



Top down influence on visuo-tactile interaction modulates neural oscillatory responses

Noriaki Kanayama ^{a,b,*}, Luigi Tamè ^c, Hideki Ohira ^d, Francesco Pavani ^{c,e}

^a Interfaculty Initiative in Information Studies, Department of General Systems Studies, University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8902, Japan

^b Japan Society for the Promotion of Science, Tokyo, Japan

^c Centre for Mind/Brain Sciences, University of Trento, Rovereto, Italy

^d Graduate School of Environmental Studies, Department of Psychology, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

^e Department of Cognitive Sciences and Education, University of Trento, Rovereto, Italy

ARTICLE INFO

Article history:

Received 3 June 2011

Revised 7 October 2011

Accepted 25 November 2011

Available online 7 December 2011

Keywords:

Crossmodal congruency effect

Gamma band activity

Theta band activity

Gamma–theta coupling

Top-down control

Time-frequency analysis

Multisensory integration

Vision

Touch

ABSTRACT

Multisensory integration involves bottom-up as well as top-down processes. We investigated the influences of top-down control on the neural responses to multisensory stimulation using EEG recording and time-frequency analyses. Participants were stimulated at the index or thumb of the left hand, using tactile vibrators mounted on a foam cube. Simultaneously they received a visual distractor from a light emitting diode adjacent to the active vibrator (spatially congruent trial) or adjacent to the inactive vibrator (spatially incongruent trial). The task was to respond to the elevation of the tactile stimulus (upper or lower), while ignoring the simultaneous visual distractor. To manipulate top-down control on this multisensory stimulation, the proportion of spatially congruent (vs. incongruent) trials was changed across blocks. Our results reveal that the behavioral cost of responding to incongruent than congruent trials (i.e., the crossmodal congruency effect) was modulated by the proportion of congruent trials. Most importantly, the EEG gamma band response and the gamma–theta coupling were also affected by this modulation of top-down control, whereas the late theta band response related to the congruency effect was not. These findings suggest that gamma band response is more than a marker of multisensory binding, being also sensitive to the correspondence between expected and actual multisensory stimulation. By contrast, theta band response was affected by congruency but appears to be largely immune to stimulation expectancy.

© 2011 Elsevier Inc. All rights reserved.

Introduction

In the past decades, many research groups have studied the neural substrates of the multisensory integration process (e.g., see Calvert et al., 2004 for reviews). These works revealed the existence of multisensory areas in cortical as well as subcortical structures (e.g., Balslev et al., 2005; Bremmer et al., 2001; Calvert and Thesen, 2004; Macaluso and Driver, 2001; Macaluso et al., 2001). Beside the well-known role of the parietal association area as key center for multisensory integration, recent works revealed that brain regions in each sensory cortex, which were traditionally considered strictly unimodal, also play a role in multisensory integration processes (e.g., Ghazanfar et al., 2005; Kayser and Logothetis, 2007; Lakatos et al., 2007; Molholm et al., 2002; Senkowski et al., 2005; for a review, Macaluso, 2006). While these seminal works substantially expanded our understanding of the network of areas involved in multisensory integration, the timing and coupling of related neural responses remain unclear. Furthermore, the issue of how neural responses to

multisensory integration change in time remained relatively less investigated. In this perspective, one recent approach to shed light on the neural basis of multisensory integration has been to examine the oscillatory responses recorded in EEG/MEG, which have been considered as a real time reflection of cortical activity (Senkowski et al., 2008). Two main oscillatory indices that appear to be modulated by multisensory integration are the gamma band response, which is high frequency activity that has been considered as an index of high cognitive processes, and the theta band response, which is relatively lower frequency activity corresponding to the ERP components.

Using the McGurk paradigm, in which the perception of a heard syllabus is modified by the concurrent presentation of a visually incongruent spoken syllable (McGurk and MacDonald, 1976), enhanced gamma band activity was observed in response to the illusory auditory perception induced by a visual deviant within a continuous stream of multisensory audiovisual speech stimuli (Kaiser et al., 2005; 2006). Similarly, in the sound-induced visual flash illusion paradigm, in which observers typically perceive two illusory flashes when a single flash of light is accompanied by a pair of rapidly presented auditory beeps (Andersen et al., 2004; Shams et al., 2005), significantly greater gamma band response for illusory than non-illusory trials has been recorded (Bhattacharya et al., 2002; Mishra et al., 2007). Using

* Corresponding author.

E-mail address: kanayama@ardbeg.c.u-tokyo.ac.jp (N. Kanayama).

audiovisual stimuli, a modulation of the gamma band response in the symbol-to-sound matching paradigm has also been described (Widmann et al., 2007). In such a task, participants learned the correspondence between the elevation of visual stimuli and the pitch of the auditory stimuli before the experiment. During the experiment, however, the visual and the auditory patterns were congruent only in 50% of the trials. Enhanced gamma band response was observed only during congruent (compared to incongruent) trials. Widmann et al. (2007) concluded that the gamma band response might reflect formation of a unitary event representation, including both the visual and auditory aspects of the stimulation, and the match of expectation and stimulation.

Further work recently expanded these results to the case of visuo-tactile combinations (Kanayama et al., 2007; 2009), adopting the so-called crossmodal congruency effect (CCE) which has typically been studied for visuo-tactile combinations (Spence et al., 2004a). In a typical visuo-tactile congruency task, participants grasp two foam cubes using the forefingers and thumbs of either hand. Vibrotactile stimulators are located at the upper or lower corners of each cube, to stimulate the forefinger and thumb of each hand. Four light emission diodes (LEDs) are also attached to the cubes, near each tactile stimulator. In this setting, a tactile stimulus and a visual stimulus are presented in each trial, concurrently and independently from one of the four possible stimulus locations. Participants are required a speeded elevation discrimination of a target stimulus in one sensory modality, while ignoring task-irrelevant distractors presented the other modality. Typically, participants are required to decide whether vibrotactile target stimuli are presented from the forefinger or thumb of either hand (the correct response is “upper” at the forefinger, and “lower” at the thumb), while simultaneously ignoring the visual distractors. Elevation discrimination performance is typically worse when visual distractors are presented from different elevations than the tactile targets (incongruent condition), compared to when visual and tactile stimuli are presented from the same elevation (congruent condition). Using this procedure in combination with the rubber-hand illusion (Pavani et al., 2000), Kanayama et al. (2007, 2009) observed a gamma band response around 250 ms post-stimulus, which occurred only in the congruent condition and was further modulated by the illusory experiences or individual differences in susceptibility to this illusion.

As just described, many multisensory tasks invoke the gamma band response. However, the function and source of this component are currently controversial. Yuval-Greenberg et al. (2008) demonstrated that the gamma band response, which was observed in one of their previous studies on semantic integration by audio-visual stimuli (Yuval-Greenberg and Deouell, 2007), could be explained by the micro-saccadic activity. Indeed, the gamma band activity reported in many studies using visual stimulus was located at the occipital site, raising the possibility that this activity stemmed primarily from ocular activity. Although the gamma band activity during visuo-tactile integration task was located at the parietal site, the topography and source estimation analysis were not conducted (Kanayama et al., 2007, 2009). Thus, it appears evident that any study investigating gamma band responses during multisensory integration should be able to clearly disentangle this neural activity from the gamma band oscillations elicited by ocular activity. This can be done by taking advantage of independent component analysis to separate the EEG components into some brain sources, and cluster analysis for statistical testing of IC activities across participants. If the target gamma band oscillation is directly related to the crossmodal integration process, a source of the component should be located in brain areas outside the occipital cortex which have previously been linked to multisensory processing (e.g., the parietal association area).

Neural oscillations related to multisensory integration processes have also been studied by focusing on EEG oscillations at different frequency bands, and their relationship with the gamma band activity. Another oscillatory response typically recorded in response to visuo-tactile

stimulation is the theta band response, starting at 100 ms post-stimulus. This component can be separated into two sub-components with different time-courses: an early theta component (100–300 ms), which has suggested to reflect the response to visuo-tactile stimulation regardless of congruency, and a late theta component (300–500 ms), which is most apparent under incongruent conditions, including the visuo-tactile conflict elicited by the CCE, and may thus reflect the cognitive costs of responding to incongruent visuo-tactile pair (Kanayama and Ohira, 2009). Recent studies also investigated the coupling between gamma band power increase and theta band phase index, and suggested a role of gamma/theta coupling on multisensory processing (Schroeder and Lakatos, 2009). Canolty et al. (2006) reported the coupling between theta and gamma oscillations in human brain measuring EEG by electrocorticogram (ECoG). The location of the theta gamma coupling was also modulated by the task (passive listening to predictable tones, auditory n-back task and many other cognitive tasks). Monto et al. (2008) recorded scalp EEG activity during somatosensory detection, and observed a correlation between phase index of very slow oscillation and amplitude index of oscillation in high frequency band. Finally, using audiovisual stimuli and electrophysiological recording in the macaque, Lakatos et al. (2008) demonstrated that the cross-sensory attentional shift has an impact on the cross-frequency coupling. These findings lead to the conclusion that the gamma band activity, which was reported in several multisensory integration studies, should be considered in combination with activity in the lower frequency bands.

