component analysis, Neurosci Biobehav Rev. 30 (2006) 808–22. https://doi.org/10.1016/j.neubiorev.2006.06.007.
- [20] S. Tremblay, D.M. Shiller, D.J. Ostry, Somatosensory basis of speech production, Nature. 423 (2003) 866–869.
- [21] T. Ito, J.-L. Caillet, P. Perrier, Stability in postural tongue control: response to transient mechanical perturbations, in: Annual Meeting of the Society for Neuroscience (Neuroscience 2018), San Diego, United States, 2018. https://hal.archives-ouvertes.fr/hal-01919048 (accessed September 26, 2019).
- [22] M.C. Tapia, L.G. Cohen, A. Starr, Selectivity of attenuation (i.e., gating) of somatosensory potentials during voluntary movement in humans, Electroencephalogr Clin Neurophysiol. 68 (1987) 226–230.
- [23] J. Vrána, H. Polácek, A. Stancák, Somatosensory-evoked potentials are influenced differently by isometric muscle contraction of stimulated and non-stimulated hand

in humans, Neurosci. Lett. 386 (2005) 170-175.

https://doi.org/10.1016/j.neulet.2005.06.005.

Figure Legends

Figure 1: Temporal sequence of visual cue, auditory recording and somatosensory stimulation in one trial

Figure 2: A: Somatosensory event-related potential responses. B: Averaged amplitude of the first negative peak (top panel) and the second positive peak (bottom panel). Error bar represents the standard error across the participants.

Figure 3: Maximum displacement due to facial skin deformation during the vowel production task. Error bar represents the standard error across the participants.







