Modulation of Auditory and Somatosensory Processing During the Planning of Speech Movements

Ludo Max[1,2], Jennifer C. Daniels[3], Karla M. Curet[3], & Kristine L. Cronin[3]

[1] University of Washington, Seattle, WA; [2] Haskins Laboratories, New Haven, CT;
[3] University of Connecticut, Storrs, CT
E-mail: ludomax@u.washington.edu

Abstract

Sensorimotor control has been shown to involve an attenuation of self-generated afferent inputs. Given that—during movements—the CNS also attenuates its response to externally-generated stimuli, the phenomenon cannot be fully explained by current suggestions that attenuation occurs when a match is detected between predicted and actual sensory consequences. In addition, various aspects of the nature and time course of such sensory modulation in the context of movement remain unknown. We therefore investigated sensory modulation using a paradigm in which externally-generated auditory or orofacial vibrotactile stimuli are delivered prior to the onset of voluntary movements. Recordings of auditory and somatosensory evoked potentials were used to examine the afferent pathways' response to stimulation during the planning phase preceding speech and finger movements. Results show that (a) the CNS modulates its response to externally-generated stimuli even before movement onset, and (b) this modulation is specific to sensory systems that are relevant for the motor task.

1 Introduction

Results from several studies indicate that the central nervous system (CNS) attenuates its response to self-generated versus externally-generated afferent inputs. For example, perceptions of ticklishness, applied force, and cold temperature are weaker when self-applied than when applied by an external agent [1,2,3,4].

In the auditory system of both human and non-human primates, cortical responses to self-generated vocalizations are also reduced in comparison with responses to played-back vocalizations [5,6,7,8,9]. Based on these findings, it has been suggested that sensory attenuation occurs when the CNS detects a match between predicted and actual sensory consequences [6].

It appears, however, that there may be at least one additional component to sensory attenuation given that the CNS also modulates its response to externally-generated stimuli that are presented during voluntary movements. For example, auditory cortex responses to tones presented during speaking are attenuated as compared with tones presented in the absence of speaking [6,8,9,10]. Other results suggest that the amplitude of auditory and somatosensory evoked potentials is also attenuated during finger movements [11,12,13,14].

Recent work has explored responses to external stimuli in the absence of movement or prior to its onset. Results suggest that sensory attenuation occurs even in these circumstances [11,15,16]. If confirmed, these findings demonstrate that sensory attenuation cannot be explained solely by the proposal that the CNS detects a match between predicted and actual sensory consequences.

Here, we report a series of experiments examining the auditory system's response to external stimuli delivered during the planning phase prior to the onset of speech movements.

2 Experiment I

The purpose of this pilot experiment was to test the overall paradigm, and to further optimize it by estimating the appropriate timing for the delivery of auditory stimuli during speech planning.

2.1 Subjects

Subjects were 10 healthy women, 23-28 years of age, with hearing thresholds at 20 dB HL or better for the octave frequencies from 250 Hz to 8 kHz.

2.2 Procedure

Condition 1 was a delayed response *speaking* task: each trial started with the visual presentation of a monosyllabic word in white characters on a black background. The color of the word changed to green 600 ms later, and this change in color served as the "go" signal to speak. Condition 2 was a *silent reading* task for which different words were presented in the same manner as in Condition 1. Condition 3 consisted of *seeing* "++++" symbols presented in the same manner. Condition 4 was an eyes-closed *rest* condition.

For each condition, tones (1 kHz, 50 ms, 10 ms rise/fall times, 75 dB SPL) were presented through insert earphones (ER-3A) 300, 400, or 500 ms after appearance of the white characters on the monitor (Figure 1). Each of these three sub-conditions was completed as a block of 90 trials.

