Left Lateralized Enhancement of Orofacial Somatosensory Processing Due to Speech Sounds

Takayuki Ito[a], **Alexis R. Johns**[a,b], and **David J. Ostry**[a,c]

[a]Haskins Laboratories, New Haven, CT

[b]University of Connecticut, Storrs

[c]McGill University, Montreal, Quebec, Canada

Abstract

Purpose—Somatosensory information associated with speech articulatory movements affects the perception of speech sounds and vice versa, suggesting an intimate linkage between speech production and perception systems. However, it is unclear which cortical processes are involved in the interaction between speech sounds and orofacial somatosensory inputs. The authors examined whether speech sounds modify orofacial somatosensory cortical potentials that were elicited using facial skin perturbations.

Method—Somatosensory event-related potentials in EEG were recorded in 3 background sound conditions (pink noise, speech sounds, and nonspeech sounds) and also in a silent condition. Facial skin deformations that are similar in timing and duration to those experienced in speech production were used for somatosensory stimulation.

Results—The authors found that speech sounds reliably enhanced the first negative peak of the somatosensory event-related potential when compared with the other 3 sound conditions. The enhancement was evident at electrode locations above the left motor and premotor area of the orofacial system. The result indicates that speech sounds interact with somatosensory cortical processes that are produced by speech-production-like patterns of facial skin stretch.

Conclusion—Neural circuits in the left hemisphere, presumably in left motor and premotor cortex, may play a prominent role in the interaction between auditory inputs and speech-relevant somatosensory processing.

Keywords

neuroimaging; speech perception; speech production

There is longstanding interest in the functional linkage between speech production and speech perceptual processing (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). In this context, the majority of studies to date have focused on motor function (or motor

---

cortex) and its role in the perception of speech sounds (D'Ausilio et al., 2009; Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007; Möttönen & Watkins, 2009; Watkins, Strafella, & Paus, 2003; Wilson, Saygin, Sereno, & Iacoboni, 2004). On the other hand, the possible role of the somatosensory system in speech perception has been largely overlooked, despite the importance of the somatosensory system in speech production and motor control. Recent psychophysical studies that involve the use of facial skin deformation have demonstrated a reciprocal interaction between orofacial somatosensation and audition in speech processing (Ito & Ostry, 2012; Ito, Tiede, & Ostry, 2009). However, virtually nothing is known about the neural correlates of orofacial somatosensory processing in the perception of speech sounds.

The brain mechanisms of somatosensory–auditory interaction in nonspeech processes have been investigated in humans (Beauchamp, Yasar, Frye, & Ro, 2008; Foxe et al., 2002; Lütkenhöner, Lammertmann, Simões, & Hari, 2002; Murray et al., 2005; Schürmann, Caetano, Hlushchuk, Jousmäki, & Hari, 2006) and in other primates (Fu et al., 2003; Kayser, Petkov, Augath, & Logothetis, 2005; Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007). However, because speech and nonspeech sounds are processed differently in the brain (Kozou et al., 2005; Möttönen et al., 2006; Thierry, Giraud, & Price, 2003), it is important to know the extent to which the cortical regions associated with somatosensory–auditory interaction in non-speech processes are likewise involved in speech processing. In addition, studies that have explored this interaction have assessed somatosensory cortical processes using stimuli such as light touch, vibrotactile stimuli, or brief tapping, which have little or no relation to somatosensory inputs during speech articulatory motion. Given the kinesthetic contribution of cutaneous mechanoreceptors to speech motor control, as shown previously using facial skin deformation (Ito & Gomi, 2007; Ito & Ostry, 2010), investigating the orofacial somatosensory system using facial skin deformation could offer a new way of understanding the functional linkage between speech production and perception in terms of a somatosensory–auditory interaction.

The cortical processing of speech and language is generally found to be left-hemisphere dominant (Damasio & Geschwind, 1984). Left lateralization effects have been examined in speech production (Chang, Kenney, Loucks, Poletto, & Ludlow, 2009; Ghosh, Tourville, & Guenther, 2008; Simonyan, Ostuni, Ludlow, & Horwitz, 2009) and in non-speech orofacial movements (Arima et al., 2011; Malandraki, Sutton, Perlman, Karampinos, & Conway, 2009). Left lateralized circuits, in left posterior superior temporal gyrus, may be responsible for auditory-motor integration in speech (see Hickok & Poeppel, 2007, for a review). Although the role of somatosensory inputs in this process is not known, there is reason to expect that left cortical circuits will similarly be predominant for speech-related somatosensory–auditory interaction.