In the present study, we aimed to examine to what extent these two oscillatory components and the coupling between them is sensitive to top-down control, or instead just sensitive to the multisensory aspect of the stimulation. Several recent studies examined modulation of CCE from top-down signals. For instance, Aspell et al. (2009) demonstrated the impact of the illusory perceptions concerning the whole body on CCE, stimulating the participant's back and using head mounted display (HMD) to show both the visual distractors and several viewing conditions of the body. Light and vibration devices were attached to the skin of their backs using tape. The positions of the devices were upper right and left (at the inner edges of the shoulder blades), lower right and left (9 cm below the upper positions). During the experiment, participants saw their backs through a HMD, but were asked to ignore the visual distractors and make instead elevation judgment of the tactile stimulus. A significantly larger CCE emerged in the body visible condition as compared to the body invisible condition (when only distractors were visible). In another condition, task irrelevant stroking was applied to the participant's body during the elevation judgment, which was spatially synchronous or asynchronous with respect to the stroking seen through HMD. The synchronous stroking condition elicited a strong bodily illusion (similar to the like the rubber hand illusion, but experienced here for the whole body), with respect to the asynchronous stroking condition. A modulation of synchrony of the task irrelevant stroking on CCE emerged, when the visual stimulus preceded the tactile one by 233 ms. Another example in which the context can modulate the strength of the CCE effect, presumably reflecting top-down modulation of multisensory processing, is a recent study by Heed and colleagues (2010). This work demonstrated that CCE can be modulated when the participant performs the CCE task while another person sits close to him/her and perform a task on the visual distractors. When the partner sat close to the participant and responded to the light stimulus, CCE decreased compared to the condition without partner. Remarkably, this CCE decrement did not emerge if the partner sat farther away or sat close but performed no task on the visual distractors. This demonstrated that the participant's performance was modulated by others' actions in a top down fashion. Other examples that CCE can be modulated in a top-down fashion have been provided (Austen et al., 2004; Pavani and Castiello, 2004; Zopf et al., 2010).

To clarify the effect only from top-down information, in the present study we adopted a variation of the visuo-tactile congruency task in which top-down influences on multisensory integration are

obtained by changing the relative proportion of congruent and incongruent trials (Shore and Simic, 2005). If an incongruent trial is frequently presented, participants can develop an expectancy that the visual stimulation is mostly at a different elevation with respect to the tactile stimulation and that they should adopt a strategy of always ignoring it. By modulating the probability of spatially congruent visuotactile stimulation, we could investigate the influences of the knowledge established by the previous experiences, as a top-down control, on the multisensory integration process. We predicted that the gamma band response, would reflect the highly cognitive visuotactile integration process, and would be observed in sensory association areas in the parietal cortex, rather than the frontal pole (ocular movement) or the occipital cortex. Also, we predict that the gamma band response would be modulated by the proportion of congruent trials, because its activity is influenced by the expectancy as a spontaneous top-down control on visuotactile processing. Instead, we predicted that the theta band activities would not be modulated by the proportion of congruent trials, because the components have merely the function of comparing the inputs from two different modalities. Finally, we expected that the coupling between these oscillations would be a factor modulating the gamma band activity.

Materials and methods

Participants

Twelve volunteer students (five women; age range: 25–33 years) participated in the experiment. All participants were right-handed, reported normal or corrected vision and normal somatosensation, and did not report any neurological or psychiatric problems. Informed consent was obtained from each participant prior to the experiment.

Stimuli and apparatus

The participants sat in a dimly lit and sound-attenuated experimental chamber, shielded to avoid line frequency interference and ambient noise. The participants grasped an expanded polystyrene cube (7 cm × 10 cm × 1 cm) with the thumb and forefinger of their left hand. Two vibrotactile stimulators (BC461-1: Oticon A/S) and two red light-emitting diodes (LEDs) were attached to the corners of the cube, as shown in Fig. 1a. Vibrotactile stimulators served the purpose of delivering tactile targets to the participant's thumb or

forefinger; red LEDs served the purpose of delivering visual distractors. At the mid-point between two red LEDs, a green LED was attached to serve as visual fixation throughout the task. The visual stimuli were applied for 50 ms by delivering a 5 V pulse via a parallel port output from a control computer. The somatosensory stimuli were applied for 50 ms by activating the vibrotactile stimulators via a sound card (Sinusoidal wave applied with a frequency of 100 Hz). An adjustable armrest was placed on the table in order to minimize the movement of the participants' left hand during the tasks. White noise was continuously presented to the participants via headphones to mask any noise generated by the operation of the vibrators.

Experimental design

The experimental design included two within-participants factors: congruency (congruent, incongruent) and proportion of congruent trials (congruent 80%: CON80, congruent 20%: CON20). There were four combinations of the presented stimuli: visual upper/tactile upper and visual lower/tactile lower were considered as congruent condition, visual upper/tactile lower and visual lower/tactile upper were considered as incongruent condition. In addition, somatosensory only conditions were added to the design (i.e., no visual stimulus/tactile upper and no visual stimulus/tactile lower). These conditions were used for the intermission task (see procedure) and were not included in the analysis. These unimodal filler blocks were adopted in order to avoid the aftereffect of the proportion of congruent trials that the participant experienced in the previous block.

Procedure

The experiment comprised eight experimental blocks. Two experimental blocks included 120 congruent trials and 30 incongruent trials (congruent 80%, incongruent 20%). In another two blocks, the proportion of the congruent trials was reversed (congruent 20%, incongruent 80%). Finally, each of the experimental blocks was preceded by one block with 80 unimodal somatosensory trials which served as fillers. The presentation order of congruency in the experimental blocks and of target location in the filler blocks was entirely randomized and unpredictable. The order of the experimental blocks was counterbalanced across participants.

At the beginning of each block the red LEDs attached to the cube were switched on for 4 s, to inform participants to prepare for the task. Each trial began with 350–450 ms of no-stimulation, followed by the simultaneous presentation of visual and tactile stimuli for 50 ms. After stimulus delivery the trial continued for 1450 ms without any stimulation, waiting for the participants' response. Regardless of whether participants delivered response, the next trial started without any feedback. In the intertrial interval both red LEDs were switched on for 50–150 ms, which inform the participants about permission to blink only one time. During any other periods, participants were asked to minimize blink. Participants were instructed to indicate the location of the tactile stimulation (upper or lower) by pressing the appropriate key using the thumb and index finger of their right hand. For upper stimulation, participants pressed a key by the index finger, for lower stimulation participants pressed a key by the thumb. To avoid any incongruency between the response finger and the stimulated finger, this response mapping was not counterbalanced across participants. Speeded and accurate responses were required. A short break was provided after each block.

Data analyses

Analysis of behavioral results

Response times (RT) and percent correct data were analyzed using a repeated measure ANOVA with two within-participants factor (congruency: congruent, incongruent; proportion congruent: Congruent

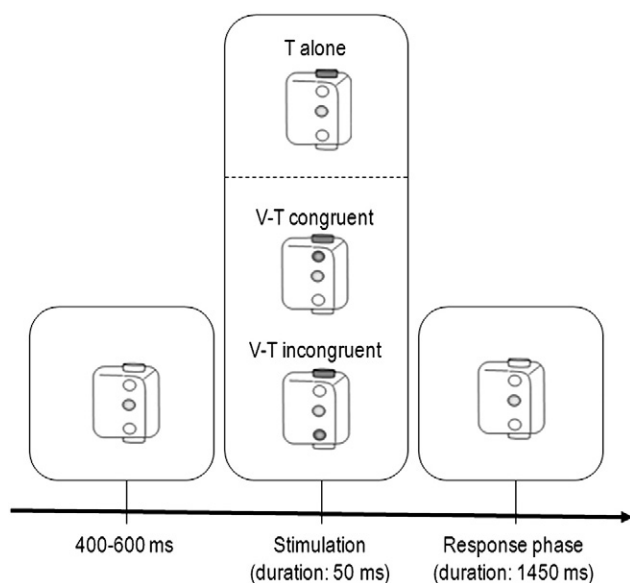


Fig. 1. The temporal sequence of the single trial. The circles indicate LEDs, and the rectangles indicate vibrators. The grey color implies "on", the white color implies "off."

80%, Congruent 20%). Trials with too fast key press (<300 ms) were regarded as an error trial and discarded from the further analyses. The visuotactile congruency effect (CCE) was calculated by subtracting congruent from incongruent condition for both RT and percent correct, and was analyzed using *t*-test between Congruent 80% and Congruent 20% blocks. Also, we conducted *t*-tests to investigate significant differences between somatosensory unimodal condition and the other conditions. The Greenhouse–Geisser method was employed for the correction of the degrees of freedom in the ANOVA, and post hoc analyses were conducted using Bonferroni adjustments of the *p*-value.

EEG recording and data analysis

EEG was continuously recorded at a sampling rate of 500 Hz from 62 Ag–AgCl electrodes referenced to the nose tip (Brain Vision; Brain Products, Germany). The electrodes were placed on Fpz, Fp1, Fp2, AF3, AF4, AF7, AF8, Fz, F1, F2, F3, F4, F5, F6, F7, F8, FCz, FC1, FC2, FC3, FC4, FC5, FC6, FT7, FT8, Cz, C1, C2, C3, C4, C5, C6, T7, T8, CPz, CP1, CP2, CP3, CP4, CP5, CP6, TP7, TP8, TP9, TP10, Pz, P1, P3, P2, P4, P5, P6, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, O2, according to the international 10–10 system. The impedance at each electrode was maintained below 5 k Ω . Eye movement was monitored from the outer canthus of the right and left eyes. The amplifier bandpass was 0.1–120 Hz, and no additional filters were applied. Trials with incorrect task responses were excluded from the behavioral and EEG analyses.

