Research report

Multisensory auditory–somatosensory interactions in early cortical processing revealed by high-density electrical mapping

John J. Foxe\textsuperscript{a,b,c,*}, Istvan A. Morocz\textsuperscript{d}, Micah M. Murray\textsuperscript{a,b}, Beth A. Higgins\textsuperscript{a}, Daniel C. Javitt\textsuperscript{a,b,c,e}, Charles E. Schroeder\textsuperscript{a,b}

\textsuperscript{a}Cognitive Neurophysiology Laboratory, Nathan Kline Institute for Psychiatric Research, Program in Cognitive Neuroscience and Schizophrenia, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA
\textsuperscript{b}Department of Neuroscience, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York, NY 10461, USA
\textsuperscript{c}Department of Psychiatry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York, NY 10461, USA
\textsuperscript{d}Department of Neurology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York, NY 10461, USA
\textsuperscript{e}Department of Psychiatry, New York University School of Medicine, 550 1st Avenue, New York, NY 10016, USA

Accepted 28 March 2000

Abstract

We investigated the time-course and scalp topography of multisensory interactions between simultaneous auditory and somatosensory stimulation in humans. Event-related potentials (ERPs) were recorded from 64 scalp electrodes while subjects were presented with auditory-alone stimulation (1000-Hz tones), somatosensory-alone stimulation (median nerve electrical pulses), and simultaneous auditory–somatosensory (AS) combined stimulation. Interaction effects were assessed by comparing the responses to combined stimulation with the algebraic sum of responses to the constituent auditory and somatosensory stimuli when they were presented alone. Spatiotemporal analysis of ERPs and scalp current density (SCD) topographies revealed AS interaction over the central/postcentral scalp which onset at approximately 50 ms post-stimulus presentation. Both the topography and timing of these interactions are consistent with multisensory integration early in the cortical processing hierarchy, in brain regions traditionally held to be unisensory. © 2000 Elsevier Science B.V.

Theme: Sensory systems

Keywords: Multisensory integration; Crossmodal integration; Event-related potential; High-density mapping; Somatosensory cortex; Auditory cortex; ERP; AEP; SEP; Human; Binding

1. Introduction

Unified perception of stimuli that are seen, felt and heard requires that regions of the brain receive convergent inputs from these different sensory modalities and that these convergent inputs are somehow integrated or ‘bound’. Such multisensory integration constitutes a fundamental component of cognition and behavior [26,25]. The means by which information from the different sensory modalities is bound together in the brain is therefore of critical interest. Detailing the anatomic organization of the brain networks that subserve this integration will be crucial to our clarification of this aspect of the ‘binding’ problem. The temporal pattern of integration across the areas comprising these networks is key to understanding the nature of the interactions amongst these areas.

The objective of this study was to examine the time-course and scalp topography of multisensory integration processes between the senses of audition and somatosensation in humans. A number of recent studies have demonstrated the behavioral relevance of auditory–somatosensory (AS) interactions [8,15,24]. For instance, Jousmaki and Hari [8] have shown that artificially altering the rubbing sounds that subjects hear when they rub their hands together can drastically alter the tactile sensations...
that subjects report. In fact, by enhancing the high-frequency component (>2 kHz) of the rubbing sound, a majority of their subjects experienced the rather extraordinary sensation of having a leaf of parchment paper interposed between their rubbing hands. This effect underscores AS multisensory integration as an important adjunct to tactile judgements of texture. Also, recent evidence has shown that there are AS interactions during both focused and divided spatial attention tasks [15,24].