We investigated somatosensory event-related potentials (ERPs) due to speech-production-like patterns of facial skin stretch, and we assessed whether speech sounds affect orofacial somatosensory processing. The focus here is on temporal processing of somatosensory signals in the presence of speech. This approach complements the existing literature that is based largely on fMRI techniques, which offer high spatial resolution but poor temporal resolution. In the present test, somatosensory ERPs were recorded in response to

somatosensory stimulation (facial skin stretch) in three background sound conditions (speech sounds, nonspeech sounds, and pink noise) and a silent control condition. We used a robotic device to generate patterns of facial skin deformation that are similar in timing and duration to those experienced in speech production. We specifically examined whether speech background sounds in particular modulate somatosensory ERPs when compared with the other background sound conditions. The findings complement our previous psychophysical tests that show that speech sounds alter orofacial somatosensory judgments associated with facial skin deformation (Ito & Ostry, 2012) and underscore the linkage between the production and perception of speech in the cortical processing of speech.

Method

Participants

Nine native speakers of American English (six women, three men) participated in the experiment. The participants were healthy young adults with normal hearing, and all reported to be right-handed. All participants signed informed consent forms approved by the Yale University Human Investigation Committee. The verification of participants' mother tongues was limited to self-report and to the assessment of the experimenter who tested the participants (the second author).

Experimental Manipulation

We examined the effects of speech sounds on somatosensory ERPs. The ERPs were recorded from 64 scalp sites in response to somatosensory stimulation (facial skin stretch) in three background sound conditions (pink noise, speech sounds, and nonspeech sounds) and a null condition (silent). The somatosensory stimulation was irrelevant to the participant's primary task, which was to detect occasional tone bursts by pressing a keyboard button and to fixate, without blinking, during presentation of a + symbol. We compared the ERP magnitudes observed in the different sound conditions.

The participant thus had a nominal task, the detection of tone bursts, and a main experimental manipulation, which was designed to be unattended. The experimental manipulation of skin stretch in different background sound environments occurred in parallel with the tone detection task, but there were no behavioral requirements whatsoever. ERPs are known to be sensitive to attentional manipulations (e.g., Woldorff et al., 1993). Accordingly, our goal was to examine ERPs under conditions that are unattended by the participant, so as to study somatosensory processing, not the effects of attention.

We programmed a small robotic device to apply skin stretch loads to evoke somatosensory ERPs. The details of the somatosensory stimulation procedure are described in our previous studies (Ito & Ostry, 2010; Ito et al., 2009). Briefly, the skin stretch was produced by using two small plastic tabs that were attached bilaterally with tape to the skin at the sides of the mouth. The skin stretch was applied upward. We applied a single cycle of a 3-Hz sinusoidal pattern (333 ms) with 4 N maximum force. This temporal pattern of facial skin stretch has been shown previously to alter auditory speech perception (Ito et al., 2009), and the facial skin deformations were also perceived differently in the context of speech auditory signals

as opposed to nonspeech sounds (Ito & Ostry, 2012). In the present study the somatosensory stimulation associated with facial skin deformation was applied during the background sound presentation. The interval between two sequential somatosensory stimuli was varied between 1,000 and 2,000 ms. We know of no other studies that have used facial skin deformation to elicit evoked responses in the somatosensory system.

For purposes of auditory stimulation, background sounds with occasional tone bursts (1000 Hz, 300-ms duration) were delivered binaurally through plastic tubes and earpieces (ER3A, Etymotic Research, Inc., Elk Grove Village, IL). The main experimental manipulation involved changes to the background sound used in four different conditions (speech, nonspeech, pink noise, and silence). Speech sounds were taken from a story spoken by male native speaker of American English. The story involved a wide variety of phonemic variation and was originally developed for dialect research (Comma Gets a Cure, International Dialects of English Archive; see www.dialectsarchive.com/comma-gets-a-cure). Nonspeech sounds comprised several identifiable nonverbal continuous sounds, such as traffic noise, construction noise, and fireworks. The sounds in the nonspeech samples were carefully chosen to preclude one imagining any particular human movement, including speaking and voicing. The pink noise condition was intended to broadly stimulate the auditory system, in comparison to the time-varying dynamical sounds that were used in the speech and nonspeech conditions. The four background sound conditions were presented in random order and switched every 10 somatosensory stimuli (two blocks; see Figure 1). We carried out 24 blocks of trials per background condition. Each block contained five somatosensory stimuli and one tone burst. The interval between blocks was self-paced and accordingly differed over the course of the experiment and between participants.

