Dynamics of Within-, Inter-, and Cross-Modal Attentional Modulation

Tetsuo Kida, Koji Inui, Emi Tanaka, and Ryusuke Kakigi
Department of Integrative Physiology, National Institute for Physiological Sciences, Okazaki, Japan

Submitted 2 September 2009; accepted in final form 2 December 2010

INTRODUCTION

Attention is often directed to sensory signals coming from different senses in the same space. However, previous studies have traditionally focused on the effect of spatial attention within single sensory modalities (within-modal spatial attention) and the effect of directing attention to one sense compared with the other senses (intermodal attention) on cortical neuronal activity. Furthermore, recent studies have been revealing that the effects of spatial attention directed to a certain location in a certain sense spread to the other senses at the same location in space (cross-modal spatial attention). The present study used magnetoencephalography to examine the temporal dynamics of the effects of within-modal and cross-modal spatial and intermodal attention on cortical processes responsive to visual stimuli. Visual or tactile stimuli were randomly presented on the left or right side at a random interstimulus interval and subjects directed attention to the left or right when vision or touch was a task-relevant modality. Sensor-space analysis showed that a response around the occipitotemporal region at around 150 ms after visual stimulation was significantly enhanced by within-modal, cross-modal spatial, and intermodal attention. A later response over the right frontal region at around 200 ms was enhanced by within-modal spatial and intermodal attention, but not by cross-modal spatial attention. These effects were estimated to originate from the occipitotemporal and lateral frontal areas, respectively. Thus the results suggest different spatiotemporal dynamics of neural representations of cross-modal attention and intermodal or within-modal attention.

Kida T, Inui K, Tanaka E, Kakigi R. Dynamics of within-, inter-, and cross-modal attentional modulation. J Neurophysiol 105: 674–686, 2011. First published December 8, 2010; doi:10.1152/jn.00807.2009. Numerous studies have demonstrated effects of spatial attention within single sensory modalities (within-modal spatial attention) and the effect of directing attention to one sense compared with the other senses (intermodal attention) on cortical neuronal activity. Furthermore, recent studies have been revealing that the effects of spatial attention directed to a certain location in a certain sense spread to the other senses at the same location in space (cross-modal spatial attention). The present study used magnetoencephalography to examine the temporal dynamics of the effects of within-modal and cross-modal spatial and intermodal attention on cortical processes responsive to visual stimuli. Visual or tactile stimuli were randomly presented on the left or right side at a random interstimulus interval and subjects directed attention to the left or right when vision or touch was a task-relevant modality. Sensor-space analysis showed that a response around the occipitotemporal region at around 150 ms after visual stimulation was significantly enhanced by within-modal, cross-modal spatial, and intermodal attention. A later response over the right frontal region at around 200 ms was enhanced by within-modal spatial and intermodal attention, but not by cross-modal spatial attention. These effects were estimated to originate from the occipitotemporal and lateral frontal areas, respectively. Thus the results suggest different spatiotemporal dynamics of neural representations of cross-modal attention and intermodal or within-modal attention.

Address for reprint requests and other correspondence: T. Kida, National Institute for Physiological Sciences, Department of Integrative Physiology, 38 Nishigo-naka, Myodaiji, Okazaki 444-8585, Japan (E-mail: nikit@nips.ac.jp)
characterized by the modulation of visual-evoked responses by directing attention to touch at the location where visual stimuli appear. This effect demonstrates a link from touch to vision. The effect of intermodal attention is characterized as the overall effect of directing attention to vision compared with touch. In particular, our study deals with two questions with respect to what the spatiotemporal dynamics of a neural representation of the link from touch to vision in spatial attention is and whether there is some difference in spatiotemporal dynamics among modulations by within-modal, intermodal, and cross-modal attention. We hypothesized that the occipitotemporal activity is modulated by cross-modal spatial attention as well as by within-modal spatial and intermodal attention because the occipitotemporal activity would involve neural activity in both unimodal and multisensory cortices (Beauchamp et al. 2004a,b, 2008; Macaluso et al. 2002a, 2005; Noesselt et al. 2007; Tanaka et al. 2009a,b). We also hypothesized that within-modal attention and intermodal attention produce similar attentional influence regarding time and space. An electroencephalographic (EEG) study reported the similarity of modulations by within-modal and intermodal attention (Talsma and Kok 2002), which is assumed to be mediated by modality-specific cortex. The present study measured the gradient of magnetic fields and performed the source analysis to demonstrate whether modality-specific cortex actually contributes to the similarity of attentional modulations. The data on tactile-evoked responses have been published elsewhere (Kida et al. 2007).

METHODS

Subjects

Recordings were obtained from 11 healthy right-handed subjects (2 women, 9 men; age range: 23 to 40 yr). All subjects gave written informed consent prior to the study, which was first approved by the Ethics Committee of the National Institute for Physiological Sciences.

Stimulation

The subjects were seated in a magnetically shielded and semidarkened room and put their hands and forearms comfortably on an obliquely oriented board in front of them. Eight kinds of stimuli differing in spatial location (left and right), sensory modality (vision and touch), and number of pulses (single and double) were presented in a random order at a random interstimulus interval (ISI, 800–1,200 ms). Visual stimuli were presented through two plastic optical fibers (diameter 1 cm, length 3 m; Multicore POF Luminous, Asahi Kasei, Japan), each on the left and right side. The duration of the single-pulse visual stimulus was 0.2 ms and its intensity was 2.5 times the sensory threshold. In the case of the double-pulse stimulus, two electrotactile pulses separated by an interval of 180 ms were applied.