EEG data analyses were conducted using EEGLAB 7.13.11b (Delorme and Makeig, 2004) under Matlab platform (version 7.4; MathWorks, Natick, MA, USA). An average reference was calculated and applied. EEG data was segmented into epochs from 600 ms pre-stimulus to 1200 ms post-stimulus with 901 data points. All epochs were visually inspected, and epochs with extreme irregular noise were discarded. Physiological artifacts, such as eye blinks and electromyographic responses, were not excluded for the ICA process. The extended Infomax ICA with natural gradient (implemented in EEGLAB) was performed to obtain 62 Independent components (ICs) from each of 12 participants. Total number of IC was thus 744 (12 participants \times 62 ICs). Visual inspection was performed on the IC waveforms for each subject, to exclude data with improbable activity. In order to control signal/noise (SN) ratio, trials were trimmed to the trial number of smallest condition (incongruent trials in CON80 blocks in most cases) by random selection for each participant. Mean trial number included in later analyses was 37.75 (SD = 5.95). A second ICA was performed using the remained data. For each IC, an equivalent current dipole was estimated using DIPFIT 2.2 (EEGLAB plug-in using Fieldtrip toolbox functions). The Montreal Neurological Institute (MNI) standard coordinate system was used for the head model. ICs located outside of the brain area were discarded (112 ICs). Event-related spectral perturbation (ERSP), inter-trial phase coherency (ITC), and EEG spectra were calculated for each of the remaining 632 ICs. For ERSP and ITC calculation, the Morlet wavelet was used. Log-spaced 96 frequencies ranging from 3 Hz to 100 Hz were calculated every 8 ms starting from 300 ms prior to and ending up to 800 ms following stimulus onset (baseline from –300 ms to –50 ms) with a linearly increasing wavelet cycle of 1 at 3 Hz and 7 at 100 Hz. Data dimensions of ERSP and ITC were reduced to 10 with principle component analysis, and IC clustering was performed using *k*-means with the criteria of dipole locations (normalization weight, 8), ITC (latency, for 0–800 ms; frequency, 3–100 Hz; normalization weight, 1), and ERSP (latency, 0–800 ms; frequency, 3–100 Hz; normalization weight, 5) to generate 15 IC clusters. 40 ICs, which did not fall within 3 SD from any cluster, were excluded as outlier components. Finally, 592 ICs were used for statistical test. Out of 15 clusters, target clusters were selected based on the criterion that at least 10 participants were included in the cluster.

For statistical test on ERSP and ITC, nonparametric permutation with pseudo *F*-test was performed to compare levels of five conditions. First, *F*-values of condition effect were computed for the each time-frequency point. Also same *F*-values were computed for shuffled data across ICs and conditions (for example, for cluster 9, 62 ICs \times 5 conditions = 310 data were shuffled). The *F*-values on the shuffled data were repeatedly calculated 10,000 times and sorted in ascending order. The *p*-value was calculated by the rank of the original *F*-values in the shuffled *F*-value distribution (If the original *F*-value is 500th from the end of the distribution, *p*-value is 500/10,000 = 0.05). Also the *p*-value was corrected by false-discovery rate (FDR) method. All time frequency points with no significant *p*-value were masked. For the time-frequency windows, which showed significant difference between five conditions, post-hoc multiple comparison using Wilcoxon signed rank test was performed on averaged or peak values across all time and frequency points. The multiple comparisons were performed for all possible combination of conditions, and *p*-values were corrected by FDR method.

The cross-frequency coupling was computed only for cluster 9, which shows significant condition differences on gamma band ERSP and theta band ITC. ERSP data was sorted by a phase of theta band oscillation in a target window for each participant and each trial. 32 phase windows from $-\pi$ to π was set. All ERSP values (96 frequencies \times 32 phases) were subtracted by the value averaged across all phase point. The positive values of this index indicate increase in a certain phase of theta oscillation compared to the averaged value. This values also statistically tested by the above mentioned nonparametric permutation test, and Wilcoxon signed rank test.

Results

Behavioral results

Mean RTs, percent of correct responses, and the congruency effects in each condition are shown in Table 1. The two-factor ANOVA on RTs demonstrated a significant main effect of congruency, $F(1, 11) = 46.14$, $p < .001$, partial eta squared = .81, and a significant interaction between congruency and proportion congruent, $F(1, 11) = 6.18$, $p < .05$, partial eta squared = .36. Overall, RTs were faster in the congruent compared to the incongruent condition (537 ms vs. 648 ms). This difference (crossmodal congruency effect in RT) was significantly modulated by proportion congruent (Table 1). Post hoc analysis of the interaction revealed that responses to congruent stimuli were faster in the 80% congruent block compared to the 20% congruent block, $t(11) = -2.35$, $p < .05$. By contrast, response speed in incongruent trials was not significantly modulated by proportion congruent, $t(11) = .96$, *n.s.* This resulted in larger visuotactile congruency effect (CCE) in the 80%

Table 1

Mean and standard error (in parenthesis) of reaction times (ms) and percent correct (%) as a function of congruency (congruent or incongruent) and proportion of congruent stimulations (80% or 20%). Congruent indicates the congruent trials, and incongruent indicates the incongruent trials. Congruency effect is the differences in performance between congruent and incongruent trials. Congruent 80% indicates the blocks in which 80% of the trials were congruent and the remaining were incongruent; Congruent 20% indicates the blocks in which 20% of the trials were congruent trials and the remaining were incongruent. The somatosensory condition is the one in which tactile targets were presented alone (i.e., without visual distractors).

		Congruent 80%	Congruent 20%
Congruent	RT	515 (39)	559 (37)
	Percent correct	93.2 (2.6)	93.8 (1.9)
Incongruent	RT	658 (41)	638 (40)
	Percent correct	67.8 (4.9)	83.4 (3.3)
Congruency effect	RT	143 (27)	79 (11)
	Percent correct	25.5 (4.8)	10.4 (3.3)
Somatosensory	RT	501 (28)	
	Percent correct	94.3 (1.6)	

congruent block compared to in the 20% congruent block ($t(11) = 2.49$, $p < .05$).

A similar analysis on percent correct revealed the main effects of congruency, $F(1, 11) = 22.45$, $p < .01$, partial eta squared = .67, and proportion congruent, $F(1, 11) = 12.56$, $p < .01$, partial eta squared = .53. Again, there was also a significant interaction between congruency and proportion congruent, $F(1, 11) = 22.15$, $p < .01$, partial eta squared = .67. Overall, percent correct was higher in congruent than incongruent trials (93.5% vs. 75.6%). This difference (crossmodal congruency effect in percent correct) was significantly modulated by proportion congruent (Table 1). Post-hoc analyses of the interaction revealed that percent correct in incongruent trials was significantly smaller in the 80% congruent block compared to the 20% congruent block, $t(11) = -4.84$, $p < .01$. By contrast, percent correct in congruent trials was not modulated by proportion congruent, $t(11) = 0.23$ n.s. This resulted in larger visuotactile congruency effect in the 80% congruent block compared to the 20% congruent block, $t(11) = 4.71$, $p < .01$. These results closely parallel those observed for the RT data, thus excluding speed-accuracy trade-offs.

Finally, we compared the unimodal somatosensory condition with the other experimental conditions using *t*-tests. For RTs, the unimodal somatosensory condition was significantly faster than all other stimulation conditions (incongruent stimulation in the 80% congruent block, $t(11) = 5.98$, $p < .001$; congruent stimulation in the 20% congruent block, $t(11) = 3.74$, $p < .01$; incongruent stimulation in the 20% congruent block, $t(11) = 6.24$, $p < .001$), except the congruent stimulation in the 80% congruent block. For percent correct, a significant difference emerged only between the unimodal somatosensory condition and the incongruent stimulation conditions in the 80% congruent block ($t(11) = 5.86$, $p < .001$), or the 20% congruent block ($t(11) = 4.38$, $p < .001$). All unreported differences were not significant.

EEG results

The clustering analysis using 632 ICs revealed 16 clusters. Cluster number 1 was excluded because it was a parent cluster including all components. The results of dipole estimation for all of the remaining clusters are shown in Table 2.

Among the remaining clusters were operated a further selection based on two criteria. First, we reasoned that all clusters could include EMG, EOG, and other artifact-like effects. Because the artifacts likely show large individual difference, artifact cluster should include, in principle, a relatively low number of participants. Accordingly, we set a minimum acceptable number of participants included in each cluster to 10 (out of the 12 participants tested in the study). This first criterion excluded clusters 2, 3, 5, 7, 8, 10, 14, and 15. Second, we conducted permutation tests and selected those clusters in which a significant difference between the experimental conditions emerged for either of the following measures: the gamma band ERSP or ITC (30–100 Hz), theta band ERSP or ITC (3–8 Hz). This second criterion excluded clusters 4, 6, 12, and 13. In the following paragraph, we present the results in details for clusters 9, 11, and 16, which we also identify based on the estimated dipole coordinates.

Right superior parietal lobule (cluster 9)

The averaged dipole coordinates of cluster 9 are shown in Fig. 2 (top part) and were located at the right superior parietal lobule (BA7). Permutation tests revealed significant effect of condition for gamma band ERSP and theta band ERSP and ITC. Within this cluster, the gamma band ERSP response (70–90 Hz, around 250 ms poststimulus) was clearly modulated as a function of stimulation congruency and proportion of congruent trials (see bar plot on the lower left in Fig. 2). By contrast, the theta band ITC (3–8 Hz, around 100 ms poststimulus) was only modulated as function of congruency.