We also presented tone bursts (1000 Hz, 300-ms duration) that were embedded in the background sounds, and their detection was the participant's primary task. The amplitude of the tone burst was 20 dB greater than that of the background sound. In order to avoid anticipation and habituation, the tone bursts were presented in randomly selected intervals between two of the five somatosensory stimuli in each block of trials (see Figure 1). The participants were asked to respond to the tone burst by pressing a button on a keyboard. We recorded auditory-evoked potentials (AEPs) in response to the tone bursts. The somatosensory stimulus immediately following the tone burst was presented after the participant's response in order to avoid overlap of the cortical potentials. The reaction time from the onset of the tone burst to the key press was recorded to evaluate the participant's performance in four background sound conditions. The participants were also asked to gaze at a fixation mark (a + symbol) in order to eliminate unnecessary eye motion and blinking during the ERP recording. The fixation mark was removed every five somatosensory stimuli (one block, see Figure 1). The details of the stimulus sequence described above are schematized in Figure 1.

EEG Acquisition and Analysis

EEG was recorded with the BioSemi ActiveTwo (BioSemi B.V., Amsterdam, the Netherlands) system using 64 scalp electrodes (512-Hz sampling). For each participant, we recorded 120 somatosensory responses and 24 auditory responses in each of four

background sound conditions. Trials with blinks and eye movement were rejected offline on the basis of horizontal and vertical electro-oculography (over ± 150 mV). More than 85% of trials per condition were included in the analysis. EEG signals were filtered using a 0.5–50-Hz band-pass filter and rereferenced to the average across all electrodes. All responses were aligned at stimulus onset. Bias levels were adjusted using the average amplitude in the prestimulus interval (−100 to 0 ms). Finally, somato-sensory ERPs were transformed to show *scalp current density* (SCD: the second derivative of the distribution of scalp potentials; Pernier, Perrin, & Bertrand, 1988). Based on the literature on auditory-somatosensory integration, we have specifically focused on the following regions of interest (ROIs) and presumed relevant electrode locations: the orofacial motor and premotor area (left: FC3, FC5, and C3; right: FC4, FC6, and C4), the auditory area (left: T7, CP5, and TP7; right: T8, CP6, and TP8), the frontal area (Fz, F1, and F2), and the parietal area (Pz, P1, and P2; see the electrode locations in Figure 2B). These specific electrodes correspond to the standard cortical locations for the regions of interest. Electrodes were further grouped with respect to ROI by showing a similar temporal pattern of ERP for all participants. In other words, the specific combination of sets of three electrodes that were used for analysis purposes was determined empirically, based on similarity in their temporal patterns and cortical mappings. We have not analyzed areas outside of these ROIs, except for purposes of determining which electrodes should be included in the specific ROIs. SCD measures were averaged over the three electrodes that comprised each ROI to reduce variation in the response source and cortical mapping.

AEPs associated with the tone bursts were processed in the same fashion, but we did not apply the SCD transformation to AEP. The first negative peak of AEP (N1) in the frontal region (Fz) was used for this analysis. Note that the maximum amplitude of the AEP is generally observed in the mid-sagittal plane, and Fz typically shows a representative response.

We focused on the amplitude of the first peak of the somatosensory SCD in the first 200 ms following the somatosensory stimulus and on the N1 peak of the AEP. We used a 20-ms time window about the peak to compute a measure of amplitude. Amplitude measures were transformed to *z* scores to exclude individual variation in potentials. Oneway repeated-measures analysis of variance (ANOVA) was applied to the *z*-transformed amplitudes of each ROI for both auditory and somatosensory analyses.

Results

We examined whether speech sounds modify the somatosensory ERPs that are elicited in response to facial skin stretch. We employed gentle stretches of the skin lateral to the oral angle using a paradigm in which the somatosensory stimulation was unrelated to the participant's primary task. We also recorded somatosensory ERPs under nonspeech, pink noise, and silent conditions as controls.