Intracranial investigations in animals have identified candidate neural structures for such auditory–somatosensory (AS) integration. AS neurons have been found in the neostriatum and globus pallidus of the rat [1], superior colliculus [19,32,33] and anterior ectosylvian cortex of the cat [29,28,31], and the posterior parietal cortex [7], superior temporal sulcus [2,6] and temporo-parietal cortex [11] of the monkey. In cats, multisensory AS neurons have also been described in the inferior colliculus, a subcortical relay to primary auditory cortices [27], although it is not clear to what extent these findings would apply to humans. Recent intracranial investigations in awake monkeys at this laboratory have shown AS co-representation in the posterior ‘belt’ region of macaque auditory cortex on the superior temporal plane [13]. Collectively, these findings suggest that AS integration may be a bottom-up (feedforward) process occurring during initial sensory transmission and in early sensory cortices, which are usually assumed to be unisensory. Other potential areas for AS integration are lateral and ventral intraparietal sulcal areas (LIP [14,21]and VIP [3]), both of which show audio–visual and somato–visual responses, but have not been directly investigated for AS function.

Since intracranial investigations in the superior temporal plane of awake macaques have shown early feedforward AS co-representation [13], we predicted integration effects during the timeframe of the earliest components of the human somatosensory and auditory-evoked responses (SEPs and AEPs). High-density event-related potential (ERP) recordings revealed multisensory AS interactions over the central/postcentral scalp region, contralateral to the side of multisensory stimulation that onset just 50 ms post-stimulus presentation.

2. Materials and methods

2.1. Subjects

Eight (one female), neurologically normal, paid volunteers, aged 22–35 (mean=26.9±4.2) participated. All subjects provided written informed consent in accordance with the Declaration of Helsinki, and the procedures were approved by the Institutional Review Board of the Nathan Kline Institute for Psychiatric Research. All subjects reported having normal hearing, and were right-handed (except for one) as assessed by the Edinburgh handedness inventory.

2.2. Stimulation

Subjects were presented with three types of stimulation during the experiment: (1) auditory-alone stimuli (1000-Hz tones, 85 dB SPL, 60 ms duration, 7.5 ms rise/fall) delivered through Sony MDR-V400 headphones; (2) median nerve electrical stimulation; and (3) combined auditory and median nerve stimulation where the above stimuli were delivered simultaneously with synchronized onsets. In the current experiment, only responses to stimuli presented to the left ear and the left median nerve are included. Stimuli of the three types were presented in blocks of 200 stimuli with order of block presentation randomly varied. Inter-stimulus interval was randomly varied from 700 to 900 ms for blocks of all three stimulus types.

For median nerve stimulation, electrical stimulation consisted of 400-μs, constant-current, square-wave pulses applied with bipolar electrodes to the skin of the forearm over the median nerve. The optimal stimulation point was located on the ventral forearm of each subject on an individual basis, typically 2–3 cm proximal to the wrist. Current settings (1–5 mA) were first adjusted to produce isolated supra-threshold activation of the appropriate distal musculature, and were then attenuated until the overt muscle activation was no longer visible, but subjects still reported a strong pulse stimulus.

Subjects took breaks as necessary between blocks in order to minimize fatigue. In the interest of minimizing eye-movements, subjects were instructed to direct their gaze at and concentrate upon pictures (still photographs) which were placed on a screen 150 cm in front of them and to refrain from making eye-movements or blinks as much as possible. Photographs were changed between each stimulation block. Subjects were instructed to ignore the auditory and somatosensory stimuli and, upon debriefing at the end of the experiment, all subjects reported that they had quickly become habituated to the stimuli so that they became essentially oblivious to the stimuli under study.