For the gamma band ERSP, we selected 3 frequency and time windows based on the significant ERSP map, 30–40 Hz 150–180 ms, 70–90 Hz 250–280 ms, and 30–50 Hz 600–800 ms. Post-hoc Wilcoxon

Table 2

The MNI coordinates of the averaged dipole location for all clusters. The gray colored rows identify the target clusters, based on the criterion that at least 10 subjects were included in this cluster. BA: Brodmann Area. Subjects and ICs: the number of subjects and IC included in the cluster, respectively. STG: Superior Temporal Gyrus, MFG: Middle Frontal Gyrus, PCG: Post Central Gyrus, MTG: Middle Temporal Gyrus, IFG: Inferior Frontal Gyrus, SPL: Superior Parietal Lobe, SFG: Superior Frontal Gyrus, MOG: Middle Occipital Gyrus. “r” and “l” preceding the abbreviation of the brain area identifies the right or left hemisphere, respectively.

Cls	x	y	z	BA	Subjects	ICs	Name
2	32	4	-41	38	9	19	rSTG
3	-70	-34	12	22	9	50	ISTG
4	-42	54	4	10	10	86	lMFG
5	-64	-13	21	43	8	32	lPCG
6	-17	-35	65	3	10	33	lPCG
7	63	-57	2	21	9	31	rMTG
8	59	6	25	9	9	38	rIFG
9	13	-62	62	7	10	62	rSPL
10	-66	9	-5	22	6	36	ISTG
11	21	64	5	10	10	40	rSFG
12	-6	32	57	6	10	38	lSFG
13	27	-95	12	18	11	27	rMOG
14	-72	-46	-18	-	8	57	
15	-27	69	1	10	7	8	lSFG
16	-36	-84	18	19	10	35	lMOG

signed rank tests revealed that there is no significant effect on the averaged values of all frequency and time windows. Peak values differed between congruent and incongruent trials only for the 70–90 Hz 250–280 ms time-frequency window, irrespective of proportion congruent (80% congruent, $z(9) = 2.80$, $p < .05$; 20% congruent, $z(9) = 2.29$, $p < .05$). Remarkably, however, which stimulation condition led to the highest peak value was reversed as a function of proportion congruent (compare black and dark gray bars in the Con80% and Con20% conditions in the Gamma ERSP bar plot at the bottom left part in Fig. 2). Specifically, the peak values for the congruent stimulation condition decreased significantly when the proportion of congruent trials was 20% compared to when the proportion of congruent trials was 80% ($z(9) = 2.70$, $p < .05$). Finally, the somatosensory only stimulation was significantly different with respect to the congruent condition in the 80% congruent block ($z(9) = 2.60$, $p < .05$), and the incongruent condition in the 20% congruent block ($z(9) = 2.50$, $p < .05$).

The theta band ERSP was not modulated by the proportion of congruent trials. We selected one time-frequency window based on the significant ERSP map, 3–5 Hz 100–150 ms. Post-hoc tests revealed no significant effect either on averaged or peak values of the time-frequency window. For the theta band ITC, we selected one time frequency window based on the significant ITC map, 3–8 Hz 100–300 ms (see Theta ITC bar plot in Fig. 2). Post-hoc tests revealed that the somatosensory only stimulation differed from all multisensory stimulation conditions (all *p* values < 0.05).

Right superior frontal gyrus (cluster 11)

The averaged dipole coordinates of cluster 11 are shown in Fig. 3 (top part) and were located at the right superior frontal gyrus (BA10). Permutation tests revealed that significant effects of condition were observed for gamma band ERSP and theta band ITC. As illustrated in the bottom part of Fig. 3, ERSP in the gamma band (> 30 Hz) was larger in congruent compared to incongruent trial in the 80% congruent block.

For the gamma band ERSP, we selected two time-frequency windows based on the significant ERSP map, 50–70 Hz 280–300 ms and 30–60 Hz 550–600 ms. Post-hoc Wilcoxon signed rank tests on averaged values revealed significant differences between congruent and incongruent stimulations only in the 80% congruent block, for both time-frequency window (50–70 Hz 280–300 ms: $z(9) = 2.80$, $p < .05$; 30–60 Hz 550–600 ms: $z(9) = 2.60$, $p < .05$). No other difference reached significance (see gamma ERSP bar plots in Fig. 3).

Right Superior Parietal Lobule Cluster

(Cluster 9)

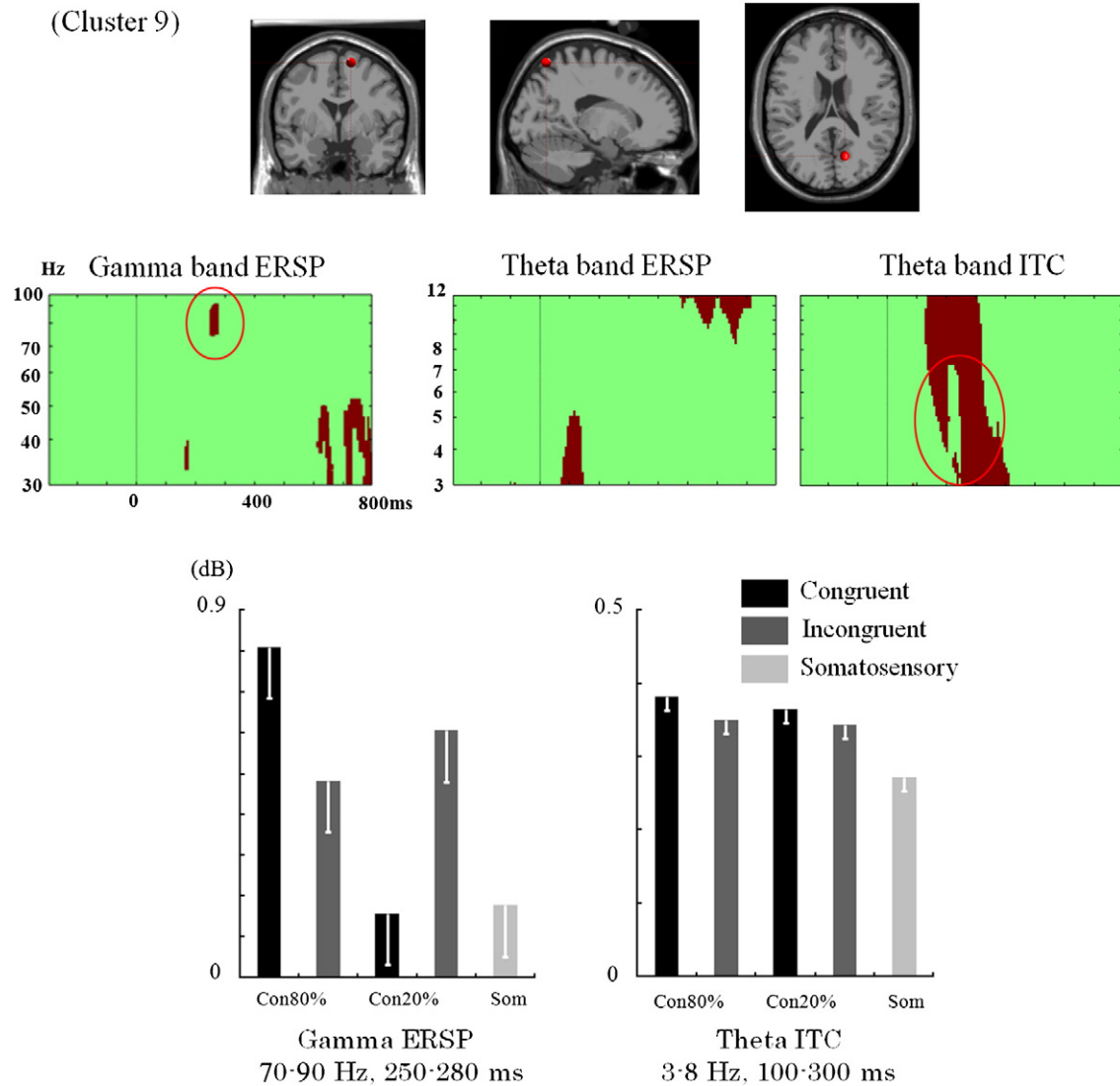


Fig. 2. Averaged dipole location (top) and ERSP and ITC maps of significant condition difference (middle, $p < .05$) for the right superior parietal lobule cluster (Cluster 9). Red colored pixels indicate the significant effect of condition. Also, the bottom bar plots indicate peak ERSP values of gamma band activity (70–90 Hz, 250–280 ms) for each condition (bottom left), and averaged ITC values of theta band activity (3–8 Hz, 100–300 ms) for each condition (bottom right). For the ERSP bar graph, y-axes represent increment of spectrum power compared to baseline (dB). For the ITC bar graph, y-axes represent the index of phase-coherency across trials (1: totally coherent, 0: totally random). Black bar: congruent condition. Gray bar: incongruent condition, Light gray bar: unimodal somatosensory condition. Error bars indicates standard errors.