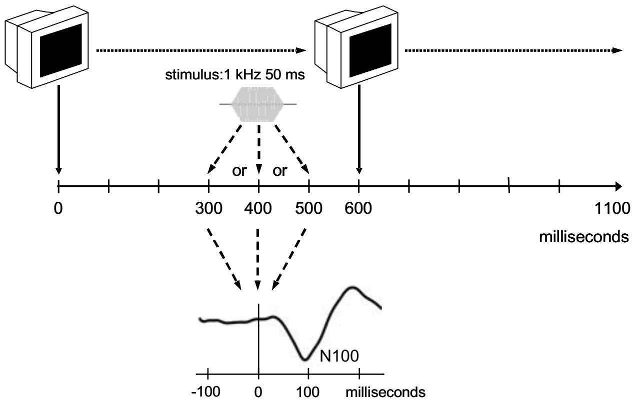

Figure 1: *Schematic representation of the time course of an individual trial in Experiment I.*

2.3 Data acquisition and processing

Two channels of auditory evoked potentials (AEPs) were recorded with a custom EEG system (NeuroScan Quik-Cap with Ag/AgCl electrodes, Grass Model 15 Neurodata Amplifier System; Measurement Computing A/D board): Cz-M1 (vertex re. left mastoid) and Cz-M2 (vertex re. right mastoid). Electro-oculography (EOG), surface electromyography (sEMG), and speech acoustics were recorded to reject trials with eye movements/blinks, muscle activity, and/or speech during the delay interval. sEMG electrodes were placed over the masseter (MASS), orbicularis oris inferior (OOI), and anterior belly of the digastric (ABD). Signals were digitized with a sampling rate of 4 KHz after analog band-pass filtering from 0.1 to 100 Hz for EEG, 10 to 300 Hz for sEMG, and 0.1 to 30 Hz for EOG. Offline, data were epoched from 100 ms before to 300 ms after stimulus onset.

2.4 Results

Results showed contamination (i.e., absence of stable baselines) of the AEP by visual evoked potentials when tones were presented only 300 ms after appearance of the word. For the 400 ms sub-condition, statistically significant ($p < .05$) differences in N100 amplitude were found for both channels. However, the conditions' grand average waveforms (highly similar to those shown in panel A of Figure 2, although the figure illustrates data from Experiment II) suggested that non-auditory components related to the task itself (speaking, reading, seeing) contributed unequal amounts of DC shift to the signals from different conditions, thereby preventing unambiguous interpretation of any differences in auditory processing.

3 Experiment II

Experiment II was designed to isolate the auditory component from simultaneous cognitive, linguistic, and visual components by subtracting time-locked EEG activity during no-tone trials from time-locked EEG activity during tone trials.

3.1 Subjects

Ten healthy women, ranging from 20 to 28 years of age, participated. All subjects passed a hearing screening as described above.

3.2 Procedure

The protocol from Experiment I was modified to include only the 400 ms sub-condition, and auditory stimuli were presented during only 1/3 of 270 trials per condition (speaking, reading, seeing).

3.3 Data acquisition and processing

Data acquisition and processing followed the steps described for Experiment I. After epoching, each subject's EEG averaged over 90 (minus rejected) no-tone trials was subtracted from the EEG averaged over 90 (minus rejected) tone trials. This subtraction procedure is illustrated in the three panels on the left side of Figure 2.

3.4 Results

As shown in Figure 2, after the subtraction, N100 amplitude was smaller for speaking vs. both seeing (t-tests: Cz-M1 p = .05; Cz-M2 p = .05) and silent reading (Cz-M1 p = .03; Cz-M2 p = .01).

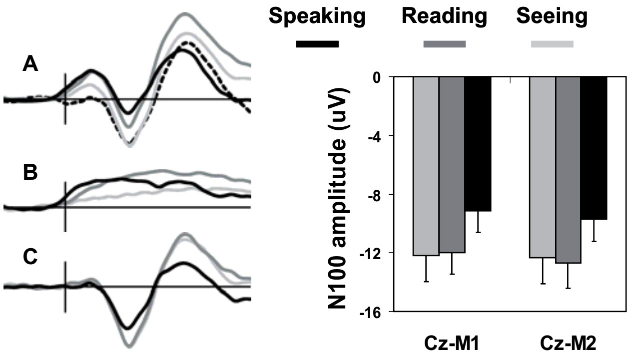

Figure 2: *Left: grand average Cz-M1 waveforms from Experiment II. A: tone trials, B: no-tone trials, C: tone trials after subtraction of no-tone trials. Right: group means for N100 amplitude after subtraction.*

4 Experiment III

Trials in Experiment III included auditory stimuli or vibrotactile stimuli to the right side of the upper lip. Stimuli were applied during the planning of speech or thumb movements.

