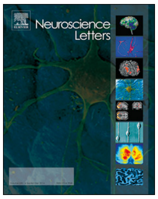
Neuroscience Letters 730 (2020) 135045

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet



Research article

Changes of orofacial somatosensory attenuation during speech production

Takayuki Ito[a,b,*], Hiroki Ohashi[b], Vincent L. Gracco[b,c]

[a] Grenoble Alpes University, CNRS, Grenoble INP, GIPSA-lab, 11 rue des Mathématiques, Grenoble Campus BP46, F-38402 Saint Martin D'heres Cedex France
[b] Haskins Laboratories, 300 George Street, New Haven, CT 06511, USA
[c] McGill University, 2001 Avenue McGill College, Montréal, QC H3A 1G1, Canada

1988

ARTICLE INFO

Keywords:
Electroencephalography
Event-related potentials
Orofacial somatosensory processing
Speech production
Sensory suppression

ABSTRACT

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

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 self-initiated 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].

In the speech motor system, the majority of studies investigating the modulation of sensory processing during speech have focused on an auditory-mediated response. The amplitude of auditory cortical potentials is reduced during speech production, the so-called speech induced suppression consistent with a down-regulation of reafferent feedback [7–9]. Prior to the onset of speech, the amplitude of the first negative-going potential (100 ms) and the first positive going potential (200 ms) are reduced compared to silent reading with the effects on the first component stronger than the second component [10,11]. For the somatosensory system, the behavioral data are consistent with a down-regulation of reafferent input. In the absence of auditory change, mechanical loads to the facial skin during speech induce rapid compensatory movement for the production of bilabial utterances with a significant cortical component [12,13]. The time-varying reduction in compensation to mechanical loads to the lips during speech production and increased vibrotactile detection thresholds during lip force generation [12,14] suggests attenuation of somatosensory input during voluntary speech and nonspeech oral behavior. However, the apparent gating of somatosensory input during speech and non-speech orofacial 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 event-related potentials (ERPs) were induced by stretching the facial skin using a computer-controlled robotic device [15]. We observed somatosensory ERP changes from a resting baseline condition, and compared those changes during the different tasks and conditions.

2. Methods

2.1. Participants

Fifteen native speakers of American English (ten for the main

---

[*] Corresponding author at: GIPSA-Lab, CNRS, Univ. Grenoble-Alpes, Grenoble-INP, 11 rue des Mathématiques, Grenoble Campus BP46, F-38402 Saint Martin D'heres Cedex, France.

E-mail address: takayuki.ito@gipsa-lab.grenoble-inp.fr (T. Ito).

https://doi.org/10.1016/j.neulet.2020.135045
Received 20 January 2020; Received in revised form 28 April 2020; Accepted 7 May 2020

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.

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 64-electrode Biosemi ActiveTwo system (256 Hz sampling rate). Eye blinks and motion was recorded using electro-oculography. For somatosensory stimulation, a computer-controlled small robotic device applied skin stretch to the facial skin to evoke somatosensory ERPs [15]. The skin stretch was produced by attaching two small plastic tabs bilaterally with tape to the skin at the corners of the mouth. We applied a single cycle of a 4-Hz sinusoid with 4 N maximum force in a rearward direction (relative to the participant). We selected this pattern of stimulation based on our previous work demonstrating changes in speech perceptual and physiological processing [16–18]. The direction of stimulation was set by focusing on the horizontal articulatory contrast in the task vowels, /i/ and /u/ described below. For each participant, we recorded 560 somatosensory ERPs associated with facial skin stretch and the speech produced by each participant was recorded with a microphone (Sennheiser, ME66) with the acosutic signal digitized at 22.05 kHz.

The participants were instructed to produce the vowels, /a/, /i/ and /u/ and their related postures under two different conditions (silently and with voicing). The vowels contrasted in terms of articulatory movement with the vowels /i/ and /u/ produced primarily in a horizontal direction with /a/ produced primarily in a vertical direction. In the voicing condition, the participants were instructed to sustain the vowels for approximately 2 s. In the posture condition, the participants were instructed to hold the posture for the utterance without producing any voice or airflow. In addition to these six conditions (three tasks × two conditions), a control condition was included in which the skin stretch was applied while the participants were at rest. Each of the seven conditions were repeated 80 times. The sequence of trials is presented in Fig. 1. A visual cue was used for the instructions. Somatosensory stimuli were applied 1.5 s after 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 an adaptation.

2.3. Data analysis

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

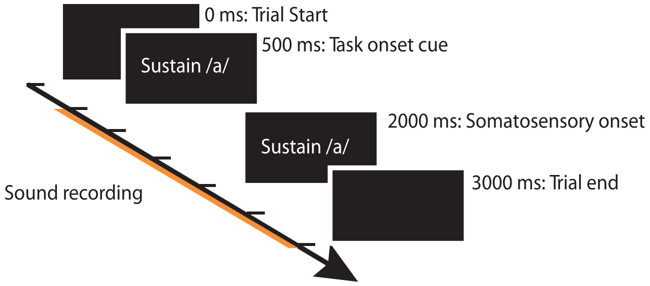


Fig. 1. Temporal sequence of visual cue, auditory recording and somatosensory stimulation in one trial.