Task

Subjects executed four different task conditions in which the direction of attention varied (left/tactile, right/tactile, left/visual, and right/visual attention conditions). The subjects were instructed to fixate on a point at a viewing distance of 40 cm and to put the second digit of each hand near the visual stimuli (i.e., near one end of two optical fibers). In the vision-relevant/attend-right condition, subjects were instructed to attend covertly to visual stimuli presented to the right visual field and to count silently the number of visual double-pulse stimuli presented there. Therefore in this condition, the visual single-pulse stimulus presented in the right visual field was regarded as a spatially attended stimulus from the task-relevant modality, whereas the visual single-pulse stimulus presented to the left visual field was regarded as a spatially unattended stimulus from the relevant modality. In each of the other three conditions, subjects were asked to attend to the designated side and to count only the number of double-pulse stimuli presented on the designated side in the relevant modality. These conditions were run in a random order. Each condition contained about 600 trials (about 250 single-pulse electrotactile, 250 single-pulse visual, 50 double-pulse electrotactile, and 50 double-pulse visual trials, presented approximately equally on the left and right sides), divided into two blocks. Each block lasted about 5 min and an experiment lasted about 50–60 min. Before each condition, subjects were given verbal instructions about the direction of attention to be directed and the type of stimuli to be counted. They had to report the number after the termination of each block.

Recordings

The MEG signal was recorded with a helmet-shaped detector array (Vectorview, ELEKTA Neuromag Yo, Helsinki, Finland), which in each of 102 locations has three sensors (306 sensors in all), two orthogonal planar gradiometers, and one magnetometer coupled to a multi-SQUID (superconducting quantum interference device). Signals were filtered with a band-pass filter of 0.03–150 Hz and digitized at a sampling rate of 1,024 Hz.

To allow the transformation of coordinates between different systems, the next procedure was carried out before the acquisition of MEG data. First, the three anatomical landmarks were digitized (nasion and bilateral preauricular points) using a three-dimensional (3D) digitizer outside the shielded room to determine a head-based coordinate system. The x-axis was fixed with the preauricular points, the positive direction being to the right. The positive y-axis passed through the nasion and the z-axis thus pointed upward. Then, four head position indicator (HPI) coils were attached on the subject’s head and their locations were digitized in relation to the three anatomical landmarks. After the fixation of the subject’s head to the helmet-shaped sensor array in the shielded room, magnetic fields generated by currents fed to the four HPI coils were measured. The locations of the HPI coils were calculated in the device-based coordinate system by a fitting procedure to allow transformation between different systems. After that, the acquisition of MEG data started.

Eye movements were monitored with a camera (ISCAN, Woburn, MA). Trials with horizontal and vertical eye movements of >0.5° or with MEG signals >3,000 ft/cm were rejected from the averaging of MEG data; 8.1 (SE, 2.7) trials (of 250 including 125 right and 125 left) were rejected for the single-pulse visual stimulation in each condition according to these criteria.

Analysis

We analyzed the MEG signals recorded from 102 pairs of two-orthogonal planar-type gradiometers. The period of analysis was from 100 ms before to 400 ms after the stimulus. The baseline correction was performed using the prestimulus samples. In the present study, neural responses to the visual single-pulse stimulus (frequently pre-
sented stimulus) were analyzed to see a pure spatial selective attentional effect on visual-evoked responses (García-Larrea et al. 1995; Kida et al. 2004, 2007). The large noise included in the recorded signal was removed using the signal space projection (SSP) technique and then the data were filtered off-line with a low-pass filter of 50 Hz.

For the sensor-space analysis, vector sum signals from each pair of gradiometers were calculated as an inner product (i.e., scalar product) by squaring the signals from the two gradiometers at each sensor’s location, summing together, and then computing a square root of the sum (called the root sum square [RSS]) as follows

\[ \text{RSS} = \sqrt{(\delta B_x/\delta x)^2 + (\delta B_y/\delta y)^2} \]

where \( B_z \) is the strength of magnetic fields and \( \delta x \) and \( \delta y \) are longitudinal and latitudinal derivatives, respectively. The baseline correction was performed using the prestimulus samples after this calculation. This measure has been successfully used to examine the distribution and modulation of cortical responses in a variety of sensory modalities including vision, audition, and touch in our laboratory (Kida et al. 2006, 2007; Tanaka et al. 2009a,b). The calculation was carried out for each location of the 102 sensors. Next, we used the RSS waveforms and isocontour map of RSS amplitude to look for a peak channel showing the greatest amplitude for each prominent response. Then, the peak amplitude and latency of prominent responses in the RSS waveform were measured at the peak channel. The RSS signal was computed after averaging individual trials and thus the present analysis is limited to phase-locked (evoked) activity. The across-sensor mean of the RSS signal (MRSS signal) was also calculated around the occipitotemporal and lateral frontal regions. The mean amplitudes of the MRSS signal were measured between 120 and 180 ms for the occipitotemporal region and between 130 and 280 ms for the lateral frontal region.