Figure 2A shows somatosensory SCD at electrode locations over the left motor and premotor cortex as representative examples of the temporal pattern. The thick red line shows the potentials in the speech sound condition, the thin green line shows the nonspeech

background, the thin blue line shows the pink noise background, and the dashed line shows the silent condition. A negative going potential starting around 50 ms occurs in all conditions. There is a further negative depression in the speech condition, typically around 155 ms at its peak. It is seen that the peak amplitude of somatosensory event-related SCD (155 ms after stimulus onset) in the presence of speech background sounds is larger than that observed in the three other background sound conditions. Figure 2B(a) summarizes the peak amplitude that is observed in all four conditions at electrode locations above left primary motor cortex. Repeated-measures ANOVA showed a reliable difference among four conditions, $F(3, 24) = 6.30$, $p < .005$, $\omega^2 = 0.37$. Pairwise comparisons with Bonferroni correction showed that the somatosensory SCD in the presence of speech background sounds was reliably different than that observed for each of the other three background sound conditions (nonspeech: $p < .01$; pink noise: $p < .03$; silence: $p < .01$). None of the remaining pairwise comparisons showed reliable differences ($p > .90$). Thus, the somatosensory SCD at electrode locations over left premotor and primary motor cortex was enhanced in the speech condition relative to the other three background sound conditions.

We also examined the first somatosensory SCD peak at the five other ROIs (see Figure 2B). We found no statistically reliable differences in somatosensory SCD as a function of background sound at electrode locations presumed to be associated with any of these ROIs — right motor region, $F(3, 24) = 0.79$, $p > .50$, $\omega^2 = -0.024$; left auditory region, $F(3, 24) = 1.78$, $p > .10$, $\omega^2 = 0.08$; right auditory region, $F(3, 24) = 1.30$, $p > .30$, $\omega^2 = 0.032$; frontal region, $F(3, 24) = 1.54$, $p > .20$, $\omega^2 = 0.057$; parietal region, $F(3, 24) = 0.10$, $p > .90$, $\omega^2 = -0.111$. The results indicate that the enhancement in SCD was present only at electrode locations above the presumed left motor and premotor cortex.

The observed enhancement of somatosensory SCD in the context of speech sounds might reflect effects that are primarily auditory in nature, in the presence of different background sounds. In order to explore this idea, we examined the amplitude of the N1 AEP that was recorded in conjunction with the presentation of the target tone bursts. Figure 3A shows AEPs at Fz in the four background conditions. Each color corresponds to a different background sound condition as in Figure 2A. Despite the relatively small number of the ERPs that were used in the average, the AEP showed clear differences around the first negative peak (N1). There was little evidence of the typical positive potentials at around 200 ms in the speech and nonspeech background sound conditions. The mean peak amplitude of N1 at Fz is summarized in Figure 3B. Repeated-measures ANOVA showed reliable differences across background sound conditions, $F(3, 24) = 6.37$, $p < .005$, $\omega^2 = 0.37$. Pairwise comparison with Bonferroni correction showed reliable differences between the pink noise and speech conditions ($p < .0001$) and between pink noise and nonspeech conditions ($p < .05$). There was a marginal difference between the speech and silent conditions ($p = .07$). Thus, although there were differences in the pattern of N1 responses in the AEP, the pattern was not comparable to the one observed for somatosensory SCD. Indeed the current N1 pattern (and also the positive potential around 200 ms) was similar to that seen for different background sounds in mismatch negativity manipulations (Kozou et al., 2005). Although mismatch negativity responses are observed independent of the attentional level, a change of attentional level in the current auditory task might also play a

role in the modulation of the N1 component in the current data, because the N1 component of AEP is quite sensitive to attention (Woldorff et al., 1993). Specifically, background sounds in the current auditory task might differently affect the participant's attention to tone bursts.

In addition to the cortical response, we examined behavioral performance in the different background sound conditions. Reaction time to the tone bursts did not show any reliable differences across the four background conditions, $F(3, 24) = 0.94$, $p > .40$, $\omega^2 = 0.0$ (see Figure 3C). This is consistent with Kozou et al. (2005), who observed no behavioral differences despite the modulation of mismatch negativity associated with speech and nonspeech sounds as background noise.

Taken together with the behavioral data, the somatosensory and auditory responses described above suggest that background sounds affect auditory and somatosensory processing differently. We thus conclude that the modulation of auditory cortical processing observed in the context of tone bursts in different background sound conditions does not explain the enhancement of somatosensory SCD due to speech sounds.

In summary, we find that somatosensory SCD associated with facial skin stretch is enhanced by speech background sounds in comparison with nonspeech sounds, pink-noise, and no-sound conditions. This change is specifically observed at electrode locations over the left motor and premotor cortex.