4.1 Subjects

Ten healthy adults (5 male) with an age range from 20 to 34 years participated. They all passed the hearing screening described above.

4.2 Procedure

With their head immobilized, subjects completed four conditions: speaking, thumb movements (clicking reversed computer mice), and two control tasks with the same visual displays but no movements. The speaking task was as described above. In the finger-movement task, instructions for two clicks (e.g., "Le Ri" for "left then right") first appeared in white characters and changed to green 600 ms later to represent the "go" signal.

Each task consisted of 270 trials presented in 3 blocks of 90 trials, with each block containing—in random order—30 trials with auditory stimulation, 30 with vibrotactile stimulation, and 30 with no stimulation. Auditory stimuli were identical to those in Experiments I and II. Vibrotactile stimuli (60 Hz, 50 ms, 10 ms rise/fall times, 2 mm) were applied by a plastic probe (10-mm diameter contact surface) driven by a mini shaker (Bruel & Kjaer model 4810 with amplifier model 2718).

4.3 Data acquisition and processing

AEPs and somatosensory evoked potentials (SEPs) were recorded using Ag/AgCl electrodes (10 mm, Grass Technologies) at positions Cz, C3, C4, Fz, F3, F4, M2, and M1 (recording reference). EOG, sEMG, and acoustics were used to reject artifacts. The microphone signal was routed to two in-series analog high-pass filters (100 Hz cut off, Wavtek model 862) to eliminate feedback from operation of the mini-shaker in the earphones.

Continuous EEG was analog band-pass filtered from 0.1-300 Hz; sEMG and EOG were filtered as described above. Offline, EEG was re-referenced to mathematically linked mastoids and digitally low-pass filtered at 100 Hz prior to epoching from -100 ms to 250 ms. Each subject's EEG averaged over 90 (minus rejected) no-stimulus trials was subtracted from the EEG averaged over 90 (minus rejected) tone or vibration trials.

4.4 Results

Results are shown in Figure 3. After subtraction, decreases in amplitude of the AEP P50 (t-test: Cz p = .05) and SEP N50 (Fz p = .02) were found when auditory or vibrotactile stimuli were delivered during speech planning vs. silent reading. No statistically significant decreases in AEP or SEP amplitudes occurred for finger movement planning.

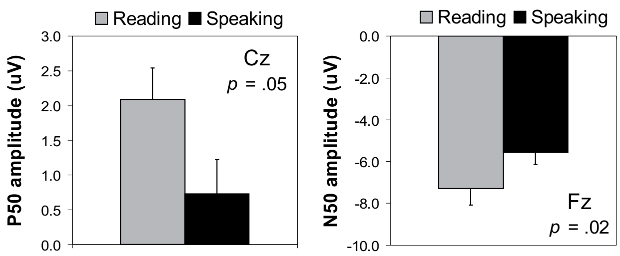

Figure 3: *Group means for amplitude of AEP P50 (left panel) and SEP N50 (right panel).*

5 Experiment IV

Experiment IV investigated whether the different AEP results in Experiments II and III (reduction in N100 vs. P50, respectively) may be related to differences in stimulus predictability (only auditory stimuli vs. randomized auditory and vibrotactile stimuli, respectively). Sensory modulation was not directly examined, however, as no-speaking trials were not included in order to limit the duration of the test sessions (990 speaking trials).

5.1 Subjects

Seven healthy adults (3 male; age 18-24 years) participated. Six subjects passed the aforementioned hearing screening at all frequencies, but one had higher thresholds at 8 KHz (left 25 dB, right 30 dB).

5.2 Procedure

Subjects performed a delayed response speaking task as described above. Auditory stimuli, identical to those described above, were presented 400 ms after appearance of the word. Presentation frequency and predictability of the stimuli were manipulated across conditions. In Condition 1 (1 block of 90 trials), tones were presented during each trial. In Condition 2 (2 blocks of 90 trials each), tones were presented during half of the trials, with tone and no-tone trials occurring in an alternating manner. In Condition 3 (2 blocks of 90 trials), tones were also presented during half of the trials, but the order of the trials was randomized. In Condition 4 (5 blocks of 90 trials), tones were presented in one fifth of the trials, with the tone trials again presented randomly among the no-tone trials. In Condition 5 (1 block of 90 trials), no tones were presented.