For the mean amplitude of the MRSS waveform, a four-way ANOVA with repeated measures was performed with spatial attention (attended and unattended), relevant modality (touch and vision), hemisphere (right and left), and stimulation (right and left) as factors. This statistical test revealed that the occipitotemporal response was observed clearly in the hemisphere contralateral to the side stimulated. Accordingly, for both the peak amplitude and the latency of the RSS waveform, a three-way ANOVA with repeated measures was performed with spatial attention (attended and unattended), relevant modality (touch and vision), and hemisphere (right and left) for the occipitotemporal response. For the lateral frontal response in the RSS waveform, a three-way ANOVA was performed with spatial attention and relevant modality because the statistical test of the MRSS signal showed that this response was found clearly for the right hemisphere in response to stimulation of the left visual hemifield. If the sphericity assumption was violated in Mauchly’s sphericity test, the Greenhouse–Geisser correction was used. The statistical significance was set at \( P < 0.05 \).

To display the effects of different types of attention, the difference waveform was calculated by subtracting the MRSS signal for spatially attended visual stimuli from that for spatially unattended visual stimuli (within-modal spatial attention effect), by subtracting the signal for spatially attended visual stimuli from that when attention was directed to touch at the location where visual stimuli appear (intermodal attention effect) and by subtracting the signal when attention was directed to touch at the location where visual stimuli appear from that when attention was directed to touch at the location where visual stimuli did not appear (cross-modal spatial attention effect).

For the source-space analysis, an equivalent current dipole source analysis was used to estimate sources generating recorded magnetic responses to visual stimuli. The sources were modeled with time-varying current dipoles (Hamalainen 1993; Sarvas 1987). The MEG signals were evaluated at successive time points by a least-squares search in a spherical volume conductor model by using 18–24 sensors around a sensor that had been used to measure the peak amplitude of RSS waveforms. Since the planar gradient peaks right above a current-dipole source, a less extensive set of sensors is needed to pick up the essential field data from a local source (in other words, the restriction of the analysis to a subset of planar sensors is very efficient in avoiding contamination from distant sources). This analysis resulted in the 3D location, moment, and direction of each equivalent current dipole (ECD) in a spherical conductor model. The goodness-of-fit (GOF) value of an ECD was calculated to indicate in percentage terms how much the dipole accounts for the measured field variance. ECDs with \( >85\% \) of the GOF value were accepted. The signal-to-noise ratio (SNR) was, on average, 15.3\% (SE = 2.2\%), which was computed by dividing the estimates of noise levels by the response amplitude and then expressed in percentage terms. The locations of ECDs were transformed to Talairach coordinates to determine corresponding cortical regions. Finally, all sensors were used to compute the time-varying multidipole model, allowing the strength of the previously found ECDs to change over the entire period of the analysis while the source locations and orientations were kept fixed.

The obtained source strength was statistically analyzed by ANOVA.

**Minimum norm estimation**

Minimum norm estimation (MNE) was used to estimate the distribution of cortical current sources on the cortical surface (see Supplemental Materials for details).1 MNE is a common method of estimating distributed electrical current sources in the brain in a particular time point (Hamalainen and Ilmoniemi 1984). Out of the many current distributions that can account for the recorded sensor data, the solution with the minimum \( L2 \) norm (i.e., the minimum total power of the current distribution) is used. First, the forward model (lead field matrix \( L \)) of all the sources is calculated. Then, the activities of all sources \( S(t) \) are computed from the data matrix \( D(t) \) using an inverse regularized by the estimated noise covariance matrix

\[ S(t) = RL^2(LRL^2 + C_0)^{-1}D(t) \]

where \( L \) is the lead field matrix of the distributed regional source model, \( C_0 \) denotes the noise correlation matrix in the sensor space, and \( R \) is a weighting matrix in the source space. Depth weighting and spatiotemporal weighting were used. For the depth weighting, the lead field of each regional source is scaled with the largest singular value of the singular value decomposition of the source’s lead field. The spatiotemporal weighting method proposed by Dale and Sereno (1993) was used. The spatiotemporal weighting tries to assign large weights to sources that are assumed to be more likely to contribute to the recorded data. The correlation \( p_i \) of the lead field of regional source \( i \) with the inverse of the data covariance matrix was computed along with the largest singular value \( \lambda_{\text{max}} \) of the data covariance matrix. The weighting matrix \( R \) is a diagonal matrix with weights:

\[ \{1/[(1 + \lambda_{\text{max}}(1 - p_i)]\}. \]

Each of the two components of a regional source receives the same weighting value. These calculations were done on a standard brain in individual subjects so that we could average the data across subjects for each of the distributed current source points in the source space for each condition. The strength of the source activity is transformed to a color-coded image of the cortical surface. The data were also pooled across different points in each area of interest.

Performance of the counting task was assessed by calculating the absolute deviation of the subject’s target count from the correct target count in each of two blocks (absolute error) and then converting the total number of absolute errors to a percentage of the total correct count (i.e., error rate; a higher error rate represents more frequent failures to count the stimulus).

---

1 The online version of this article contains supplemental data.
RESULTS

Behavioral performance

The rate of count error was very low [0.96% (SE, 0.64) in the touch-relevant/attend-right condition, 1.41% (0.73) in the touch-relevant/attend-left condition, 0.96% (0.64) in the vision-relevant/attend-right condition, and 1.41% (0.73) in the touch-relevant/attend-left condition], indicating that participants almost perfectly executed the counting task in all conditions. A two-way ANOVA [direction of attention (left vs. right) × relevant sensory modality (vision vs. touch)] also showed that the error rates were not significantly affected by direction of attention \([F(1,10) = 0.94, P = 0.36, \text{n.s.}]\) and relevant sensory modality \([F(1,10) = 0.002, P = 0.96, \text{n.s.}]\) and there was no interaction \([F(1,10) = 0.002, P = 0.96, \text{n.s.}]\).