2.3. Measurements and analyses

High-density EEG recordings were acquired from 64 scalp electrodes (see Fig. 1) referenced to an electrode placed on the tip of the nose (filtered online with a pass-band of 0.05–100 Hz; digitally sampled at 1000 Hz; impedances <3 kΩ). Trials with blinks and large eye movements were rejected offline on the basis of horizontal (HEOG) and vertical (VEOG) electro-oculogram. An artifact rejection criterion of ±60 μV was used to reject trials with excessive EMG or other noise transients. Accepted trials were epoched from −50 ms pre-stimulus to 300 ms post-stimulus. The average number of accepted
sweeps per condition was \(722 \pm 221\) with the lowest number of trials for any condition being 429 and the highest, 1157. The baseline was defined as the mean voltage over the 10–20-ms post-stimulus latency window (to avoid incorporating the stimulus artifact from electrical stimulation into baseline calculation). Averages were made for the three stimulation types used: (1) auditory-alone; (2) somatosensory-alone; and (3) ‘simultaneous’ auditory–somatosensory stimulation. The averages for auditory-alone and somatosensory-alone were then ‘summed’ and all comparisons were made between the ‘simultaneous’ and ‘summed’ responses. Multisensory integration was defined as any significant difference between these responses, on the premise that as long as responses to the constituent auditory and somatosensory components are processed by unisensory neural structures, their responses will sum linearly and be equivalent to the summed responses of the stimulus-alone conditions. When this linearity no longer holds, it is reasonably concluded that a multisensory area has been activated and integration of the somatosenory and auditory inputs is occurring. It should be noted that this methodology will not be sensitive to areas of purely multisensory convergence wherein responses to two sensory modalities might occur but would sum linearly.

Scalp current densities (SCD) (second spatial derivative of the potential) were computed from the spherical spline interpolation of the surface voltage recordings made from all 64 channels. This technique eliminates the influence of the reference electrode and emphasizes local contributions to the surface map, providing for better visualization of approximate locations of intracranial generators.
3. Results

Inspection of the group-averaged voltage waveforms (see Fig. 1) over right central scalp (the expected scalp projection of both auditory and somatosensory cortices for left-sided stimulation) revealed a series of ERP components (N20, P30, P50, and N65) in the early time-period following auditory–somatosensory bi-sensory stimulus presentation (Figs. 1 and 2A).

A robust difference was noted between the responses to simultaneous AS stimulation and the summed responses from the separately presented auditory and somatosensory constituents. This interaction effect appeared to onset in the 40–50 ms time-range during the peak to falling phase of the P50 positivity and was sustained through the N65 component. The dashed red difference wave (summed minus simultaneous) in Fig. 2A illustrates this interaction effect.

To verify the presence of interactions between auditory and somatosensory stimulation, ERPs elicited by simultaneous presentations were statistically compared to the ‘summed’ ERP derived from the separate presentation conditions, using two-way repeated measures analysis of variance (ANOVA), with factors of stimulus type (summed and simultaneous) and electrode (C4 and C6). The α level criterion was set at $P<0.05$. Amplitude measures were taken for four successive latency windows (40–49, 50–59, 60–69 and 70–79 ms) by deriving an area measure between the ERP waveforms and the 0 μV baseline. Significant effects were found for the latter three periods (50–59 ms, $F(1,7)=10.413, P=0.015$; 60–69 ms, $F(1,7)=11.417, P=0.012$; 70–79 ms, $F(1,7)=7.085, P=0.032$), but not for the earliest window (40–49 ms, $F(1,7)=0.822, P=0.395$).

Scalp current density maps (Fig. 2B) of the difference waves at 65 ms show the topographic distribution characteristic of the early phase of the interaction effect. A lateral negative focus over right central/post-central scalp can be seen with a concurrent positive focus over the vertex. The polarity of these foci is arbitrary since it is based on the direction of the subtraction from which the difference maps are derived. This scalp distribution is consistent with a contribution from neural activity within the hand-representation of the somatosensory cortices, in the region of the postcentral gyrus. By 80 ms, the positive focus over the vertex has diminished in amplitude and an additional positive focus has emerged over the right inferior occipitotemporal region while the topographic distribution of the right lateral negativity has shifted anteriorly (Fig. 2C). This later positive/negative distribution reverses polarity in the region of the superior temporal plane, consistent with a contribution from neural activity within the region of posterior auditory cortices.

For a comprehensive treatment of the neural basis of event-related potentials and the resolving power of high-density recordings, readers are referred to Vaughan and Arezzo [30] and Simpson and co-workers [22,23].