Discussion

This article reports the neural correlates of the interaction between orofacial somatosensory processing and speech sound processing. Speech sounds enhance the somatosensory cortical potentials associated with the facial skin deformation that typically occurs in conjunction with speech production. The enhancement in the present study was specific to the likely area above the left motor and premotor cortex. This finding should be considered in conjunction with our previous psychophysical finding that speech sounds alter somatosensory processes associated with facial skin deformation (Ito & Ostry, 2012). The two taken together suggest that speech sounds are integrated with somatosensory inputs in the cortical processing of speech. Because auditory inputs directly affect somatosensory processing even in the absence of actual motor execution, speech sounds may possibly serve to tune the speech production system even when there is no actual execution of movement.

We found that increases in SCD (i.e., current source density or surface Laplacian: the second derivative of the distribution of scalp potentials over the scalp) occurred specifically in the probable area above motor and premotor cortex. SCD provides an estimate of the cortical surface potential that is generally a more accurate representation of the underlying current source than the raw scalp potential (Nunez & Westdorp, 1994). Therefore, an area near to motor cortex is likely the source of the increase in somatosensory–auditory interaction processing. However, given the proximity of motor and somatosensory areas and the poor spatial resolution of EEG, it is possible that the source activations may be in somatosensory rather than cortical motor areas. Evidence of sound processing in somatosensory cortex

(Beauchamp & Ro, 2008; Lütkenhöner et al., 2002) would be consistent with this possibility. More generally, although observations of speech processing in cortical motor areas may be taken as support for the involvement of the motor apparatus in speech perception, evidence for localization of speech sound processing in somatosensory cortex would fit with the idea that sound processing is referenced to the somatosensory system, rather than movement—in effect, a somatosensory theory of speech perception. A further possibility, given previous findings that the motor and premotor cortex play a role in speech perception (D'Ausilio et al., 2009; Meister et al., 2007), is that both motor and somatosensory cortical areas are involved in the neural processing of speech sounds.

Our finding also suggests that the superior temporal region, a center of somatosensory–auditory interaction in nonspeech processing (Foxe et al., 2002; Murray et al., 2005), may not play a prominent role in somatosensory cortical processing with speech sounds. Although we cannot rule out the involvement of the superior temporal region, an area closer to motor cortex seems to play a more important role in speech sound processing in the context of orofacial somatosensory function. The high-resolution EEG (i.e., 64 or more electrodes) that we used has improved the spatial resolution of scalp-recorded data. Nevertheless the source of the present somatosensory–auditory interaction is still uncertain. Further investigation is required.

The enhancement of the somatosensory SCD in the speech sound condition may be due to a change in facilitation in cortical sensorimotor areas resulting from speech sound inputs. Indeed, it has been shown that the excitability of face area motor cortex can be increased by speech sounds (Fadiga et al., 2002; Watkins et al., 2003). Second somatosensory cortex (SII) also shows activation in the presence of nonspeech auditory inputs but in the absence of somatosensory input (Beauchamp & Ro, 2008). Activity is also observed in SII in the context of combined somatosensory and auditory inputs in comparison to the activity observed in response to somatosensory inputs alone (Lütkenhöner et al., 2002). These observations suggest that hearing speech sounds may alter the excitability of the somatosensory system.

The somatosensory enhancement due to speech sounds was lateralized in the left hemisphere. Left lateralization is traditionally observed in the cortical processing of speech and language (Damasio & Geschwind, 1984), particularly in sensorimotor integration (Hickok & Poeppel, 2007). In particular, in audio-motor and visual-motor interaction in speech processing, changes in motor cortical excitation due to viewing speech movements and hearing speech sounds are lateralized to the left hemisphere (Möttönen, Järveläinen, Sams, & Hari, 2005, Watkins et al., 2003). Evidence for lateralization in orofacial motor function has been obtained during movement preparation, where greater signal change was observed in the left somatosensory cortex than in the right (Kell, Morillon, Kouneiher, & Giraud, 2011). The effect reported by Kell and colleagues was marginally stronger for speech movements than for a nonspeech orofacial motor task. Likewise, Thierry et al. (2003) demonstrated that cortical circuits in the left hemisphere are more active for hearing speech sounds than for environmental sounds, such as the nonspeech sounds used in the current study. These observations suggest that the left cortical circuits may also be predominant in somatosensory–auditory interaction in speech processing. Taken together

with our present results, neural circuits in left hemisphere, presumably in the left sensorimotor regions of the orofacial system, may play a prominent role in the interaction between auditory inputs and speech-relevant somatosensory processing.