Response profile

Figure 1 indicates data from an individual subject. The responses to stimulation of the left and right visual fields in the spatially attended/vision-relevant condition are shown on the top row and the superimposed waveforms from different four attention conditions are shown on the bottom-most row. There were consistently peaks of the RSS waveform in most subjects as follows: a peak around the occipitotemporal area at about 120–180 ms (mean, 148.8 ms; M150ot) in the contralateral hemisphere (i.e., the left hemisphere when the
right visual field was stimulated and the right hemisphere when the left visual field was stimulated) and a peak around the frontal area at 130–250 ms (mean, 177.7 ms; M180f) in the right hemisphere when the left visual field was stimulated. The distribution and focus of these magnetic responses are also shown in 3D and two-dimensional (2D) maps in Fig. 1.

**Occipitotemporal response (M150ot)**

Figure 2 (left two columns) indicates the waveforms of MRSS from 11 occipitotemporal sensors. These bilateral responses peaked at 140 and 160 ms in the contralateral and ipsilateral hemispheres, respectively. The difference waveform is also displayed in Fig. 3. Figure 4 (first row) shows the mean amplitude of MRSS between 120 and 180 ms corresponding to the range of M150ot. A four-way ANOVA indicated main effects of spatial attention \([F(1,10) = 29.6, P < 0.001]\) and relevant modality \([F(1,10) = 25.7, P < 0.001]\), with amplitude enhanced by spatial attention and by intermodal attention to vision. The effect of spatial attention was even observed when touch was task-relevant \((P < 0.001)\), i.e., the effect of cross-modal spatial attention. There was significant interaction between stimulus side and hemisphere \([F(1,10) = 22.5, P < 0.005]\), such that the amplitude for the left visual hemifield stimulus was greater in the right hemisphere than that in the left \((P = 0.004)\), but that for the right visual hemifield stimulus was greater in the left hemisphere \((P = 0.005)\), i.e., the response was greater in the hemisphere contralateral to the side stimulated. There was also significant interaction between relevant modality and hemisphere \([F(1,10) = 5.5, P < 0.05]\), such that directing attention to vision compared with touch enhanced the amplitude in both hemispheres \((P = 0.1\) and 0.01 for the left and right hemispheres, respectively), although the degree of enhancement was greater in the right hemisphere \((2.8 \pm 0.6 \mu T/cm)\) than that in the left \((1.3 \pm 0.4 \mu T/cm)\). There was also significant interaction among relevant modality, hemisphere, and stimulus side. This three-way interaction indicates that the enhancement of the amplitude by directing attention to vision compared with touch was significant in the right hemisphere for both visual hemifield stimuli \((P = 0.001\) and 0.02 for the left and right visual hemifield stimuli, respectively), whereas in the left hemisphere it was not significant for the left visual hemifield stimulus \((n.s., P = 0.19)\), but was significant for the right visual hemifield stimulus \((P = 0.04)\). In addition, the enhancement in the right hemisphere was greater for the left visual hemifield stimulus than that for the right stimulus. Thus this three-way interaction indicates both a right-hemisphere dominance and a contralateral-hemisphere dominance regard-

![Fig. 2](image-url)

**Fig. 2.** Attentional modulation at the occipitotemporal and lateral frontal sensors. The bottom row shows grand-averaged waveforms of the across-sensor mean of the RSS (MRSS) signal in different conditions. SE is shown as shaded area. The sensors used to calculate MRSS (purple) are shown in the top row.
ing the effect of intermodal attention to vision compared with touch. There was also significant interaction among spatial attention, hemisphere, and stimulus side $[F(1,10) = 13.7, P < 0.005]$, such that the enhancement of the amplitude by spatial attention was significant in the hemisphere contralateral to the side stimulated ($P = 0.01$ and 0.05 in the left and right hemispheres contralateral to the side stimulated, respectively), whereas it did not reach the level of significance in the ipsilateral hemisphere ($P = 0.08$ and 0.09 in the left and right hemispheres ipsilateral to the side stimulated, respectively). Thus this interaction indicates the dominance of the contralateral hemisphere regarding the effect of spatial attention. There was also significant interaction between spatial attention and relevant modality $[F(1,10) = 6.9, P < 0.01]$, such that the enhancement by spatial attention was greater when vision was relevant ($2.4 \pm 0.5 \text{ fT/cm}$) than when touch was relevant ($1.4 \pm 0.2 \text{ fT/cm}$). Finally, there was significant interaction among spatial attention, relevant modality, and hemisphere $[F(1,10) = 5.8, P < 0.05]$. Regarding this three-way interaction, significant enhancement of the amplitude by spatial attention was observed at all the low levels, but the enhancement was greatest in the right hemisphere when vision was task-relevant ($2.9 \pm 0.6 \text{ fT/cm}$), compared with the other levels ($1.6 \pm 0.5$ and $1.5 \pm 0.5 \text{ fT/cm}$, for the left hemisphere when vision or touch was relevant, respectively; $1.4 \pm 0.4 \text{ fT/cm}$ for the right hemisphere when touch was relevant). Thus this three-way interaction indicates that the summed effect of within-modal spatial and intermodal attention was dominant in the right hemisphere.