4. Discussion

This study revealed the presence of early interactions between the senses of somatosensation and audition in human subjects. Interaction effects were manifest by just 50 ms post-stimulus presentation. Scalp current density (SCD) topographic analysis revealed a central/post-central scalp distribution for the early phase of the interaction effect, contralateral to the side of stimulation. This distribution was largely consistent with interaction effects in somatosensory cortices in the region of the postcentral gyrus. By 70–80 ms, a shift in scalp topography suggested a contribution from a second generator. This second topographic distribution was consistent with additional generators in auditory cortices of the posterior superior temporal plane. In sum, these findings predict a set of generators involving both somatosensory and auditory cortices in AS integration. Since these two topographic distributions overlap considerably in the temporal domain and onset latency is not assessed here, we are not inclined to make any strong claims about the temporal sequence of activations at this juncture.

Concerning the first prediction, multisensory AS function has not been previously associated with early somatosensory cortices, but evidence for multisensory visual–somatosensory interactions have been found. Intracranial recordings in somatosensory cortices [34] have shown visual responsiveness during a short-term visuo–somatosensory delayed match-to-sample task. In this study, visual responses were found in the anterior parietal somatosensory areas of the monkey (Broadman’s area 3a, 3b, 1 and 2) when a visual cue indicated which of two subsequently presented tactile objects was the correct choice in a forced-choice task. Thus, visual–somatosensory co-representation is seen in early somatosensory cortices and we might predict that similar auditory inputs would be found in these areas, consistent with the topography of the interaction effect seen in the present results. A recent investigation, more directly related to the current study, used magneto-encephalographic (MEG) recordings [10] in humans to assess the effects of continuous auditory (music) stimulation on the somatosensory evoked field (SEF) to median nerve stimulation. Unlike the results of the current study, no effect of auditory stimulation on the early contra-stimulation components of the SEF (1 M, 16–20 ms; 2 M, 20–30 ms; 3 M, 35–45 ms or 4 M) was found, whereas visual stimulation significantly modulated both the 3 M and 4 M. However, they did find a significant AS integration effect on a later SEF component over the ipsilateral hemisphere (MI, 70–100 ms), manifesting as an attenuation of this response in about half of their subjects. The use of continuous stimulation paired with phasic somatosensory stimulation may not have been the ideal stimulus configuration for investigating multisensory integration properties. That is, neural integration processes between sensory modalities have been demonstrated to be highly dependent on both temporal concordance [18] as...
Fig. 2. Multisensory auditory–somatosensory interactions. (A) Group-averaged voltage waveforms (n=8) recorded from a right central electrode site (C4) are plotted. The response to simultaneously presented auditory–somatosensory stimulation in the left hemifield (blue trace, simultaneous) is compared to the derived response from summing the responses to auditory-alone and somatosensory-alone stimulation (green trace, summed). A subtraction waveform (red dashed trace, ‘Summed’ minus ‘Simultaneous’) shows the time-course of the multisensory interaction effect. (B) Scalp current density maps (derived from 64-channel voltage recordings) of the difference waveform at 65 ms, seen from the right side of the head (left map) and the top of the head (right map). A negative (blue isocontours) focus over lateral post-central right scalp can be seen, with a concurrent positive focus (red isocontours) over central scalp. Electrode site C4 (traces in A) is indicated by the black and white disk. Isocontour lines are calibrated to indicate 0.05 μV/cm² increments. Note that the polarity of foci is arbitrary, being dependent upon the direction of the subtraction. (C) Scalp current density maps of the difference waveform at 80 ms. The right side-view shows the emergence of an occipito-temporal positive focus and an anterior shift of the lateral negative focus, suggesting the addition of another generator in the region of the posterior superior temporal plane.
well as spatial concordance [16,17] of the bimodal stimuli which give rise to the perceptions of unified stimulation or ‘binding’. Subjects in the Lam et al. [10] experiment are unlikely to have perceived continuous auditory stimulation as strongly related to the somatosensory pulses.