Speech sounds specifically affect somatosensory cortical processes produced by speech-production-like patterns of facial skin stretch. However, Möttönen et al. (2005) showed that, in contrast, speech sounds did not change magneto-enchalographic somatosensory potentials due to simple lip tapping. This difference is conceivably due to the difference in stimulus patterns between the current facial skin stretching and lip tapping. Simple pressure sensations such as those due to lip tapping are not associated with any particular articulatory motion in speech. On the other hand, deformation of the facial skin can provide kinesthetic input in conjunction with speech articulatory motion, because the facial skin stretch stimulation affects the control process of speech articulatory motion (Ito & Gomi, 2007) and motor adaption (Ito & Ostry, 2010). Like other sensory integration processes, somatosensory–auditory interaction in speech processing likely follows the general rule of multisensory integration that the integration is tuned to the specific types and pattern of sensory inputs that are associated with each other in the task (Stein & Meredith, 1993). Hence, somatosensory inputs that are similar to those experienced in speech production can interact effectively with speech sound processing. The interaction appears to be quite specific. Nonspeechlike patterns of facial skin stretch (9-Hz stretch, which is beyond the normal speaking rate) do not affect the perception of speech sounds, whereas speechlike patterns of skin stretch (3-Hz stretch, which is within the normal speech range) do affect speech perception (Ito et al., 2009). Consequently, speech sounds specifically affect the somatosensory processing that occurs in association with speech-production-like patterns of facial skin stretch. The somatosensory–auditory interaction in speech processing is narrowly tuned to the specific pattern and type of sensory signals, and this specificity is consistent with a tight cortical linkage between speech production and perception.

Acknowledgments

References

Arima T, Yanagi Y, Niddam DM, Ohata N, Arendt-Nielsen L, Minagi S, Svensson P. Corticomotor plasticity induced by tongue-task training in humans: A longitudinal fMRI study. Experimental Brain Research. 2011; 212:199–212. [PubMed: 21590261]

Beauchamp MS, Ro T. Neural substrates of sound-touch synesthesia after a thalamic lesion. Journal of Neuroscience. 2008; 28:13696–13702. [PubMed: 19074042]

Beauchamp MS, Yasar NE, Frye RE, Ro T. Touch, sound and vision in human superior temporal sulcus. NeuroImage. 2008; 41:1011–1020. [PubMed: 18440831]

Chang SE, Kenney MK, Loucks TM, Poletto CJ, Ludlow CL. Common neural substrates support speech and nonspeech vocal tract gestures. NeuroImage. 2009; 47:314–325. [PubMed: 19327400]

Damasio AR, Geschwind N. The neural basis of language. Annual Review of Neuroscience. 1984; 7:127–147.

D'Ausilio A, Pulvermüller F, Salmas P, Bufalari I, Begliomini C, Fadiga L. The motor somatotopy of speech perception. Current Biology. 2009; 19:381–385. [PubMed: 19217297]

Fadiga L, Craighero L, Buccino G, Rizzolatti G. Speech listening specifically modulates the excitability of tongue muscles: A TMS study. European Journal of Neuroscience. 2002; 15:399–402. [PubMed: 11849307]

Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Murray MM. Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study. Journal of Neurophysiology. 2002; 88:540–543. [PubMed: 12091578]

Fu KMG, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Schroeder CE. Auditory cortical neurons respond to somatosensory stimulation. Journal of Neuroscience. 2003; 23:7510–7515. [PubMed: 12930789]

Ghosh SS, Tourville JA, Guenther FH. A neuro-imaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. Journal of Speech, Language, and Hearing Research. 2008; 51:1183–1202.

Hickok G, Poeppel D. The cortical organization of speech processing. Nature Reviews Neuroscience. 2007; 8:393–402.

Ito T, Gomi H. Cutaneous mechanoreceptors contribute to the generation of a cortical reflex in speech. NeuroReport. 2007; 18:907–910. [PubMed: 17515799]

Ito T, Ostry DJ. Somatosensory contribution to motor learning due to facial skin deformation. Journal of Neurophysiology. 2010; 104:1230–1238. [PubMed: 20592121]

Ito T, Ostry DJ. Speech sounds alter facial skin sensation. Journal of Neurophysiology. 2012; 107:442–447. [PubMed: 22013241]

Ito T, Tiede M, Ostry DJ. Somatosensory function in speech perception. Proceedings of the National Academy of Sciences, USA. 2009; 106:1245–1248.