We also analyzed the RSS signal to examine attentional modulation of activity that better reflects the underlying signal source. Table 1 shows the peak amplitude of the RSS signal. For the peak amplitude, a three-way ANOVA indicated significant main effects of spatial attention $[F(1,10) = 58.1, P < 0.001]$ and relevant modality $[F(1,10) = 29.5, P < 0.001]$, with amplitude enhanced by spatial attention or by intermodal attention to vision. The effect of spatial attention was also observed when touch was task-relevant ($P < 0.005$). Furthermore, there was significant interaction between spatial attention and relevant modality $[F(1,10) = 6.5, P < 0.05]$, such that the enhancement by spatial attention was greater when vision was relevant ($15.3 \pm 3.3 \text{ fT/cm}$) than when touch was relevant ($10.0 \pm 2.1 \text{ fT/cm}$). The peak latency of the response was not changed by spatial attention or relevant modality.
interaction between relevant modality and hemisphere dominance in this latency range. There was also significant interaction indicates the contralateral and right hemisphere stimulus side in the left hemisphere (n.s., \( P < 0.15 \)), such that the interaction was greater for the left visual hemifield stimulus than for the right visual stimulus in the right hemisphere (\( P = 0.005 \)), but was not different between the stimulus side in the left hemisphere (n.s., \( P = 0.15 \)). This interaction indicates the contralateral and right hemisphere dominance in this latency range. There was also significant interaction between relevant modality and hemisphere \( F(1,10) = 9.7, P < 0.01 \), such that the amplitude was enhanced in the right hemisphere by directing attention to vision compared with touch (\( P = 0.001 \)), but was not changed in the left hemisphere (\( P = 0.1 \)).

For the amplitude of M180f in the RSS waveform (Table 1), a two-way ANOVA indicated a main effect of spatial attention \( F(1,10) = 25.4, P < 0.01 \) and relevant modality \( F(1,10) = 17.2, P < 0.01 \), with amplitude enhanced by spatial attention and when vision was relevant. There was also a significant interaction between spatial attention and relevant modality \( F(1,10) = 7.7, P < 0.05 \), such that spatial attention effect had an effect when vision was relevant (\( P < 0.01 \)), but not when touch was relevant (n.s., \( P = 0.25 \)).

**Frontal response (M180f)**

A response over the frontal area in the right hemisphere was found at 130–280 ms after the left visual field stimulus (M180f), as shown in Fig. 2 (right two columns). Figure 4 (bottom) shows the mean amplitude of MRSS between 130 and 280 ms corresponding to the range of M180f. There was a main effect of hemisphere \( F(1,10) = 5.1, P < 0.05 \), with greater amplitude in the right hemisphere than that in the left. There was also a main effect of relevant modality \( F(1,10) = 53.3, P < 0.001 \), with amplitude enhanced by directing attention to vision compared with touch. There was significant interaction between hemisphere and stimulus side \( F(1,10) = 37.1, P < 0.005 \), such that the amplitude was greater for the left visual hemifield stimulus than for the right visual stimulus in the right hemisphere (\( P = 0.005 \)), but was not different between the stimulus side in the left hemisphere (n.s., \( P = 0.15 \)). This interaction indicates the contralateral and right hemisphere dominance in this latency range. There was also significant interaction between relevant modality and hemisphere \( F(1,10) = 9.7, P < 0.01 \), such that the amplitude was enhanced in the right hemisphere by directing attention to vision compared with touch (\( P = 0.001 \)), but was not changed in the left hemisphere (\( P = 0.1 \)).

For the amplitude of M180f in the RSS waveform (Table 1), a two-way ANOVA indicated a main effect of spatial attention \( F(1,10) = 25.4, P < 0.01 \) and relevant modality \( F(1,10) = 17.2, P < 0.01 \), with amplitude enhanced by spatial attention and when vision was relevant. There was also a significant interaction between spatial attention and relevant modality \( F(1,10) = 7.7, P < 0.05 \), such that spatial attention effect had an effect when vision was relevant (\( P < 0.01 \)), but not when touch was relevant (n.s., \( P = 0.25 \)).

**Source analysis**

Estimated current dipoles were displayed on the MR images (Fig. 5). Sources for M150ot and M180f were estimated in the occipitotemporal and lateral frontal regions, respectively. Areas responsible for the occipitotemporal sources were the superior and middle temporal gyri (STG/MTG, area 21/22/39/40/41), middle occipital gyrus (MOG, areas 19/39), and temporoparietal junction (TPJ, including angular and supramarginal gyrus, areas 39/40), whereas those for the frontal regions were the middle and inferior frontal gyrus (MFG/IFG). Other sources were also observed in the cuneus (areas 19/18), precuneus (area 7), superior and inferior parietal sulci (areas 39/40), and inferior temporal regions (including the inferior temporal and fusiform gyri, areas 19/37) in a limited number of subjects. Eventually, four to six signal sources were found in individuals as indicated in Table 2.

The time course of the ECD moment calculated by the multidipole source analysis and its peak moment are shown in Figs. 6 and 7. For the source strength (ECD moment) of the occipitotemporal response in the hemisphere contralateral to the side stimulated, a three-way ANOVA with repeated measures was performed with hemisphere (right and left), relevant modality (vision and touch), and spatial attention (attended and unattended). There was a main effect of hemisphere \( F(1,10) = 8.3, P < 0.05 \), with greater source strength in the right hemisphere than that in the left. There was also a main effect of relevant modality \( F(1,10) = 25.7, P < 0.001 \), with source strength increased by intermodal attention to vision compared with touch. Furthermore, there was a main effect of spatial attention \( F(1,10) = 37.1, P < 0.001 \), with source strength increased by spatial attention. A significant interaction was observed between relevant modality and spatial attention, such that the enhancement by spatial attention was greater when vision was relevant (6.1 ± 1.1 nAm) than when touch was relevant (3.9 ± 0.8 nAm).