For the theta band (3–8 Hz) in the 100–300 ms time window, the ITC decrease was detected only in the somatosensory condition (see Theta bar plot in Fig. 3). The somatosensory only condition differed with respect to the congruent condition in the 80% congruent block ($z(9) = 2.80$, $p < .05$), the incongruent condition in the 80% congruent block ($z(9) = 2.60$, $p < .05$), and the incongruent condition in the 80% congruent block vs. somatosensory block ($z(9) = 2.70$, $p < .05$).

Left middle occipital gyrus (cluster 16)

The averaged dipole coordinates of cluster 16 are shown in Fig. 4 (top part) and were located in the left middle occipital gyrus (BA19). Permutation test revealed that significant effect of conditions was observed only for theta band ERSP. In the early time window (50–150 ms), the ERSP decrease of theta band oscillation was observed only in the congruent condition when proportion congruent was 80% (Fig. 4, bottom, left). In the later time window (300–700 ms), the decrease of theta band ERSP in congruent trials occurred regardless of the proportion of congruent trials (Fig. 4, bottom, right).

For the theta band ERSP, we selected two time-frequency windows based on the significant ERSP map, 3–8 Hz 50–150 ms, 3–8 Hz 300–700 ms. For the 3–8 Hz 50–150 ms time-frequency window, post-hoc tests confirmed that the response to congruent stimulation in the 80% congruent block was significantly lower with respect to all other multisensory stimulation conditions (all p values $< .05$; see bar plot on the bottom left in Fig. 4). For the late time-frequency window (3–8 Hz 50–150 ms), a significant difference between congruent and incongruent stimulation emerged regardless of the proportion of congruent trials (80% congruent: $z(9) = 2.80$, $p < .05$; 20% congruent: $z(9) = 2.50$, $p < .05$). As illustrated in Fig. 4, ERSP increases in the incongruent condition are remarkable compared to the congruent condition regardless of probability.

Cross frequency coupling

In the right superior parietal lobule cluster (cluster 9), the significant effects of condition were observed on both of the ERSP values in the gamma band and the ITC (phase) values in the theta band. We

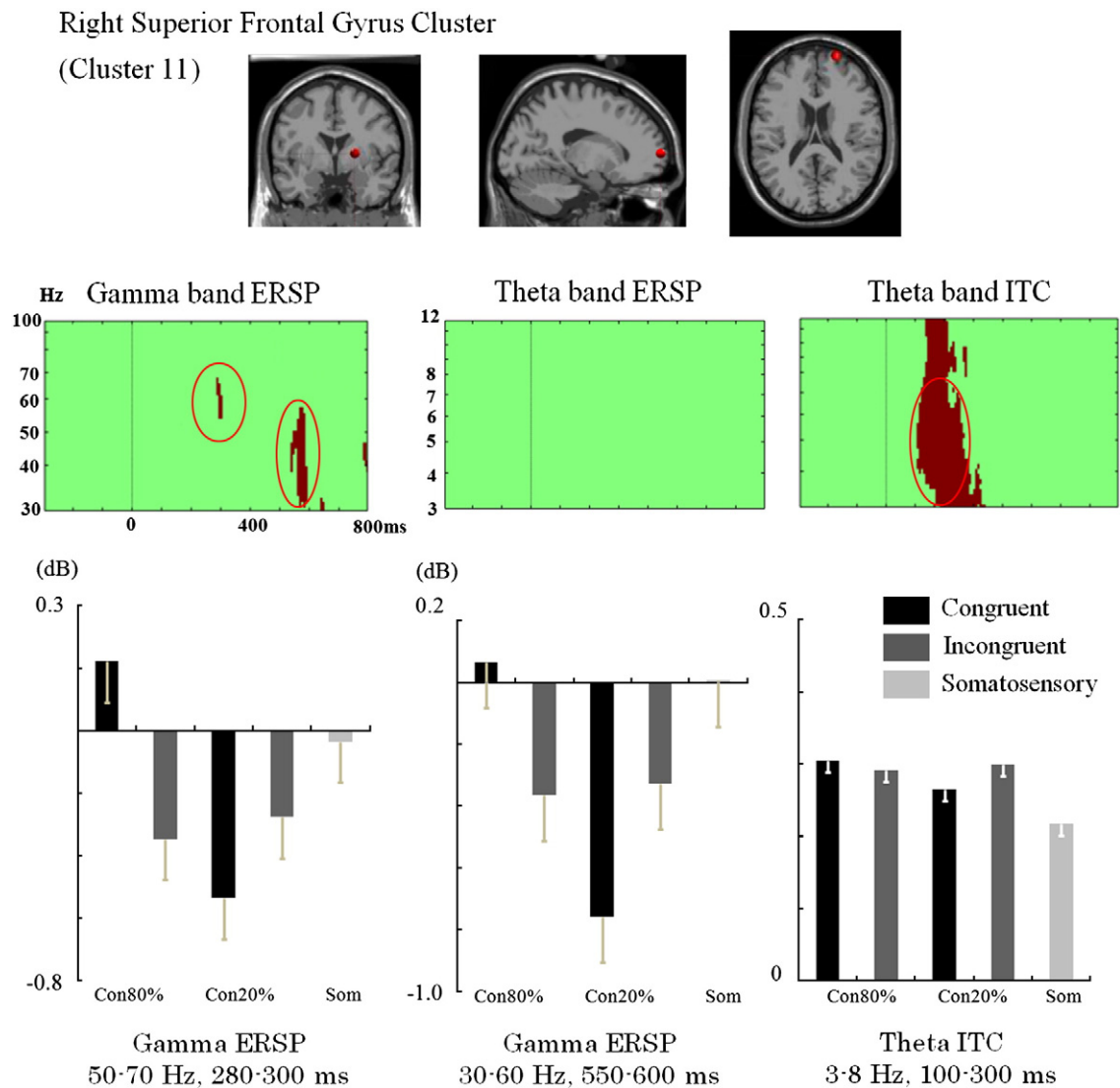


Fig. 3. Averaged dipole location (top) and ERSP and ITC maps of significant condition difference (middle, $p < .05$) for the light superior frontal gyrus cluster (Cluster 11). Red colored pixels indicate the significant effect of condition. Also, the bottom bar graphs indicate averaged ERSP values of gamma band activity (50–70 Hz, 280–300 ms) for each condition (bottom, left), averaged ERSP values of gamma band activity (30–60 Hz, 550–600 ms) for each condition (bottom, center), and averaged ITC values of theta band activity (3–8 Hz, 100–300 ms) for each condition (bottom, right). For the ERSP bar graphs, y-axes represent increment/decrement of spectrum power compared to baseline (dB). For the ITC bar graph, y-axes represent the index of phase-coherency among trials (1: totally coherent, 0: totally random). Gray bar: incongruent condition, Light gray bar: unimodal somatosensory condition. Error bars indicates standard errors.

analyzed the cross frequency coupling, using ERSP values in the gamma band rearranged by the phase of the theta band activity.

Fig. 5 revealed that the significant difference of ERSP was observed in the phase-frequency map. We selected 3 phase windows for post-hoc tests, around $-\pi/2$ (-2 to -1 radian), around $-\pi/4$ (-1 to 0 radian), around $\pi/2$ (1 to 2 radian). For the around $-\pi/2$ phase, significant differences emerged for congruent trials in the 80% congruent block compared to the 20% congruent block ($z(9) = 2.70$, $p < .05$), and between congruent and incongruent trials in the 20% congruent block ($z(9) = 2.80$, $p < .05$). Furthermore, significant differences emerged between congruent trial in the 80% congruent block vs. somatosensory trials ($z(9) = 2.80$, $p < .05$), and incongruent trial in the 20% congruent block vs. somatosensory trials ($z(9) = 2.70$, $p < .05$). For the around $-\pi/4$ phase, no significant effect was found by the post-hoc tests, despite the numerical differences in the average values. For the around $\pi/2$ phase, there was a significant difference between congruent and incongruent trials in the 80% congruent block ($z(9) = 2.80$, $p < .05$). Furthermore, the congruent trial in the 80% congruent block differed also with respect to congruent trial in the 20% congruent block ($z(9) = 2.80$, $p < .05$), incongruent trial in

the 20% congruent block ($z(9) = 2.80$, $p < .05$), and somatosensory trials ($z(9) = 2.40$, $p < .05$).

Discussion

In this study, we aimed to investigate the neural correlates of the top-down modulations on the visuotactile congruency effect. We measured RTs, percent of correct responses, and EEG activity during a vertical discrimination task of tactile sensations at the fingers, while changing the proportion of the trials in which a concurrent visual distractor occurred at congruent or incongruent locations with respect to the tactile targets. Four main findings emerged: first, the proportion of congruent trials in the block modulated the visuotactile congruency effect measured in behavior; second, the power of gamma band activity was strongly modulated as a function of proportion of congruency; third, the late theta band activity related to the congruency effect was largely insensitive to expectancy of congruency; and fourth, cross-frequency coupling was modulated by the proportion of congruent trials.