5.3 Data Processing and Acquisition

Thirty-two EEG channels, EOG, and sEMG were digitized at 4 KHz using Ag/AgCl active electrodes connected to a DC amplifier (BioSemi Active-Two system). EEG was offline low-pass filtered with a cut-off of 100 Hz and epoched from -50 to 300 ms.

5.4 Results

No statistically significant differences in P50 or N100 amplitude were found in any pair-wise comparisons of the five conditions.

6 Discussion

Combined, findings indicate that the sensory processing of externally-generated auditory or orofacial somatosensory stimuli is already modulated during the planning phase preceding speech movements. Modulation was observed more than 200 ms before movement onset. Moreover, given that no auditory or orofacial somatosensory modulation occurred during the planning of thumb movements, findings suggest that the response to externally-generated stimuli is modulated only for sensory modalities and receptive fields that are task-relevant. Overall, these data are consistent with a theoretical model of sensorimotor control in which an optimal feedback controller [17] implements early, direct adjustments in sensory gains on the basis of an efference copy of planned motor commands and a forward internal model of the effector systems.

7 References

[1] Blakemore S, Wolpert D, Frith C (1998) Central cancellation of self-produced tickle sensation. Nature Neurosci 1:635-640

[2] Weiskrantz L, Elliot J, Darlington C (1971) Preliminary observations on tickling oneself. Nature 230:598-599

[3] Shergill S, Bays P, Frith C, Wolpert D (2003) Two eyes for an eye: the neuroscience of force escalation. Science 301:187

[4] Van Doorn G, Richardson B, Wuillemin D, Symmons M (2005) Modifications of magnitude estimations in thermotactile perception during self-generated and externally generated movements. Perception 34:231-235

[5] Curio G, Neuloh G, Numminen J, Jousmäki V, Hari R (2000) Speaking modifies voice-evoked activity in human auditory cortex. Hum Brain Mapp 9:183-191

[6] Houde J, Nagarajan S, Sekihara K, Merzenich M (2002) Modulation of the auditory cortex during speech: an MEG study. J Cognitive Neuroscience 14:1125-1138

[7] Müller-Preuss P, Ploog D (1981) Inhibition of auditory cortical neurons during phonation. Brain Res 215:61-76

[8] Eliades S, Wang X (2003) Sensory-motor interaction in primate auditory cortex during self-initiated vocalization. J Neurophysiol 89:2194-2207

[9] Eliades S, Wang X (2005) Dynamics of auditory-vocal interaction in monkey auditory cortex. Cereb Cortex 15: 1510-1523

[10] Numminen J, Salmelin R, Hari R (1999) Subject's own speech reduces reactivity of the human auditory cortex. Neurosci Lett 265:119-122

[11] Cohen L, Starr A (1987) Localization, timing, and specificity of gating of somatosensory evoked potentials during active movement in man. Brain 110:451-467

[12] Papakostopoulos D, Cooper R, Crow H (1975) Inhibition of cortical evoked potentials and sensation by self-initiated movement in man. Nature 258:321-323

[13] Rushton D, Rothwell J, Craggs M (1981) Gating of somatosensory evoked potentials during different kinds of movement in man. Brain 104:465-491

[14] Tapia M, Cohen L, Starr A (1987) Attenuation of auditory-evoked potentials during voluntary movement in man. Audiol 26:369-373

[15] Voss M, Ingram J, Haggard P, Wolpert D (2006) Sensorimotor attenuation by central motor command signals in the absence of movement. Nature Neurosci 9:26-27

[16] Shimazu H, Kaji R, Murase N, Kohara N, Ikeda A, Shibasaki H, Kimura J, Rothwell JC (1999) Pre-movement gating of short-latency somatosensory evoked potentials. NeuroReport 10:2457-2460

[17] Scott S (2008) Inconvenient truths about neural processing in primary motor cortex. J Physiol 586:1217-1224