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 to evaluate any reliable differences from the resting baseline condition.

2.4. Estimating the displacement of the skin stretch under the different conditions

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

3. Results

The effect of facial skin stretch was quantified using somatosensory ERPs during the speech and orofacial postures. We measured the amplitude of the first negative peak and the following positive peak of the somatosensory ERPs across the conditions and tasks and compared the ERPs to the base-line (rest) condition. Fig. 2A presents the somatosensory ERPs at the mid-sagittal frontal area (around Fz) for the different tasks. Each color in the figure represents a different vowel for the voicing or a different posture; the average ERP for the rest condition is superimposed (black-dashed line). Overall, the ERP response exhibits a negative peak around 140 ms followed by a positive peak around 250 ms for all tasks and conditions. For the first negative peak in the voicing condition (top panel in Fig. 2A), the ERP amplitude for /i/ is similar to the control (rest) condition while the ERP for /a/ and /u/ are reduced. For the posturing task (bottom panel in Fig. 2A), the ERP amplitude for /u/ is slightly reduced from rest, with the other two postures are similar to the control condition. For the following positive peak, the amplitudes for all vowels were reduced relative to the rest condition. For the posturing task, the peak amplitude of /u/ was similar to the rest condition, but the other two the amplitudes were reduced.

The peak amplitudes relative to the resting baseline condition are summarized in Fig. 2B and C. For the negative peak (Fig. 2B), a two-way ANOVA resulted in significantly different peak amplitudes for the vowel condition [$F(2,45) = 6.30$, $p < 0.01$, $\eta_p^2 = 0.219$], but not for the different tasks [$F(1,45) = 2.456$, $p > 0.1$] with no reliable interaction [$F(2,45) = 0.08$, $p > 0.9$]. Pairwise *t*-test with Bonferroni correction demonstrated that the amplitude in /u/ was different from /a/ ($p < 0.05$) and /i/ ($p < 0.001$) with no difference between /a/ and /i/ ($p > 0.9$). In addition, the amplitude relative to the resting response was significantly different from zero in /u/ ($p < 0.05$), but not in /i/ ($p > 0.9$) and in /a/ ($p > 0.8$). The results indicate that the