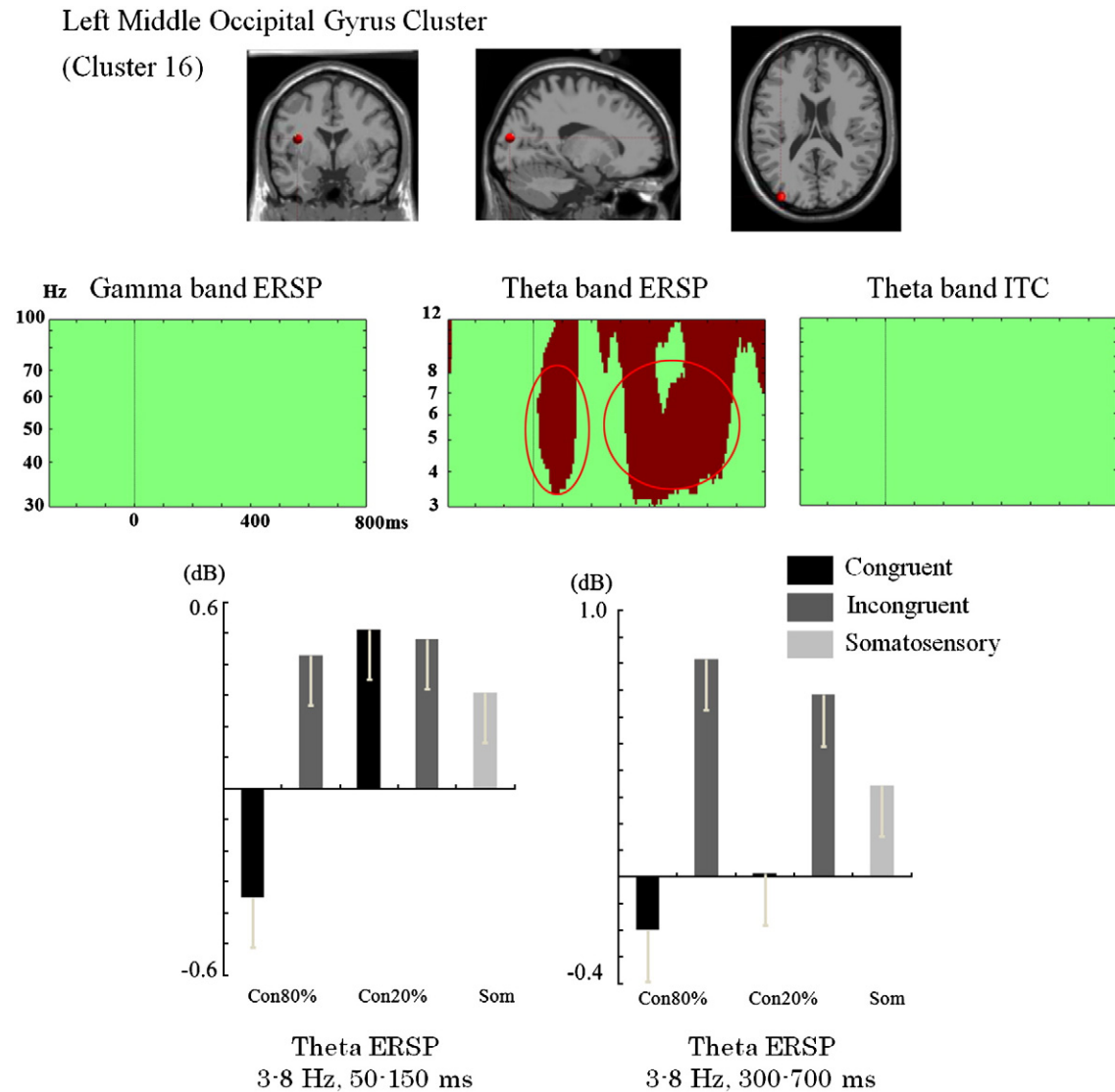


Fig. 4. Averaged dipole location (top) and ERSP and ITC maps of significant condition difference (middle, $p < .05$) for the left middle occipital gyrus cluster (Cluster 16). Red colored pixels indicate the significant effect of condition. Also, the bottom bar graphs indicate averaged ERSP values of theta band activity (3–8 Hz, 50–150 ms) for each condition (bottom, left), and averaged ERSP values of theta band activity (3–8 Hz, 300–700 ms) for each condition (bottom, right). For the ERSP bar graphs, y-axes represent increment/decrement of spectrum power compared to baseline (dB). Gray bar: incongruent condition, Light gray bar: unimodal somatosensory condition. Error bars indicates standard errors.

Top-down modulation on behavioral results

The behavioral data confirmed that the visuotactile congruency effect is strongly elicited by this visuotactile congruency paradigm. Many previous studies (Austen et al., 2004; Brozzoli et al., 2009; Heed et al., 2010; Spence et al., 2004a; Zopf et al., 2010; for a review, see Spence et al., 2004b) documented this multisensory integration effect using the visuotactile congruency paradigm, and our study fully replicated this basic finding. Most important, we found that the visuotactile congruency effect was modulated by the proportion of congruent trials in the block. Visuotactile congruency effects in RTs and percent correct were larger in the Congruent 80% block compared to the Congruent 20% block, which indicates that high-frequency presentation of the incongruent trials establishes a strong expectancy of the frequent spatial conflict between vision and touch and, in turn, this changes the way in which the multisensory stimuli are processed. One possibility is that participants implemented more effective inhibition strategies when the block comprised a large proportion of spatially incongruent distractors. Clearly participants could not predict where the next target was going to appear (upper or lower location). However, they could surely predict that in most cases, during the 20% congruent block, the tactile

and visual stimuli would not appear in the same location in space. This likely generated an expectancy of conflict that, in turn, could have produced more effective distractor inhibition.

Before discussing further this interpretation of the effect, an alternative account should be considered. Namely, the possibility that in 20% congruent blocks participants merely adopted the strategy to produce a response that was always opposite to the elevation distracting LED. Such a response strategy would have surely produced a high percent correct in the frequent incongruent trials. However, it would have also decreased percent correct in the infrequent congruent trials. In particular, the percentage of correct responses to the congruent trials in the congruent 20% block should have been worse compared to that observed for congruent trials in the congruent 80% block. Our results clearly show that this was not the case. Percent correct in response to congruent stimulations did not differ as a function of proportion congruent (93.2% vs 93.8%). Based on this result, we can suggest that the participants have not employed the strategy to select the opposite elevation to the visual information on the majority of trials.

One previous study (Shore and Simic, 2005) had previously documented limited or negligible effects of the proportion of congruent