**TABLE 1.** Mean value of the peak amplitude (fT/cm) of RSS signal

<table>
<thead>
<tr>
<th>Hemisphere</th>
<th>Stimulus Side</th>
<th>Relevant Modality</th>
<th>Spatial Attention</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A</td>
<td>49.1</td>
<td>33.3</td>
</tr>
<tr>
<td>M150ot</td>
<td></td>
<td></td>
<td>U</td>
<td>38.6</td>
<td>28.1</td>
</tr>
<tr>
<td>M180f</td>
<td></td>
<td></td>
<td>A</td>
<td>55.0</td>
<td>40.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>U</td>
<td>36.2</td>
<td>26.0</td>
</tr>
</tbody>
</table>

V, vision; T, touch; A, spatially attended; U, spatially unattended.

FIG. 5. Locations of sources underlying M150ot (white dot) and M180f (black dot). A dot represents a source in a single subject. The source generating M150ot was located in the occipitotemporal cortex including the middle occipital gyrus (MOG), superior and middle temporal gyrus (STG/MTG), and temporoparietal junction (TPJ), whereas that of M180f was in the lateral frontal cortex including the middle and inferior frontal gyrus (MFG/IFG). Some additional sources were also found in the cuneus, precuneus, and intraparietal lobule (not shown here).
For the source strength of the occipitotemporal response ipsilateral to the side stimulated, a two-way ANOVA with repeated measures was performed with spatial attention and relevant modality, separately in the right and left hemispheres. For the source strength of the right occipitotemporal response to the right visual hemifield stimulus, there was a main effect of relevant modality $[F(1,7) = 66.6, P < 0.001]$, with source strength increased by intermodal attention to vision. There was also significant interaction between spatial attention and relevant modality $[F(1,7) = 14.6, P < 0.01]$, such that spatial attention enhanced source strength when vision was relevant but not when touch was relevant. For the source strength of the left occipitotemporal response to the left visual hemifield stimulus, there was a main effect of relevant modality $[F(1,7) = 18.1, P < 0.005]$, with source strength increased by intermodal attention to vision compared with touch. There was also a main effect of spatial attention $[F(1,7) = 14.2, P < 0.01]$, with source strength increased by spatial attention.

For the source strength of the lateral frontal response in the right hemisphere to the left visual hemifield stimulus, a two-way ANOVA was performed with spatial attention (attended and unattended) and relevant modality (vision and touch). There was a main effect of relevant modality $[F(1,7) = 19.8, P < 0.005]$ and spatial attention $[F(1,7) = 14.2, P < 0.01]$, with source strength increased by intermodal attention to vision and by spatial attention. There was also significant interaction between relevant modality and spatial attention $[F(1,7) = 14.2, P < 0.01]$, such that source strength was enhanced by spatial attention when vision
was relevant \((P < 0.05)\) but not when touch was relevant \((n.s., P = 0.28)\). For the source strength of the lateral frontal response in the right hemisphere to the right visual hemifield stimulus, a two-way ANOVA did not show any significant result. For the left frontal responses, we did not find enough sources (five for the left stimulus and six for the right stimulus) and thus did not perform a statistical evaluation of the responses.

Further, we performed a source analysis with the minimum norm estimation (MNE) (Figs. 8 and 9 and Supplementary Materials). Overall, response enhancement by within-modal spatial attention and intermodal attention was widely distributed over occipitotemporal regions up to lateral frontal regions. In contrast, response enhancement by cross-modal spatial attention originated in the occipitotemporal cortex, especially the TPJ. The results from the distributed source estimation support findings of RSS (MRSS) and dipole source analyses. Thus the source analysis shows that these extrastriate and some multisensory areas and lateral frontal cortices produced attentional modulation of the recorded magnetic fields. Notably, the TPJ is expected to

**Fig. 8.** Results of minimum norm estimation (MNE). The distribution of source strength averaged across all the subjects is mapped on the surface of the brain at different times of interest \((TOIs, 120–170 \text{ and } 130–280 \text{ ms})\). The source strength of the response to the left visual hemifield stimulus is shown here. Also see supplemental materials for details.

**Fig. 9.** Topography of attentional modulation in the MNE analysis. The effects of within-modal spatial, intermodal, and cross-modal spatial attention (WMA, IMA, CMA) were calculated from the difference between attentional conditions in the source space. The source strength of the response in the contralateral hemisphere to the left or right visual hemifield is shown here.
be important for modulation by cross-modal spatial attention.

**DISCUSSION**

The present study used MEG to examine attentional modulation of neural responses to visual stimuli in a visual–tactile spatial attention task. The results revealed the different spatiotemporal dynamics of within-, inter-, and cross-modal attentional modulation.

**Effect of within-modal spatial attention**

M150ot was enhanced in amplitude by spatial attention when vision was task-relevant, indicating an effect of within-modal spatial attention. Most EEG studies have reported no modulation of the C1 component peaking at a latency of about 70 ms and consistent modulation after a latency of 100 ms (e.g., P1 component) by within-modal spatial attention (Hillyard et al. 1998; Luck et al. 1990, 1996; Mangun and Hillyard 1988; Teder-Salejarvi and Hillyard 1998). The C1 and P1 components have been estimated to be generated from striate and extrastriate cortices, respectively (Mangun et al. 2001; Martinez et al. 1999; Woldorff et al. 2002).