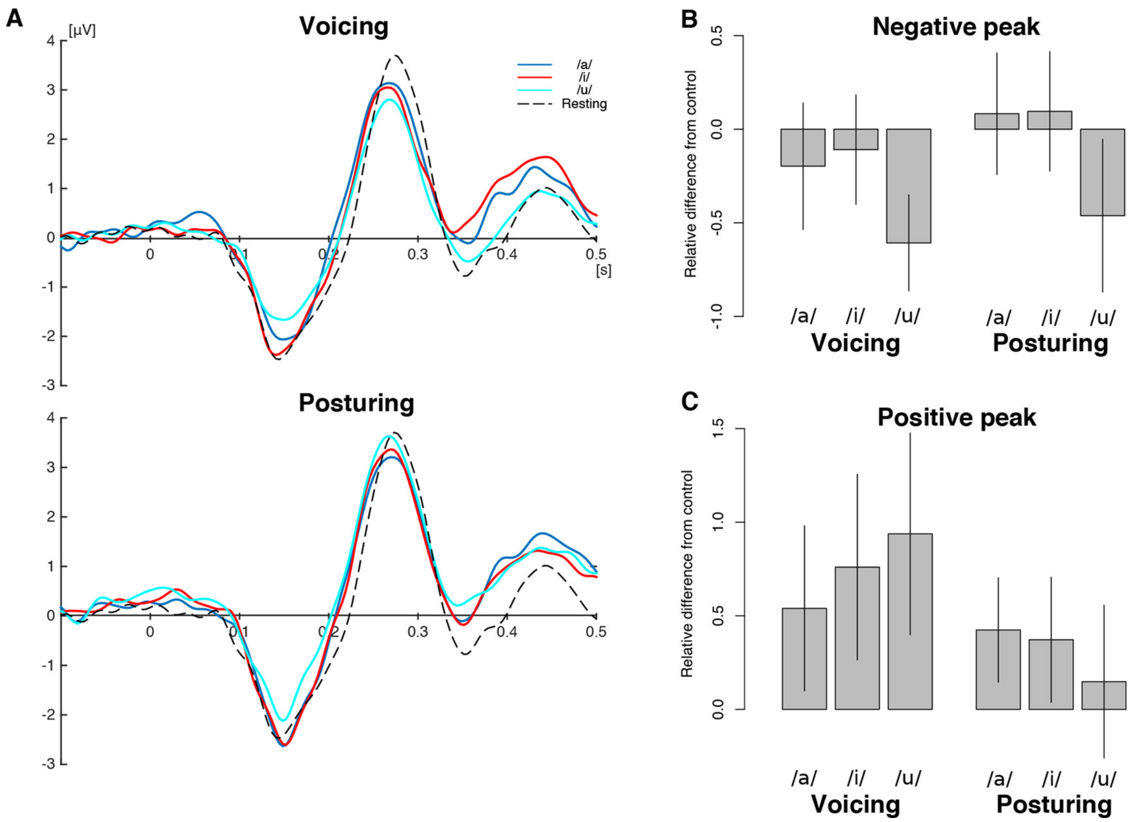


Fig. 2. (A) Somatosensory event-related potential responses. (B) Averaged amplitude of the first negative peak (top panel) and the second positive peak (bottom panel). Error bar represents the standard error across the participants.

somatosensory ERP was reduced in the production for /u/, but not for /a/ and /i/ in both conditions (voicing and posturing) and suggest that the somatosensory attenuation can be induced depending on the vowel 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 different conditions and tasks are presented in Fig. 3. The averaged value represented the skin was stretched around 6.5 mm in all tasks including speaking manner and vowels. One-way ANOVA indicated no reliable difference across all tasks [$F(5,20) = 1.116$, $p > 0.3$].

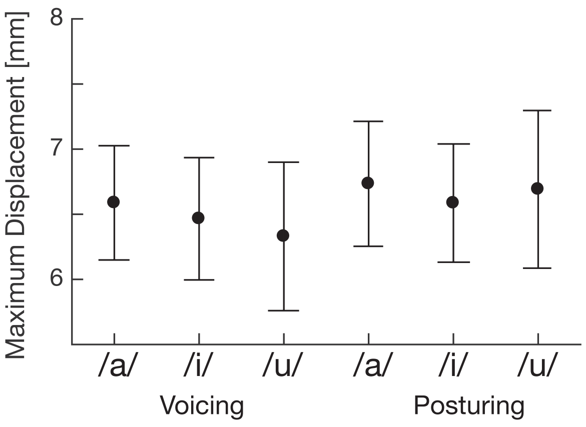


Fig. 3. Maximum displacement due to facial skin deformation during the vowel production task. Error bar represents the standard error across the participants.

4. Discussion

We found that somatosensory ERPs elicited with the facial skin deformation were attenuated during speech production. This attenuation depended on the specific vowel and production manner. The amplitude of the first negative peak of the somatosensory ERP for the vowel /u/ was reliably attenuated from the resting condition, but not for the vowels /a/ and /i/. This attenuation was consistent across the production conditions, suggesting that the somatosensory attenuation of the orofacial system may be specific to the vocal tract configuration, at least in terms of the early feedback processing. On the other hand, the effect of voicing was seen on the following positive peaks. The peak amplitude was attenuated in the voicing condition, but not in the posturing condition. This attenuation was consistent across the vowels compared to the first negative peak. 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 may reflect different sources or mechanisms of orofacial somatosensory processing.

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

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

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.

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

5. Conclusion

We found orofacial somatosensory ERPs are attenuated during speech production. This attenuation was modulated depending on speech vowels and articulatory conditions. The results suggest that the processing of orofacial somatosensory inputs can be regulated differently relative to phonetic identity and speaking manner.

CRediT authorship contribution statement

Takayuki Ito: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition. **Hiroki Ohashi:** Methodology, Software, Investigation, Writing - review & editing. **Vincent L. Gracco:** Conceptualization, Methodology, Writing - review & editing, Funding acquisition.

Acknowledgements

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

References

[1] 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, Exp. Neurol. 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) 516–524.

[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 (6) (1991) 2465–2503, https://doi.org/10.1093/brain/114.6.2465 Pt.

[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.

[6] Y. Lei, M.A. Perez, Cortical contributions to sensory gating in the ipsilateral somatosensory cortex during voluntary activity, J. Physiol. (Lond.) 595 (2017) 6203–6217, https://doi.org/10.1113/JP274504.

[7] G. Curio, G. Neuloh, J. Numminen, V. Jousmäki, R. Hari, Speaking modifies voice-evoked 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) 16110–16116, 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. Neurophysiol. 54 (1985) 418–432.

[13] T. Ito, H. Gomi, Cutaneous mechanoreceptors contribute to the generation of a cortical reflex in speech, Neuroreport 18 (2007) 907–910, 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, J. Vis. Exp. (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, Sci. Rep. 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, Front. Psychol. 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) 1245–1248, 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) 808–822, 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, Annual Meeting of the Society for Neuroscience (Neuroscience 2018), San Diego, United States, (2018) (accessed September 26, 2019), https://hal.archivesouvertes.fr/hal-01919048.

[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.