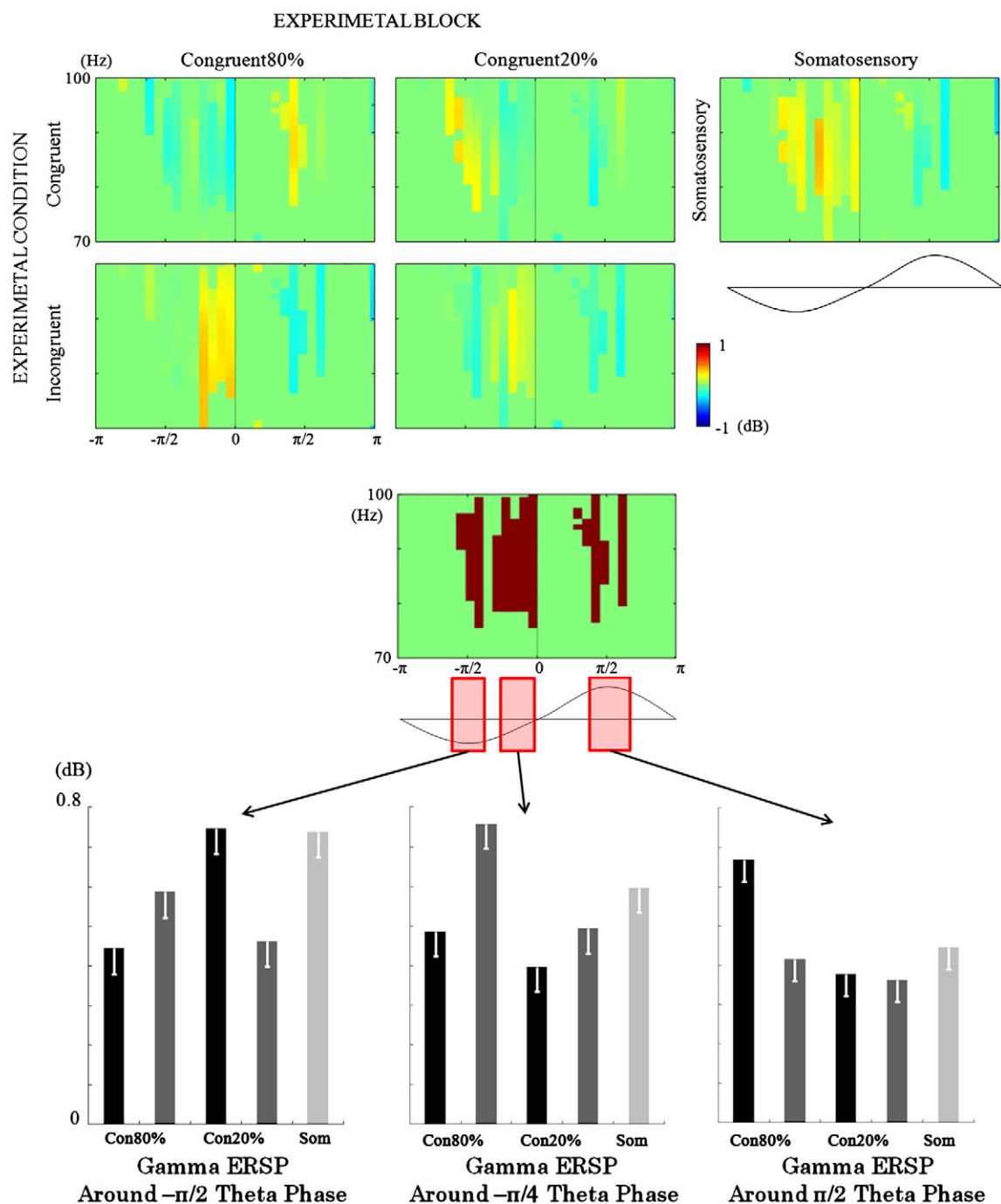


Fig. 5. The ERSP maps of gamma frequency oscillation (70–100 Hz), classified by the phase of the theta frequency oscillation (3–8 Hz) for Cluster 9 (top). The maps in the top line indicate the ERSP for each condition. X-axes indicate the phase of the Theta band oscillation ($-\pi$ to π), while y-axes indicate the frequencies (70–100 Hz). Color indicates increment/decrement of spectrum power compared to baseline (dB). The plotted ERSP values were subtracted by the value averaged across all phase point. The values indicate relative increment/decrement in a certain phase of theta oscillation compared to the averaged value. The map in the middle line is a statistical map. Red colored pixels indicate the significant effect of condition. Also, the bottom bar graphs indicate averaged ERSP values of gamma band activity in the specific theta phase (left: around $-\pi/2$ (-2 to -1 radian), center: around $-\pi/4$ (-1 to 0 radian), right: around $\pi/2$ (1 to 2 radian)) for each condition. Gray bar: incongruent condition, Light gray bar: unimodal somatosensory condition. Error bars indicate standard errors.

trials across blocks on the visuo-tactile congruency effect. Specifically, Shore and Simic (2005) showed no significant difference between blocks of high and low proportion of congruent trials (75% versus 25%; Experiment 1), despite a trend toward a proportion difference in CCE. This finding was confirmed also when a larger discrepancy in the proportion of congruent trials between blocks was introduced (89% versus 11%; Experiment 2). However, a final experiment (Experiment 3) using a longer SOA condition (visual stimulus 100 ms prior to tactile stimulus, unlike Experiment 1 and 2 in which the SOA was 30 ms) demonstrated more errors for the high proportion congruent

condition; still no proportion effect emerged on RTs. Shore and Simic (2005) concluded that the proportion effect was very limited.

In our study the paradigm was modified in order to conduct EEG recording. First, it was substantially simplified. The stimulation was limited only to left hand, making it easier for the participant to focus attention on touch. Second, the number of trials was considerably increased. Each block included 150 trials (120 trials for the high probability condition and 30 trials for the low probability condition), whereas in Shore and Simic's study, one block included only 64 trials (48 trials for the high probability condition and 16 trials for the

low probability condition). Thus, participants had longer blocks to practice distractor inhibition strategies. To confirm the latter assumption, we additionally tested whether the modulation of CCE as a function of proportion congruent was larger for the second 75 trials of the block compared to the first 75 trials of the block. The modulation of CCE as a function of proportion difference was calculated by subtracting the CCE of congruent 80% block from the CCE of congruent 20% block. T-tests revealed a significant increase of the proportion difference of CCE on RTs (first half: 42.37 ms, second half: 82.32 ms, $t(11) = 2.43$, $p < .05$). A similar analysis on percent correct data showed no significant difference between trials from the first (17.59%) and the second half of the block (19.45%, $t(11) = 0.26$, n.s.). Nonetheless, the numerical trend indicates that the differential CCE was larger for the second compared to the first half.

In addition to supporting our account of the discrepancy between the present study and the one conducted by Shore and Simic (2005), this additional analysis highlights that any inhibition strategy likely evolved during the block, and it is compatible with the hypothesis of increased distractor inhibition in the 20% congruent blocks. This inhibition could emerge as more effective suppression of the response evoked by the visual distractors, similar to what has been observed in other behavioral paradigms such as the Simon effect. However, it can also emerge as inhibition of the actual visual stimulation, which could be down-weighted in the multisensory integration. Finally, a combination of response and stimulus inhibition can also be hypothesized. The present study cannot disentangle between these options. However, the modulation of gamma band response discussed in the next paragraph seems to provide initial support to inhibition occurring at the stimulation rather than response level. Specifically, the modulation of gamma band activity as a function of expectancy in parietal and frontal areas may suggest that endogenous selection of the stimulation might have been involved (as suggested in the neural-functional model of attentional selection proposed by Corbetta and Shulman, 2002).

ICA clustering of EEG components at each frequency band

The EEG results in our study also reveal proportion effects on the cross sensory integration process. The proportion effects observed in the behavioral results were reflected in the gamma band response around 280 ms post-stimulus. While some studies reported that gamma band activity could stem from the ocular activity, the independent component analysis in our study showed that gamma band activity clearly originated from cortical sources. Specifically, one is the right superior parietal lobule (cluster 9) and the other is the right frontal superior frontal gyrus (cluster 11). It should be emphasized that gamma band activity in both these cortical sites was modulated as a function of congruency and proportion congruent, albeit in a different manner. The activity in cluster 9 was more closely related to the behavioral result, whereas the activity in cluster 11 was modulated by proportion congruent but less corresponded to the behavioral result. This confirms the independence of the parietal gamma activity from the activity of the different source, and supports the assumption that the parietal gamma band activity reflects the neural activity related to the cross sensory integration process.

Many studies have related the gamma band activity to the multisensory integration process (Kaiser et al., 2006; Kanayama et al., 2007, 2009; Sakowitz et al., 2001). In our study, it was revealed that the gamma band response in the parietal site, which was previously limited to the congruent condition (Kanayama et al., 2007; 2009), was also observed in the incongruent condition. In the congruent 20% block, we found greater gamma band activity in the incongruent condition compared to the congruent one. This suggests that, in the block in which the incongruent condition was more frequently presented, the gamma band activity was reallocated to the most frequent spatial arrangement. From the behavioral data, the decrement of CCE

in the 20% congruent block consisted of both worse performance in the congruent condition and better performance in the incongruent condition compared to 80% congruent block. This is the first time we see a complete reversal of the gamma band response to congruent and incongruent trials. In Kanayama et al. (2009), for instance, no such reversal emerged in behavioral or EEG data, regardless of whether rubber hands were present or absent. This EEG result parallels closely to the behavioral result. Given that it occurred in a region of the brain (the right superior parietal lobule) that is important for top-down control and association process of multisensory stimuli, greater gamma band activity may be considered a response to the highly expected multisensory combination. Further, a study using a symbol-to-sound matching paradigm demonstrated that the expected congruent sounds elicited a gamma band power increase compared to the unexpected sounds (Widmann et al., 2007). This result suggests that the gamma band response could reflect the matching between the expectation and the current sensory input, including both visual and auditory aspects of the stimulation. In our study, we expand the result of Widmann et al. (2007) in two ways: 1) we show that this is most evident in the cluster localized in the right superior parietal lobule; and 2) we show it for vision and touch, instead of vision and audition.

Regarding the theta band responses, we observed an effect of proportion congruent only in the occipital cortex (cluster16) and for the early theta band. The significant decrease of early theta band was observed only in the 80% congruent block. The result shows that the early activity in the visual cortex was induced when the stimulation was incongruent (regardless of expectancy), or when it was congruent but unexpected. In this respect, it may reflect some extended processing of the visual stimulation occurring within the occipital cortex whenever a conflict occurred (either in the stimulation or in the relationship between the stimulation and expectancy). Late theta ERSP modulation was also observed at the occipital cortex (cluster 16). The increase of late theta ERSP was found at the incongruent condition regardless of the proportion effect. Throughout the results, the late theta band activity did not show the proportion effect, which suggests that this component is immune to the top-down modulation. This indicates that the visual sensory area needs more activation during the presentation of distracting visual stimulus in the late time window, and the additional process is purely stimulus-driven. In line with the previous study (Kanayama and Ohira, 2009), late theta ERSP reflects the refined comparison process between multisensory stimuli for tactile localization.

For the ITC indices, we could not observe the clear proportion effect throughout all clusters. General ITC increase was observed when the stimulus was presented. In our study, ITC modulation by experimental condition was not observed at the contralateral central site (cluster 6), which suggests that neural activities in the primary somatosensory area elicited by the sensory input were equated to each other. Meanwhile, the ITC increase by bimodal stimulation in the early theta component was observed both in the frontal site (cluster 11) and parietal site (cluster 9) during 200–300 ms after stimulus presentation. Given that the ITC increase reflects the beginning of the process (phase-resetting; Brandt, 1997), this result suggests that both frontal and parietal sites could start the one function on the multisensory integration process at the same timing. Given that fronto-parietal network has been considered as a reflection of the interaction between the different modalities (Salmi et al., 2009), ITC increase in the same time course in the bimodal condition compared to the unimodal condition could be explained by the cooperation for the integration process of the different sensory modalities. Also this modulation was observed regardless of the proportion effect, which suggests that this fronto-parietal network was required to the processing on plural stimuli, and not modulated by the top-down control.

Regarding the coupling between gamma band ERSP and theta band phase index, it appears that the theta phase plays a role on the

modulation of gamma band ERSP as a function of the proportion of congruent trials. In the Congruent 80% block, the gamma band activity in the congruent condition was significantly locked to the $\pi/2$ phase (see the bar plot at the bottom right). In contrast, in the Congruent 20% block, the gamma band activity in the congruent condition is significantly locked at the $-\pi/2$ phase (see the bar plot at the bottom left). Coupling between amplitude at the gamma band and phase at the low frequency band indicates the modulation of excitability of the gamma band response by the phase at the low frequency band. These findings suggest that the proportion effect was supported by the modulation of the cross-frequency coupling between the amplitude of the gamma band oscillation and the phase of the theta band oscillation.