Therefore these authors assumed that the feedforward response in V1 is not affected by attention and the first stage for attentional modulation is V2 (Woldorff et al. 2002) and V3 (Martinez et al. 1999). In addition, delayed attentional modulation (later than 150 ms) is observed in human V1, which is considered due to delayed feedback from extrastriate visual areas (Di Russo et al. 2003; Martinez et al. 2001; Noesselt et al. 2002). A similar delayed attentional modulation is observed in the monkey V1 (Mehta et al. 2000a,b; Schroeder et al. 2001). Most MEG studies have also reported an attentional enhancement of relatively late processing (100 ms) in several cortical areas, including the inferotemporal, middle temporal, and occipital regions when attention was directed to a space (Noguchi et al. 2007) or to a face (Furey et al. 2006; Lueschow et al. 2004) and during a visual search (Hopf et al. 2000). Thus in many EEG and MEG studies, attentional modulation starts at around 100 ms after visual input (but see Poghosyan and Ioannides 2008), consistent with the present study. Also, the present sensor-space and source-space analyses indicated that the effect of within-modal attention on M150ot was generated in extrastriate and/or multimodal areas. Our previous EEG and MEG studies using dipole source estimation have also shown that the middle occipital gyrus (MOG) and surrounding areas responded to the onset, offset, and change of face stimuli or a star-shaped visual stimulus or changes of visual stimuli during multisensory stimulation, supporting the present findings (Tanaka et al. 2008, 2009a,b). In recent years, converging evidence provided by invasive macaque work, human fMRI and MEG (Beauchamp et al. 2004a,b, 2007, 2008; Driver and Noesselt 2008; Macaluso and Driver 2005; Tanaka et al. 2009b) suggests that in the occipitotemporal region there are trimodal integration areas like the posterior part of the superior temporal sulcus (STS) and the TPJ. Our ECD estimation actually localized the signal source in these regions in some subjects. Additional analysis by distributed source estimation ($L_2$-minimum norm estimation) showed modulation of the multiple cortical areas in the occipitotemporal region by within-modal spatial attention. Thus it is assumed that the modulation by within-modal spatial attention originated from enhanced activity in these cortical regions.

M180f was also enhanced by within-modal spatial attention. The present analyses showed that the effect of within-modal attention on M180f is generated in the lateral frontal areas, including the MFG and IFG. These frontal cortices have also been considered to underlie a wide variety of higher-order cognitive functions, including voluntary and involuntary attention (Corbetta and Shulman 2002; Kanwisher and Wojciulik 2000) or executive control function (Miller 2000), task set (Sakai 2008), adaptive coding (Duncan 2001), and long-term memory (Simons and Spiers 2003), depending on the experimental task used. Our previous MEG study has localized a neural response evoked by a visual change at around this latency in the MFG/IFG (Tanaka et al. 2009b), suggesting that the MFG/IFG activity is part of a network for the detection of change along with the MOG activity. An fMRI study has also suggested the involvement of MOG, temporoparietal junction (TPJ), and lateral frontal regions in the detection of change (Downar et al. 2000). The network of attentional modulation we observed may manifest modulation of the change detection system. Thus the effect of within-modal spatial attention was distributed across distant cortical areas and the timing of the effect was slightly later in the frontal areas than in the occipitotemporal areas.

**Effect of intermodal attention**

The M150ot and M180f responses were both enhanced in amplitude when vision was relevant compared with when touch was relevant, representing an effect of intermodal attention directed to vision. EEG studies have reported intermodal attentional modulation of responses to visual stimuli (Alho et al. 1992; Eimer and Schroger 1998; Talsma and Kok 2001, 2002; Woods et al. 1992). In early studies, it was unclear whether the observed attentional effects were a pure measure of intermodal attention or whether they reflect spatial attention because stimuli were presented from different locations. Eimer and Schroger (1998) presented visual and auditory stimuli at the same location in a transient attention task and observed the effect of intermodal attention on visual responses that are estimated to originate from modality-specific cortex (Eimer and Schroger 1998). Another EEG study also suggested that intermodal attention effect is mediated by a modality-specific cortex (Talsma and Kok 2001). Moreover, this EEG study found similarity between the effects of intermodal attention and within-modal attention.

In the present study, visual and tactile stimuli were presented at a spatially congruent location and effects of intermodal attention on the occipitotemporal and frontal responses were clearly observed. Also, the effect of intermodal attention was similar to that of within-modal attention regarding time and distribution, consistent with previous studies. EEG studies have suggested the contribution of modality-specific areas to intermodal attentional effects and similarity between within-modal and intermodal attention effects, merely on the basis of the distribution of the potential or the current source density on the scalp. In contrast, we found a similar contribution of
modality-specific areas to within-modal and intermodal attentional effects by recording planar gradients of magnetic fields and estimating the source activity, which strengthens the previous finding. Furthermore, the present results show that, in addition to a primarily modality-specific cortex like the MOG, the effect of intermodal attention could be also mediated by a distant cortical process that includes lateral frontal areas.