Kayser C, Petkov CI, Augath M, Logothetis NK. Integration of touch and sound in auditory cortex. Neuron. 2005; 48:373–384. [PubMed: 16242415]

Kell CA, Morillon B, Kouneiher F, Giraud A. Lateralization of speech production starts in sensory cortices: A possible sensory origin of cerebral left dominance for speech. Cerebral Cortex. 2011; 21:932–937. [PubMed: 20833698]

Kozou H, Kujala T, Shtyrov Y, Toppila E, Starck J, Alku P, Näätänen R. The effect of different noise types on the speech and non-speech elicited mismatch negativity. Hearing Research. 2005; 199:31–39. [PubMed: 15574298]

Lakatos P, Chen C, O'Connell MN, Mills A, Schroeder CE. Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron. 2007; 53:279–292. [PubMed: 17224408]

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. Psychological Review. 1967; 74:431–461. [PubMed: 4170865]

Lütkenhöner B, Lammertmann C, Simões C, Hari R. Magnetoencephalographic correlates of audiotactile interaction. NeuroImage. 2002; 15:509–522. [PubMed: 11848694]

Malandraki GA, Sutton BP, Perlman AL, Karampinos DC, Conway C. Neural activation of swallowing and swallowing-related tasks in healthy young adults: An attempt to separate the components of deglutition. Human Brain Mapping. 2009; 30:3209–3226. [PubMed: 19247994]

Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M. The essential role of premotor cortex in speech perception. Current Biology. 2007; 17:1692–1696. [PubMed: 17900904]

Möttönen R, Calvert GA, Jääskeläinen IP, Matthews PM, Thesen T, Tuomainen J, Sams M. Perceiving identical sounds as speech or nonspeech modulates activity in the left posterior superior temporal sulcus. NeuroImage. 2006; 30:563–569. [PubMed: 16275021]

Möttönen R, Järveläinen J, Sams M, Hari R. Viewing speech modulates activity in the left SI mouth cortex. Neuro-Image. 2005; 24:731–737. [PubMed: 15652308]

Möttönen R, Watkins KE. Motor representations of articulators contribute to categorical perception of speech sounds. Journal of Neuroscience. 2009; 29:9819–9825. [PubMed: 19657034]

Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Foxe JJ. Grabbing your ear: Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. Cerebral Cortex. 2005; 15:963–974. [PubMed: 15537674]

Nunez PL, Westdorp AF. The surface Laplacian, high resolution EEG and controversies. Brain Topography. 1994; 6:221–226. [PubMed: 8204409]

Pernier J, Perrin F, Bertrand O. Scalp current density fields: Concept and properties. Electroencephalography and Clinical Neurophysiology. 1988; 69:385–389. [PubMed: 2450736]

Schürmann M, Caetano G, Hlushchuk Y, Jousmäki V, Hari R. Touch activates human auditory cortex. NeuroImage. 2006; 30:1325–1331. [PubMed: 16488157]

Simonyan K, Ostuni J, Ludlow CL, Horwitz B. Functional but not structural networks of the human laryngeal motor cortex show left hemispheric lateralization during syllable but not breathing production. Journal of Neuroscience. 2009; 29:14912–14923. [PubMed: 19940187]

Stein, BE.; Meredith, MA. The merging of the senses. MIT Press; Cambridge, MA: 1993.

Thierry G, Giraud AL, Price C. Hemispheric dissociation in access to the human semantic system. Neuron. 2003; 38:499–506. [PubMed: 12741995]

Watkins KE, Strafella AP, Paus T. Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia. 2003; 41:989–994. [PubMed: 12667534]

Wilson SM, Saygin AP, Sereno MI, Iacoboni M. Listening to speech activates motor areas involved in speech production. Nature Neuroscience. 2004; 7:701–702.