Conclusion

In conclusion, we identified a clear proportion effect on the multisensory integration process. The multisensory integration process was modulated by the expectancy for the next stimulus pattern as a top-down control process. Theta band oscillation reflected the cortical processing invoked by the sensory property (for example, spatial arrangement, temporal order, stimulus intensity), because the activity in many cortical areas was differentiated by condition (congruent, incongruent, somatosensory) but not by proportion. The cortical network of theta band oscillation represents the characteristics of the presented stimuli in one trial, so it is immune to the proportion of the condition in one block. Congruency of the presented multisensory stimuli was represented by the gamma band ERSP increase, but we postulate that this system is adaptive to the circumstance by modulating the coupling with the theta band oscillation. Given that the gamma band ERSP was modulated by the proportion effect, the default stimulus pattern in the task might be altered when the spatially incongruent pattern of the multisensory stimuli was frequently presented in a block. Developmental studies have demonstrated that the contingency of the visual and somatosensory sensation is likely developmentally acquired (Amsterdam, 1972; Miyazaki and Hiraki, 2006). It is likely that the multisensory integration system in the sensory association area is changeable to the context, and actively tuned to the most frequent and predictable pattern. This assumption could be supported by the cross frequency coupling between theta (low frequency band) phase and gamma (high frequency band) power, which is considered as a sensory selection system (Schroeder & Lakatos, 2009). Altogether, we could postulate that top-down control on the sensory processes could be reflected by the theta gamma coupling in the parietal area.

Funding

Japan Society for the Promotion of Science for Young Scientists (18-06718 to N.K.)

Acknowledgments

We thank Mr. Massimo Vescovi for the cooperation for the experimental devices and settings, and also the anonymous reviewers. This study is supported by the Japan Society for the Promotion of Science for Young Scientists, 18-6718, and by a PRIN grant to FP. We also wish to acknowledge the support of the Provincia autonoma di Trento and the Fondazione Cassa di Risparmio di Trento e Rovereto.

References

Amsterdam, B., 1972. Mirror self-image reactions before age two. *Dev. Psychobiol.* 5, 297–305.
 Andersen, T.S., Tiippana, K., Sams, M., 2004. Factors influencing audiovisual fusion and fusion illusions. *Brain Res. Cogn. Brain Res.* 21, 301–308.

Aspell, J.E., Lenggenhager, B., Blanke, O., 2009. Keeping in touch with one's self: multisensory mechanisms of self-consciousness. *PLoS One* 4, e6488.
 Austen, E., Soto-Faraco, S., Enns, J., Kingstone, A., 2004. Mislocalizations of touch to a fake hand. *Cogn. Affect. Behav. Neurosci.* 4, 170–181.
 Balslev, D., Nielsen, F.A., Paulson, O.B., Law, I., 2005. Right temporoparietal cortex activation during visuo-proprioceptive conflict. *Cereb. Cortex* 15, 166–169.
 Bhattacharya, J., et al., 2002. Sound-induced illusory flash perception: role of gamma band responses. *Neuroreport* 13, 1727–1730.
 Brandt, M., 1997. Visual and auditory evoked phase resetting of the alpha EEG. *Int. J. Psychophysiol.* 26, 285–298.
 Bremmer, F., Schlack, A., Shah, N.J., Zafiris, O., Kubischik, M., Hoffmann, K., Zilles, K., Fink, G.R., 2001. Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296.
 Brozzoli, C., Pavani, F., Urquizar, C., Cardinali, L., Farnè, A., 2009. Grasping actions remap peripersonal space. *NeuroReport* 20, 913–917.
 Calvert, G.A., Thesen, T., 2004. Multisensory integration: methodological approaches and emerging principles in the human brain. *J. Physiol. Paris* 98, 191–205.
 Calvert, G.A., Spence, C., Stein, B.E. (Eds.), 2004. *The Handbook of Multisensory Processes*. MIT Press, Cambridge, MA.
 Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., et al., 2006. High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313, 1626–1628.
 Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
 Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.
 Ghazanfar, A.A., Maier, J.X., Hoffman, K.L., Logothetis, N.K., 2005. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.* 25, 5004–5012.
 Heed, T., Habets, B., Sebanz, N., Knoblich, G., 2010. Others' actions reduce crossmodal integration in peripersonal space. *Curr. Biol.* 20, 1345–1349.
 Kaiser, J., et al., 2005. Hearing lips: gamma-band activity during audiovisual speech perception. *Cereb. Cortex* 15, 646–653.
 Kaiser, J., Hertrich, I., Ackermann, H., Lutzenberger, W., 2006. Gamma-band activity over early sensory areas predicts detection of changes in audiovisual speech stimuli. *NeuroImage* 30, 1376–1382.
 Kanayama, N., Ohira, H., 2009. Multisensory processing and neural oscillatory responses: separation of visuotactile congruency effect and corresponding electroencephalogram activities. *Neuroreport* 20, 289–293.
 Kanayama, N., Sato, A., Ohira, H., 2007. Crossmodal effect with rubber hand illusion and gamma-band activity. *Psychophysiology* 44, 392–402.
 Kanayama, N., Sato, A., Ohira, H., 2009. The role of gamma band oscillations and synchrony on rubber hand illusion and crossmodal integration. *Brain Cogn.* 69, 19–29.
 Kayser, C., Logothetis, N.K., 2007. Do early sensory cortices integrate cross-modal information? *Brain Struct. Funct.* 212, 121–132.
 Lakatos, P., Chen, C.M., O'Connell, M.N., Mills, A., Schroeder, C.E., 2007. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292.
 Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., Schroeder, C.E., 2008. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113.
 Macaluso, E., 2006. Multisensory processing in sensory-specific cortical areas. *Neuroscientist* 12, 327–338.
 Macaluso, E., Driver, J., 2001. Spatial attention and crossmodal interactions between vision and touch. *Neuropsychologia* 39, 1304–1316.
 Macaluso, E., Frith, C.D., Driver, J., 2001. Multimodal mechanisms of attention related to rates of spatial shifting in vision and touch. *Exp. Brain Res.* 137, 445–454.
 McGurk, H., MacDonald, J., 1976. Hearing lips and seeing voices. *Nature* 264, 746–748.
 Mishra, J., et al., 2007. Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *J. Neurosci.* 27, 4120–4131.
 Miyazaki, M., Hiraki, K., 2006. Delayed intermodal contingency affects young children's recognition of their current self. *Child Dev.* 77, 736–750.
 Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2002. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cogn. Brain Res.* 14, 115–128.
 Monto, S., Palva, S., Voipio, J., Palva, J.M., 2008. Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *J. Neurosci.* 28, 8268–8272.
 Pavani, F., Castiello, U., 2004. Binding personal and extrapersonal space through body shadows. *Nat. Neurosci.* 7, 14–16.
 Pavani, F., Spence, C., Driver, J., 2000. Visual capture of touch: out-of-the-body experiences with rubber gloves. *Psychol. Sci.* 11, 353–359.
 Sakowitz, O.W., Quiroga, R., Schürmann, M., Başar, E., 2001. Bisenory stimulation increases gamma-responses over multiple cortical regions. *Brain Res. Cogn. Brain Res.* 11, 267–279.
 Salmi, J., Rinne, T., Koistinen, S., Salonen, O., Alho, K., 2009. Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. *Brain Res.* 1286, 155–164.
 Schroeder, C.E., Lakatos, P., 2009. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18.
 Senkowski, D., Talsma, D., Herrmann, C.S., Woldorff, M.G., 2005. Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Exp. Brain Res.* 166, 411–426.
 Senkowski, D., Schneider, T.R., Foxe, J.J., Engel, A.K., 2008. Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci.* 31, 401–409.
 Shams, L., Ma, W.J., Beierholm, U., 2005. Sound-induced flash illusion as an optimal percept. *Neuroreport* 16, 1923–1927.
 Shore, D.J., Simic, N., 2005. Integration of visual and tactile stimuli: top-down influences require time. *Exp. Brain Res.* 166, 509–517.

- Spence, C., Pavani, F., Driver, J., 2004a. Spatial constraints on visual-tactile cross-distractor congruency effects. *Cogn. Affect. Behav. Neurosci.* 4, 148–169.
- Spence, C., Pavani, F., Maravita, A., Holmes, N., 2004b. Multisensory contributions to the 3-D representation of visuotactile peripersonal space in humans: evidence from the crossmodal congruency task. *J. Physiol. Paris* 98, 171–189.
- Widmann, A., Gruber, T., Kujala, T., Tervaniemi, M., Schröger, E., 2007. Binding symbols and sounds: evidence from event-related oscillatory gamma-band activity. *Cereb. Cortex* 17, 2696.
- Yuval-Greenberg, S., Deouell, L.Y., 2007. What you see is not (always) what you hear: induced gamma band responses reflect cross-modal interactions in familiar object recognition. *J. Neurosci.* 27, 1090–1096.
- Yuval-Greenberg, S., Tomer, O., Keren, A.S., Nelken, I., Deouell, L.Y., 2008. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* 58, 303–305.
- Zopf, R., Savage, G., Williams, M.A., 2010. Crossmodal congruency measures of lateral distance effects on the rubber hand illusion. *Neuropsychologia* 48, 713–725.