**Effect of cross-modal spatial attention**

A significant enhancement by spatial attention was found even when touch was task-relevant, thereby suggesting a cross-modal link from touch to vision. Our previous study has already demonstrated the effect of cross-modal spatial attention on tactile-evoked responses around the sylvian fissure (Kida et al. 2007). Therefore these attentional influences are considered to be bidirectional. Eimer and colleagues (Eimer and Driver 2000; Eimer and Schröger 1998; Eimer et al. 2001, 2002) have extensively studied cross-modal links in spatial attention using EEG. In the early studies using a sustained attention paradigm, they found links from audition to vision and from touch to vision, but not from audition to vision or from vision to touch (Eimer and Driver 2000; Eimer and Schröger 1998). However, later studies using Posner’s paradigm successfully found cross-modal links among vision, audition, and touch (Eimer 2001; Eimer and Driver 2001; Eimer et al. 2001, 2002). The present study showed that cross-modal attention has an influence on the M1500μt, but not on the M1800μt. The source analysis showed the involvement of the occipitotemporal areas, especially the TPJ, in the effect of cross-modal spatial attention. Response modulation in TPJ by cross-modal spatial attention without significant modulation in MOG is supported by recent fMRI and macaque studies (Beauchamp et al. 2004a,b; Macaluso et al. 2005; Noesselt et al. 2007). Thus our finding is consistent with the notion that spatial attention modulates neural activity in the occipitotemporal region, even when the eliciting stimulus comes from a task-irrelevant modality (Eimer and Schröger 1998; Eimer et al. 2002). In contrast, the effect of spatial attention on the later frontal response, which is considered to be largely multimodal, was restricted to the currently task-relevant modality, suggesting that the effect of spatial attention in the task-relevant modality extends to the later processing in frontal areas.

There appears to be a difference in modulation between tactile-evoked responses in our previous study (Kida et al. 2007) and the visual-evoked responses reported here. Previously, we found that the effect of cross-modal (vision-to-touch) attention on a tactile-evoked response around the sylvian fissure did not differ in magnitude from that of within-modal spatial attention. By contrast, in the present study, the enhancement of a visual-evoked response by cross-modal (touch-to-vision) spatial attention was smaller than that by within-modal attention. This biased effect of cross-modal attention in relation to within-modal attention on tactile- and visual-evoked responses provides evidence for a separable-but-linked system underlying spatial attention (Spence and Driver 1996). Three possible systems for the organization of spatial attention have been proposed: 1) a completely supramodal attentional system, which acts commonly on any sense; 2) a completely unimodal system where attention acts individually on each sense; and 3) a separable-but-linked system where there is modality-specific attentional systems, which interplay across different senses. A study with transcranial magnetic stimulation (TMS) provided evidence for the separable-but-linked system by demonstrating modality-specific involvement of the intraparietal sulcus in spatial attention (Chambers et al. 2004). By contrast, fMRI and EEG studies have provided evidence for a supramodal system underlying attentional control (Eimer and van Velzen 2002; Macaluso et al. 2002b). Macaluso and colleagues reported that attending to visual or tactile stimuli produced common activation patterns in the superior and inferior parietal lobule (SPL, IPL) and superior temporal gyrus (STG) in the vicinity of the temporoparietal junction (TPJ). An EEG study using a similar experimental paradigm also found that electrophysiological correlates of the supramodal system, as indexed by anterior negativities and occipital positivities, appeared contralateral to the cue side, irrespective of cue modality during the cue–target interval (Eimer and van Velzen 2002). Furthermore, another EEG study demonstrated a hybrid version including the two systems: 1) separate neural correlates of the separable-but-linked system for attentional modulation of stimulus processing and 2) a supramodal system for attentional control (Eimer et al. 2002). Our findings provide the neural representation of the separable-but-linked system, as indexed by modulation of tactile and visual processing in the sylvian fissure and occipitotemporal region by spatial attention.

Attentional control is another important mechanism, which has been tested by measuring prestimulus attentional modulation. It is an interesting question whether there is a stimulus-independent attentional control system, which roughly corresponds to the pure supramodal attentional system described earlier. Eimer and colleagues reported event-related potential (ERP) modulation associated with supramodal attentional control during the cue–target interval (Eimer and van Velzen 2002; Eimer et al. 2002). In addition, frequency-specific attentional modulation (Fries et al. 2001; Lakatos et al. 2008, 2009; Siegel et al. 2008; Wyart and Tallon-Baudry 2008) might provide new insight into mechanisms of supramodal attentional control. However, since the present study used a sustained attention task where eight different stimuli were presented at a relatively short, random interstimulus interval in a random order, it would be difficult to examine prestimulus attentional modulation. Accordingly, the spatiotemporal dynamics of the neuronal activity underlying stimulus-independent attentional control remains unclear in humans. The existing literature implies two attentional systems: the supramodal system for attentional control and a separable-but-linked system for attentional modulation of sensory processing.

**Conclusion**

The present MEG study revealed the effect of within-modal and cross-modal spatial and intermodal attention on occipitotemporal and frontal cortices responsive to visual input. The effects of within-modal spatial and intermodal attention were represented at different times in distant regions of the brain including the occipitotemporal and frontal cortices, whereas the effect of cross-modal attention from touch to vision was represented at around 150 ms in the occipitotemporal cortex. Based on the present and previous reports, there is a biased effect of cross-modal spatial attention between vision and touch with vision dominant. The present findings are consistent
with the hypothesis of a separable-but-linked attentional system, but do not exclude the possibility of a stimulus-independent (or supramodal) attentional system.

ACKNOWLEDGMENTS

We thank Y. Takeshima for help in devising, constructing, and maintaining the equipment used in this study.

GRANTS

This study was supported by Ministry of Education, Culture, Sports, Science and Technology of Japan Grant-in-Aid for Young Scientists (B) 18700327 and by Japan Society for the Promotion of Science (JSPS) Fellows Science and Technology of Japan Grant-in-Aid for Young Scientists (B) 20700143.

DISCLOSURES

The authors have no conflicts of interest, financial or otherwise, to declare.

REFERENCES