Concerning the second prediction, a recent investigation in the belt region of the auditory cortex of macaque monkeys [13] has shown AS co-representation in the auditory association areas of the superior temporal plane, immediately posterior to primary auditory cortex. In this study, both the auditory and somatosensory inputs to this area appear to be feedforward in that initial activation to stimuli of both modalities is seen in the input layer (lamina 4). The possibility of direct somatosensory inputs to auditory cortices may also derive support from magnetoencephalographic (MEG) results in congenitally deaf humans, where it has been found that auditory cortices can reorganize to represent tactile inputs [12]. These results suggest that relatively latent somatosensory inputs are unmasked when primary auditory inputs are lesioned, although some caution should be exercised in extending these findings to the normal population. Further, the timing results from the current study are consistent with recent findings from an auditory–visual integration study in humans which showed remarkably early modulations of the ERP over visual cortices (as early as 40 ms post-stimulation) by concurrent audio–visual stimulation during an object recognition task [5]. Both the monkey intracranial results of Lindsley and colleagues [13] and the early timing effects reported in the current study and by Giard et al. [5] suggest that the cross-modal convergence that underlies multisensory integration occurs early in the cortical processing stream. These results dispute the widely held view that multisensory integration occurs only in specialized higher-order association regions.

It should be noted that the electrical stimulation we use is a relatively unnatural form of somatosensory stimulation, which bypasses the sensory receptors. To our advantage, responses to electrical stimulation tend to be of larger amplitude and to be somewhat better synchronized than the response to ordinary cutaneous stimulation. However, intracranial recordings in monkeys have shown that the laminar response profile and onset latencies are highly similar between these stimulation types (see Ref. [20]). Critically, electrical stimulation, as with normal cutaneous stimulation, is still governed by the constraints of normal anatomy. Also, it should be noted that a limitation of the ERP technique is its inability to determine whether effects seen at the scalp surface represent inhibitory or excitatory activations. Accordingly, ongoing intracranial investigations in monkeys at our laboratory are attempting to detail the nature of these early interaction effects [13]. Future studies will investigate the AS integration processes revealed in these data, in the context of behavioral and selective attention tasks (e.g., Ref. [4]), to determine the extent to which these processes are under top-down modulatory influence. Indeed, recent evidence has suggested the existence of AS interactions during focused and divided spatial attention tasks [15,24]. It will also be important to investigate the effect of spatial concordance or disparity between AS bimodal stimuli upon these integration properties, as spatial relationships of differing sensory inputs have been shown to be a main organizing principle in multisensory integration [9,17,25,33]. Further, more detailed analyses involving source localization in individual subject anatomy [22] should provide for more exact localization of the neural generators manifesting AS interaction.

5. Conclusion

This study provides evidence for multisensory interactions between the senses of audition and somatosensation in human subjects. High-density 64-channel event-related potential recordings revealed significant differences between the responses to simultaneous bimodal auditory–somatosensory (AS) stimulation and the summed responses from auditory-alone and somatosensory-alone stimulation. These differences were seen by just 50 ms post-stimulus presentation over the scalp contralateral to the side of stimulation. Scalp-current density mapping of the topography of this interaction effect revealed a distribution over central/post-central scalp, consistent with a generator(s) in somatosensory cortices in the region of the post-central gyrus. Slightly later (70–80 ms), the scalp topography of the interaction effects suggested the addition of generator(s) in the region of posterior auditory cortices. These results, allied with findings from monkey intracranial studies in our laboratory, suggest that multisensory AS integration occurs early in the cortical processing stream.

Acknowledgements

Sincere appreciation to Robert Lindsley for technical support. Work supported by grants from the NIMH (MH1143 to JF; MH49334 and MH01439 to DCJ).

References