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE. Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proceedings of the National Academy of Sciences, USA. 1993; 90:8722–8726.
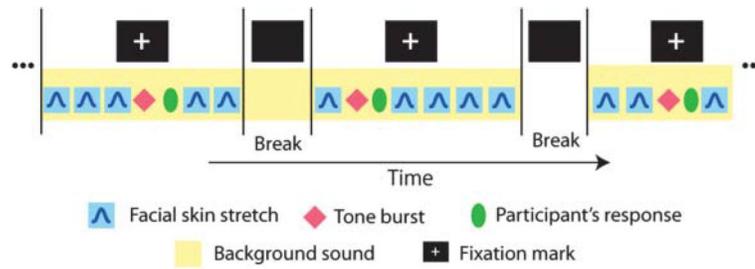
Figure 1.
Schematic representation of the experimental sequence. Somatosensory stimuli and tone bursts were presented in conjunction with background sounds. Blocks of trials involved five somatosensory stimuli and one tone burst each. The interval between successive stimuli was from 1,000 to 2,000 ms. The background sound changed every 10 somatosensory stimuli.
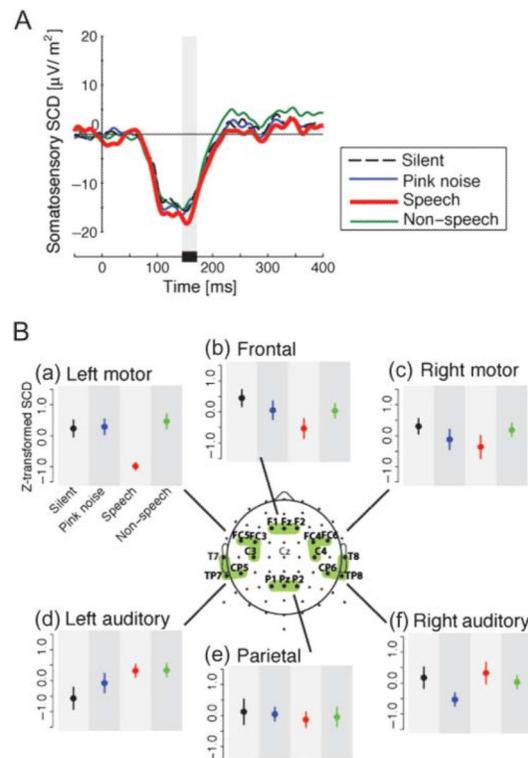
Figure 2.
Somatosensory scalp current density (SCD) due to facial skin stretch in four background sound conditions (speech, non-speech, pink noise, and silence). Panel A: Temporal pattern of somato-sensory SCD in the area above left motor and premotor cortex. Each color corresponds to a different background sound condition. The gray bar shows the time window for the calculation of the SCD amplitude measure. Panel B: Differences in *z*-score magnitudes associated with the first peak of the somatosensory SCD in each of six regions of interest. Error bars give standard errors across participants. Each color corresponds to different background sound conditions as in Panel A.
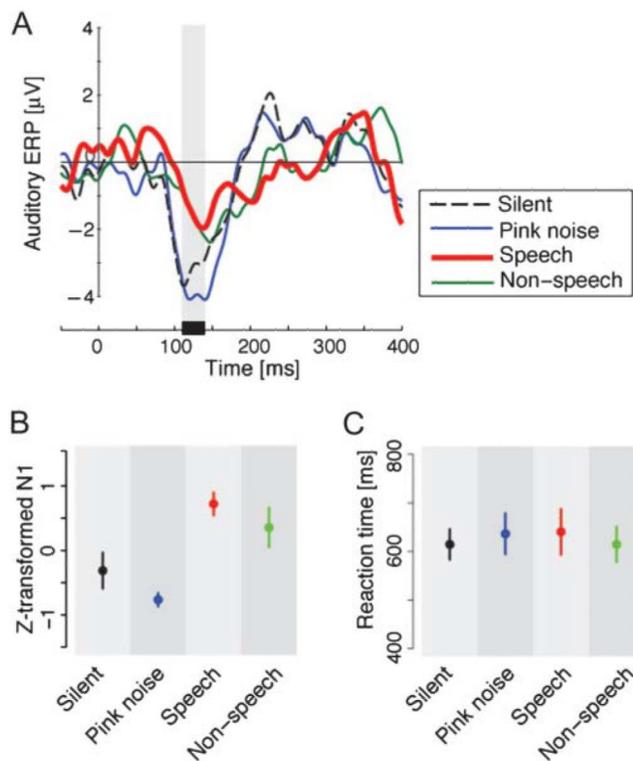
Figure 3.
Panel A: Temporal pattern of auditory-evoked potentials (AEPs) at Fz due to tone-burst presentation. The shaded area shows the time window for the AEP amplitude calculation. Panel B: Peak amplitude of N1 in the AEP. Error bars give standard errors across participants. Panel C: Differences in reaction time to tone bursts. Error bars give standard errors.